

ТЕМА № 11. Пространственные распределения.

Слайд 2. Методы анализа пространственных распределений точек: плотность.

Пространственное распределение - это расстановка, порядок, концентрация или рассеянность, соединенность или бессвязность многих объектов в пределах заключающего их географического пространства. Если мы сможем найти способы измерения этих отношений, то сможем найти пути выявления и понимания возможных механизмов, которые создают эти распределения.

Методы изучения распределений очень разнообразны. Одни из них - вычислительно просты, другие - очень сложны и требуют много машинного времени. Одни применяются очень часто, другие, хотя и чрезвычайно полезны, не получили широкого распространения.

Первое что мы с вами рассмотрим - это распределения точек.

Простейшей мерой точечного распределения является плотность точек. Она определяется как результат деления числа точек на общую площадь, на которой они расположены. Плотности населения, застройки, деревьев и т.д. широко используются как меры компактности точек. Сравнивая плотности подобных объектов в разных областях, мы можем сравнивать механизмы, которые действуют в этих областях. Или мы могли бы сравнивать точки в одном и том же месте, но в разные моменты времени, чтобы увидеть изменения плотности во времени.

Карты плотности точек обычно представляют собой растровые изображения. Значение в каждой ячейки растра является отношением числа точек в некоторой окрестности ячейки к площади этой окрестности. Форма и размер окрестности задается пользователем. Окрестность может быть представлена кругом, квадратом, прямоугольником или даже клином. Площадь окрестности вычисляется в единицах измерения входных данных. Выбирая небольшой размер окрестности, пользователь получает более разрешенную картину плотности. При увеличении размера окрестности происходит усреднение числа точек по площади изучения.

Слайд 3. Пространственные распределения точек.

Помимо общей плотности распределения, нас может интересовать еще и его форма. Точечные распределения встречаются в одном из четырех возможных вариантов. Распределение является **равномерным**, если число точек на единицу площади в каждой малой подобласти такое же, как и в любой другой подобласти. Если точки расположены в

узлах сетки, разделенные одинаковыми интервалами по всей области, то равномерное распределение называется **регулярным**. В других случаях равномерно распределенные точки располагаются в **случайном** порядке по всей рассматриваемой области.

Бывают случаи, когда точки собраны в тесные группы, такое распределение называется **сгруппированным или кластерным**.

Слайд 4. Методы анализа пространственных распределений точек: анализ квадратов.

Равномерные точечные распределения определяются на основе отношений между одинаковыми подобластями, называемыми квадратами. Это очень распространенный метод анализа дискретных зоологических и агрономических данных. Точками здесь могут быть отдельные растения, муравейники и т.д. Если каждый квадрат содержит примерно одинаковое число точек, то распределение является равномерным. Равномерные распределения редко встречаются среди биологических явлений, так как живым организмам свойственно мигрировать в сторону большей концентрации питательных веществ, лучшего орошения, определенного типа почвы и т.д. Если распределение действительно равномерное, то мы можем предположить, что нет существенного механизма, управляющего расположением объектов.

В стандартном методе анализа квадратов используется нулевая гипотеза о равномерности распределения: мы предполагаем, что примерно одно и то же число объектов будет находиться в каждой подобласти, равное общему числу объектов, поделённому на количество подобластей. Для проверки равномерности распределения может использоваться относительно простой статистический показатель, который называется критерием χ^2 (хи-квадрат) и выражается формулой:

$$\chi^2 = \sum [(Q-E)/E],$$

где Q - наблюдаемое число точек в квадрате, E – ожидаемое (т.е. среднее) число точек в квадрате, суммирование проводится по всем квадратам. Результат этого вычисления может быть сравнен с табулированными критическими величинами. Если полученное число незначительно отличается от ожидаемого, то распределение является равномерным; заметное отличие говорит о некоторой неравномерности, что может означать, наличие какого-то процесса, формирующего эту неравномерность. И пользователь может попытаться построить гипотезу, отвечающую на вопрос: «Если распределение не равномерно, то какой механизм может быть ответственен за это?»

Чаще всего наблюдаемые нами точечные распределения связаны с другими показателями (покрытиями) карты той же области исследования. Эти возможно связанные покрытия могут быть и точечными, и площадными, и линейными. В нашем примере с биологией это могли бы быть параметры почв. Это приведет нас к сравнению точек одного покрытия с объектами другого. Подобные сравнения называются операциями наложения и будут рассмотрены нами позже.

Помимо информации о равномерном распределении анализ квадратов может дать кое-что ещё. Например, отношение дисперсии к среднему (математическому ожиданию). Здесь также используется критерий χ^2 , который вычисляется как произведение отношения дисперсии к среднему на число подобластей за вычетом одной. Высокие значения χ^2 указывают на большой разброс между числом точек в каждой области и средним для всей области, то есть на то, что мы имеем кластерное (групповое) распределение. И наоборот, малые значения χ^2 означают, что распределение более равномерное. Промежуточные значения указывают на то, что распределение более тесно связано с некоторым случайным процессом, где некоторые квадраты имеют несколько большее, а другие - несколько меньшее число, чем среднее

$$\chi^2 = (n-1) d^2 / \bar{x},$$

d^2 – дисперсия, \bar{x} – среднее, n – число квадратов.

В любом случае, если результаты анализа говорят, что если распределение не является статистически случайным (т.е. если оно либо равномерное, либо кластерное), то вы можете попытаться определить возможную причину, разумно выбрав набор показателей для сравнения с точечным покрытием.

Слайд 5. Методы анализа пространственных распределений точек: анализ ближайшего соседа.

До сих пор мы описывали точечные распределения количеством точек в пределах подобластей. Другими словами, мы рассматривали распределение точек посредством сравнения областей, которые они занимают. Однако, также поучительно рассмотреть локальные отношения внутри пар точек. Чаще всего это делается другим методом анализа точечных распределений - **анализом ближайшего соседа**, общепринятой процедурой определения расстояния от каждой точки до ее ближайшего соседа (РБС) и сравнения этой величины со средним расстоянием между соседями. Среднее РБС дает меру разреженности точек в распределении. Это ценно само по себе, так как в некоторых случаях точечные объекты могут конфликтовать, если они расположены слишком близко друг к другу. Например, мы знаем, что многим животным

требуется определенное жизненное пространство, и когда оно перекрывается с пространством другого представителя того же вида, возможен конфликт.

Как и в анализе квадратов, мы можем сравнить среднее РБС с тремя возможными распределениями — регулярным, случайным и кластерным. Этот метод может быть описан в общем для каждого из этих случаев как вычисление индекса, с которым вы можете сравнить свои результаты:

Индекс случайного распределения - $1/[2 \times (n/S)^{1/2}]$,

Индекс максимальной рассеянности – $1.07453/(n/S)^{1/2}$,

n – число точек, S - площадь, n/S – плотность точек.

Наконец, для **критерия максимальной сгруппированности** мы можем принять **РБС = 0**, когда точки расположены одна под другой.

Простое сравнение рассчитанного РБС с тремя индексами даст вам понятие о том, каково это распределение.

Давайте рассмотрим, как это работает на примере. Допустим, у нас есть шесть точек, расположенных в пределах площади в 25 квадратных единиц. Ближайшие соседи этих точек и расстояния до них записаны в таблице справа.

Среднее РБС этих данных составляет примерно 1.4. Для случайным образом распределенных данных индекс составит - 1.02. Наше среднее РБС несколько больше, чем этот индекс.

Критерий максимальной рассеянности точек составляет ≈ 2.19 . Такого было бы значение, если бы наше распределение точек было идеально равномерным. Наше среднее РБС намного меньше этого, но и намного больше, чем 0, который соответствует идеально сгруппированному распределению. Таким образом, мы нашли, что наше распределение где-то между истинно равномерным и случайным. Другими словами, оно начинает принимать более регулярную конфигурацию, но пока все еще довольно случайное.

Слайд 6. Полигоны Тиссена.

Точечные распределения могут также характеризоваться с помощью полигонов Тиссена (называемых также диаграммами Дирихле и диаграммами Вороного). Они основаны на идее, что мы можем нарастить полигоны вокруг точек, чтобы показать их возможные зоны влияния.

Создание полигонов Тиссена довольно просто концептуально, но может стать запутанным, если количество точек велико. Чтобы понять, как их строить, давайте вначале разберемся, что эти фигуры должны представлять. Если у нас есть несколько точечных объектов, мы можем представить себе, что каждая точка окружена одиночным неправильным многоугольником. Но многоугольник имеет одно важное свойство - любая точка внутри него находится ближе к очерченной точке, чем любая другая точка покрытия. И наоборот, каждая точка вне полигона ближе к некоторой иной, нежели к очерченной. Другими словами, граница каждого полигона дает окружающей точке наименьшую возможную область влияния. Каждая точка покрытия будет иметь свой собственный полигон Тиссена, показывающий область исключительно ее влияния. Теперь давайте подумаем, как мы могли бы сделать это.

Возьмем простой набор точек (рисунок в верхней части слайда). Образование полигонов Тиссена можно представить как результат роста мыльных пузырей с центром в каждой из точек. В конце концов границы пузырей превращаются в прямые линии, а сами пузыри - в многоугольники. Стороны этих многоугольников ориентированы перпендикулярно линиям, соединяющим соседние точки. Причем длины двух отрезков, получившихся с обеих сторон границы одинаковы.

Зачем же нужны полигоны Тиссена? Они названы в честь климатолога Тиссена (А.Н. Thiessen), который пытался проинтерполировать сильно неравномерные распределения климатических данных. Иначе говоря, он пытался описывать и анализировать точечные данные с помощью площадных символов и аналитических методов. Таким образом, если у нас есть несколько разбросанных точек, и мы хотим охарактеризовать регионы, основанные на этих точках, то используем полигоны Тиссена. Поскольку мы считаем, что в каждом полигоне влияние очерченной точки абсолютно мы можем обращаться с этими данными как с полигональным покрытием.

Слайд 7. Распределения полигонов.

Мы можем начать анализировать распределения областей во многом подобно тому, как мы делали это с точками - через определение плотности полигонов на единицу площади нашей области изучения. Опять же, помимо плотности полигонов, нас может интересовать расположение и формы распределений, создаваемые группами полигонов, которые могут подсказать причины таких расположений. Кроме этого,

площадные объекты могут быть соединены друг с другом, или удалены на некоторое определенное расстояние.

Статистик соединений. Вид распределения, образуемого полигонами, можно также оценивать с помощью статистического показателя (статистика) соединений (общих границ).

Соединение - это общая граница двух смежных полигонов. *Статистик соединений* подсчитывает количество соединений в полигональном распределении и характеризует структуру соединений каждого покрытия.

Рассмотрим простейшее полигональное покрытие, в котором есть всего 2 вида полигонов- темные и прозрачные. Это могут быть зоны растрового покрытия или полигоны с двумя различными значениями атрибута векторного покрытия.

Всего между полигонами имеются 22 соединения (т.е. общих участков границ). На слайде показаны три возможных варианта взаимного расположения полигонов. Первый случай: $m/m = 8$, $n/n = 10$, $m/n = 4$ - 8 соединений между темными полигонами, 10 - между прозрачными и 4 - между темными и прозрачными. Эти числа показывают, что между темными и прозрачными полигонами имеется мало соединений, большинство прозрачных полигонов соединены друг с другом, и большинство темных полигонов соединены друг с другом. Другими словами, полигоны сгруппированы (кластерное распределение), подобно тому, что мы прежде наблюдали с точками. Следующий рисунок (правее) показывает совершенно другой набор чисел: $m/m = 0$, $n/n = 3$, $m/n = 19$ - здесь большинство соединений (19 из 22) - между полигонами разных классов, т.е. мы имеем разреженное распределение. На последнем рисунке - промежуточный случай: $m/m = 3$, $n/n = 4$, $m/n = 15$ - оба числа соединений однородных полигонов низки, но не так, как на предыдущем рисунке. Таким образом, здесь мы имеем дело, скорее всего, со случайным распределением.

Также как и в случае с точечными объектами случайность говорит о том, что расположение полигонов, скорее всего не зависит от какой-либо причины. И, наоборот - в двух других случаях такая причина наверняка существует.

Распределение соединений также можно оценивать с помощью критерия χ^2 . Но этот показатель подразумевает, что мы знаем, каким должно быть ожидаемое распределение в условиях случайности. Если бы мы знали подобные распределения для полигонов (на основе числа соединений), то могли бы сравнивать их точно таким же способом.

Но как мы узнаем ожидаемое случайное распределение соединений, с которым могли бы сравнить имеющиеся значения? Имеются разные подходы к решению этой задачи. Один из подходов, называемый свободным отбором, предполагает, что мы можем

определить ожидаемую частоту соединений либо на основе теоретического знания моделируемой ситуации, либо исходя из известных распределений для других областей исследования.

Слайд 8. Измерение пространственных распределений в среде ArcGIS.

В последние версии ArcGIS были добавлены несколько наборов инструментов, позволяющих анализировать пространственные распределения разных типов объектов. Инструменты объединены в общий набор **Spatial Statistics Tools**. **Spatial Statistics Tools** включает Инструмент построения центроидов для точечных распределений (**Central Feature**), Анализ ближайшего соседа для точечных распределений (**Average Nearest Neighbor**), Инструменты измерений распределений полигонов (**High/Low Clustering** и др.). ArcGIS также позволяет рассчитывать полигоны Тиссена, но пока только для одного формата данных (покрытий ARC/INFO). Последнее не является препятствием, так как ArcGIS легко проводит операции экспорта / импорта данных.

Слайд 9. Методы анализа пространственных распределения линий: плотность линий.

Поскольку линии в отличие от точек имеют пространственную протяженность, анализ их распределений несколько сложнее. Одни исследователи изучали распределения длин линий, другие рассматривали интервалы между линии, во многом подобно анализу ближайшего соседа в точечных распределениях.

Мы рассмотрим эти и другие меры распределений линий и начнем с простейшей меры — плотности линий.

Мы определили плотность безразмерных точек как отношения их числа к занимаемой ими площади. Плотность двумерных полигонов определялась как отношение суммарной площади класса к площади всей карты. Подобным же образом, для определения плотности одномерных линий мы будем использовать отношение суммы их длин к площади покрытия. Выразаться оно может в метрах на гектар или километрах на квадратный километр.

Исследование распределений линейных объектов является полезной процедурой при региональных геологических исследованиях. Предметом изучения геологов являются такие природные линейные феномены, как линеаменты. Очень часто для оценки геологической и тектонической ситуации области исследования

используют такой показатель, как плотность линеаментов. Косвенным образом плотность линеаментов характеризует макротрещиноватость осадочного чехла и степень неотектонической активности.

Расчет плотности линий (в том числе и линеаментов) аналогичен расчету плотности точек. Как правило плотность линейных объектов— это растровый набор данных. Значение в каждой ячейки растра является отношением длины линейных объектов пределах в некоторой окрестности ячейки к площади этой окрестности. Окрестность обычно представляется кругом, радиус круга задается пользователем. Программа выбирает все линии, которые попадают в очерченный круг (если линия выходит за пределы радиуса расчета, то выбирается только та часть линии которая лежит внутри круга) затем суммирует их длины и делит на площадь круга. Выбирая небольшой размер окрестности, пользователь получает более разрешенную картину плотности. При увеличении размера окрестности происходит усреднение плотности линий по площади изучения.

Слайд 10. Методы анализа пространственных распределения линий (продолжение)

Другой способ изучения распределений линейных объектов основан на расчете интервалов между линиями и во многом подобен анализу ближайшего соседа в точечных распределениях. Распределение пар линий может быть определено во многом подобно тому, как мы поступали с точками, хотя вычисления несколько усложняются, так как, в отличие от точек, линии имеют размерность. Может показаться, что просто выбрать центр каждой линии и провести анализ ближайшего соседа для этих точек. Однако, вследствие того, что линии имеют различные длины, эта процедура не даст нам правдивой картины распределения самих линий. С точки зрения статистики часто считается полезным делать случайную выборку. Существует несколько подходов к получению случайной выборки.

1. Для анализа ближайших соседей среди линейных объектов на каждой линии выбираем случайную точку (или на каждом сегменте линии, если они - не прямые). Далее, опускается перпендикуляр из этой точки к ближайшей линии (на слайде – рисунок слева). Затем измеряем эти расстояния и подсчитываем среднее РБС. Как со всеми РБС, мы должны иметь возможность оценить эту величину по отношению к случайному. Как и в случае точек используем статистический критерий. Этот критерий работает для большинства распределений линий, будь линии прямыми или изогнутыми, но имеет и некоторые ограничения. Если линии очень извилисты, этот подход - менее чем успешен.

Кроме того, чтобы критерий был полезен, линии должны быть, по меньшей мере, в полтора раза длиннее среднего расстояния между ними.

2. Методы пересечения линий являются альтернативой при анализе распределения линий. Один простой подход состоит в том, чтобы преобразовать двухмерный паттерн в одномерную последовательность прочерчиванием выборочной линии через карту и учетом пересечений этой линии с линиями покрытия. Существуют, по меньшей мере, два способа создания таких линий. Первый - случайно выбрать пару точек и соединить их линией. Второй метод состоит в проведении луча из случайной точки под случайным углом. После того, как линия проведена, может быть рассмотрено распределение интервалов между пересечениями ее с линиями покрытия с использованием стандартных статистических методов анализа данных.

Вместо одиночной можно провести зигзагообразную линию, которая пересекает покрытие два или три раза. Зигзагообразный путь, часто называемый случайным обходом, также создаст серию пересечений, расстояния между которыми могут быть проанализированы статистическими методами.

Слайд 11. Направленность линейных и площадных объектов.

Линейные и площадные объекты могут характеризоваться не только распределением по ландшафту, но и ориентацией. Такие объекты как осадочные напластования, русла ледников, переносимая водой галька, цепи валунов, оставленные ледниками, ограждения, сети улиц, ветровал деревьев в лесу имеют определенную ориентацию, которая часто указывает на породившую их силу.

Но когда мы анализируем ориентацию, у нас может возникнуть ситуация выбора между двумя встречными направлениями. Если линейный объект является улицей с односторонним движением, то ориентация ее самой не говорит нам о направлении, в котором должен двигаться транспорт. Поэтому, кроме ориентации нам нужно знать и о направленности.

В традиционном статистическом анализе ориентации линий используют диаграмму направлений (роза направлений). Розы направлений полезны для визуальной оценки, но для численного анализа более полезны измерения, получаемые непосредственно из самого линейного покрытия. Численным выражением направленности линейных объектов является **равнодействующий вектор**. Поясним смысл и способ расчета равнодействующего вектора на простом примере.

Для демонстрации анализа направлений возьмём большое количество деревьев, поваленных прямолинейным ветром. Каждое дерево может быть отображено как линейный объект покрытия, при этом записываются координаты вершины и основания каждого дерева, давая нам ориентацию каждого дерева. Метеорологи хотят выяснить общее направление ветра по поваленным деревьям, но эти деревья не имеют единой для всех ориентации, поэтому нашей первой задачей является определение равнодействующего вектора поваленных деревьев.

Равнодействующий вектор является результатом геометрической суммы всех векторов. Рисунок внизу показывает равнодействующий вектор R , полученный из трех векторов A , B и C .

Аналогично, можно рассчитать и равнодействующий вектор для поваленных деревьев. Здесь важен один момент: длина дерева вносит существенный вклад в результирующий вектор. Чтобы каждое дерево вносило одинаковый вклад следует, приравнять его длину единице. Если векторы не нормализовать на их длину, тогда длину дерева следует рассматривать как весовой коэффициент для его направления. Для данного примера более длинные деревья, скорее всего, будут представлять среднее направление.

Слайд 12. Направленность линейных объектов (продолжение).

Рисунок в верхней части слайда показывает два случая равнодействующего вектора R с тремя исходными. Когда векторы расположены близко к одному направлению, равнодействующий вектор будет длинным, в то время как при широко разбросанных исходных векторах - значительно более коротким.

Таким образом, мы имеем не только среднее направление лесовала, но и меру компактности распределения: чем компактнее распределение, тем длиннее эта линия.

Для сравнения длины равнодействующего вектора в данном месте с другим местом, нам следует, опять же, нормализовать данные. **Нормализованная длина равнодействующего вектора получается делением длины равнодействующего вектора R на сумму длин образующих его векторов.**

Это безразмерная величина в диапазоне от 0 до 1, напоминающая дисперсию в традиционной статистике, так как является мерой пространственного разброса вокруг среднего значения. Правда, она выражает этот разброс "наоборот": большие значения соответствуют более близкой ориентации векторов, меньшие - большему

разбросу. Таким образом, большое значение этого показателя в нашем примере с деревьями означало бы значительное преобладание одного из направлений ветра, а малое значение говорило бы о наличии завихрений или отсутствии явно преобладающего направления.

Остается одна проблема, связанная с ориентацией и направлением. Как мы знаем, одни линейные объекты могут иметь определенное направление (реки), другие же - нет (лесозащитные полосы), хотя определенная ориентация присутствует в обоих случаях. Например, один исследователь, собирающий данные о направлении трещин пластов горных пород, может указывать, что некоторые из них ориентированы на север, а другой - что на юг. Тогда, при анализе данных, собранных этими исследователями, может оказаться, что при определении длины среднего равнодействующего вектора, исходные векторы будут, так сказать, взаимно уничтожать друг друга.

К счастью, для этого есть остроумное решение. Крумбейн обнаружил, что при удвоении значения угла, независимо от исходного направления, записывается одно и то же значение. Допустим, мы имеем объект, ориентированный с северо-запада (315°) на юго-восток (135°). После удвоения мы получим: $315^\circ \times 2 = 630^\circ$ ($630^\circ - 360^\circ = 270^\circ$) и $135^\circ \times 2 = 270^\circ$. Такой способ выражения направлений повлияет на формулы для вычисления среднего направления, нормализованной длины равнодействующего вектора, поэтому, чтобы получить действительные значения, их нужно будет модифицировать.

В заключении следует отметить, что все эти простые меры направленности и разброса могут быть проверены на случайность и наличие тренда стандартными процедурами проверки статистических гипотез.

Слайд 13. Измерение плотности и направлений в среде ArcGIS.

В ArcGIS также имеются инструменты для измерений плотности и направлений объектов. Построение карт плотности точечных и линейных объектов реализуется через набор инструментов модуля **Spatial Analyst**. Расчет равнодействующего вектора выполняет инструмент **Linear Direction Mean** из набора инструментов **Spatial Statistics Tools**.

Кроме этого ArcGIS имеет мощный аппарат для создания, редактирования и анализа сетей, о которых будет идти речь далее.

Слайд 14. Связность линейных объектов: гамма- индекс.

Важным аспектом пространственного расположения линий является их способность образовывать сети. Сети имеют самые разнообразные формы как естественные, так и созданные человеком. Среди них: автомобильные и железные дороги, телефонные линии, реки, даже

лесозащитные полосы могут играть роль сети, позволяющей мелким млекопитающим перемещаться по ландшафту, - список можно долго продолжать. И хотя мы можем интересоваться плотностью и ориентацией объектов, образующих сеть, нам нужна также и возможность анализировать реальные связи, образованные этими объектами и степень связанности между различными точками сети. Вы наверняка сталкивались с ситуациями, когда в городе нет прямой дороги от места до места, и приходится ехать круглым путем. Здесь мы сталкиваемся с недостатком **связности** в сети.

Связность является мерой сложности сети. Имеются несколько методов для определения этой характеристики. Наиболее общими являются **гамма-индекс** и **альфа-индекс**.

Гамма-индекс g является отношением числа существующих связей между парами узлов сети, L , к максимально возможному числу связей в том же наборе узлов, L_{\max} . Очевидно, что векторно-топологическая модель данных лучше всего подходит для этих вычислений. Определить же L_{\max} не так трудно, как может поначалу показаться, оно однозначно определяется числом узлов V . Например, если мы имеем три узла, то возможны лишь три связи. Если мы добавим еще один узел, то сможем добавить еще три связи, а всего их будет шесть. И если мы полагаем, что не образуются новые пересечения, то максимальное число связей будет каждый раз увеличиваться на три. То есть, $L_{\max} = 3(V - 2)$.

Гамма-индекс тогда определяется как

$$g = L / L_{\max} = L / 3(V - 2)$$

Он принимает значения от 0 (нет ни одной связи) до 1 (все возможные связи присутствуют).

Рисунок на слайде показывает два варианта сети с 16-ю узлами. На левом рисунке имеется 15 связей, что дает связность $g = 15 / 3(16-2) = 0.36$. На рисунке справа имеется 20 связей, поэтому $g = 20 / 3(16-2) = 0.48$. Образно говоря, первая сеть связана примерно на треть, а вторая - наполовину.

Большее количество связей в сети облегчает передвижение по ней, что важно, например, для специалистов по транспортному планированию.

Слайд 15. Связность линейных объектов: альфа-индекс.

Важной характеристикой сетей помимо связности является наличие в ней контуров, позволяющих перемещаться от узла к узлу разными маршрутами. В качестве примера можно привести кольцевые автодороги вокруг крупных городов, позволяющие снизить нагрузку транзитного транспорта на уличную сеть.

В качестве меры соединенности узлов контурами альтернативных маршрутов используется так называемый **альфа-индекс (α)**. Он является отношением имеющегося в сети числа контуров к максимально возможному числу контуров в этой сети. Известно, что сеть без контуров имеет связей на одну меньше, чем число узлов: $L = V - 1$. На рисунке слева вы видите такую сеть, - в ней 16 узлов и 15 связей. Она обладает минимальной связностью, в том смысле, что в ней имеется наименьшее возможное число связей при заданном числе узлов, причем каждый узел имеет по меньшей мере одну связь.

Добавление какой-либо связи создает контур, т.е. когда сеть содержит контуры $L > V - 1$. Число же имеющихся контуров можно определить как $L - (V - 1)$.

Далее, так как максимальное число связей в сети определяется как $3(V - 2)$, то максимальное число контуров будет $L_{\max} - (V - 1) = 3(V - 2) - (V - 1) = 2V - 5$. Отсюда альфа-индекс $\alpha = (L - (V - 1)) / (2V - 5)$. Диапазон значений альфа-индекса - от 0 (сеть без контуров) до 1 (сеть с максимальным числом контуров).

Теперь мы можем вычислить альфа-индекс для сетей представленных на слайде

$$\alpha_1 = (15 - 16 + 1) / (2 \cdot 16 - 5) = 0$$

$$\alpha_2 = (20 - 16 + 1) / (2 \cdot 16 - 5) = 0.19$$

Таким образом, в первой сети есть только один вариант для перемещения из одной точки в другую, для другой сети возможны несколько альтернативных маршрутов разной длины.

Слайд 16. Сетевой анализ: геометрические сети.

Мы с вами только, что рассмотрели простейшие меры, характеризующие сети. Анализ сетей этим не ограничивается и подразумевает более сложные процедуры и расчеты, имеющие важное значение для решения многих практических задач, связанных с сетями. Прежде, чем обратиться к задачам сетевого анализа, рассмотрим основные элементы и характеристики геометрических сетей.

Сети состоят из двух основных компонентов: ребер и соединений. Ребра и соединения в сети являются топологически связанными друг с другом. *Ребро* – это тип элемента сети, который имеет определенную длину, и через который движется поток определенного предмета потребления. Линии электропередачи, трубопроводы и русла

потоков могут служить примерами ребер. *Соединение* имеет место на пересечении двух и более ребер и обеспечивает передачу потока между двумя ребрами. Примерами могут служить предохранители, трансформаторы, служебные втулки и точки слияния потоков.

Связность сети основана на совпадении элементов геометрии объектов: координаты точек, представляющих соединения, совпадают с координатами концевых точек ребер. Совпадающая геометрия – основное свойство таких сетей, отсюда и название – геометрические сети.

Направление потока. Сетевые структуры данных часто применяются для моделирования систем реального мира, в которых определены направления перемещения по сети. Например, перемещение потока электроэнергии в электрических сетях происходит в направлении от электростанции и генераторов к потребителям. В водопроводных сетях направление потока может быть не так хорошо определено, как в электрических сетях, тем не менее, поток воды может распространяться от насосных станций к потребителям, или от потребителей к очистным сооружениям.

Направление потока в сети определяется связностью сети, расположением источников и приемников сети, а также включенным или отключенным состоянием объектов.

Источники и приемники проводят поток через сеть. Источники – это объекты-соединения, из которых поток расходится по ребрам сети. Например, в сети распределения воды насосные станции могут рассматриваться как источники, так как именно они проталкивают воду по трубам в разные стороны от насосных станций. Приемники – это объекты-соединения, которые собирают в себя поток из ребер сети. Например, в речной сети устье реки можно рассматривать как приемник, так как сила гравитации направляет всю воду именно к нему. Поток всегда движется по направлению от источников к приемникам. Так как направление потока может быть установлено только с помощью приемников или только с помощью источников, обычно в сети определяют только одни из этих объектов.

Важно помнить, что при установлении направления потока учитывается **отключенность** объектов. Любой объект в геометрической сети может быть включен или отключен. Включенный объект позволяет потоку проходить через него, в то время как отключенный объект не позволяет. Следовательно, отключение объекта будет означать, что направление потока не может быть установлено не только для отключенных объектов, но и для всех объектов, которые связаны с источниками и приемниками только через отключенный объект.

Для сети может быть также указан набор **весовых коэффициентов**. Под весовым коэффициентом (весом) понимается величина стоимости перемещения по элементу сети. Например, в водопроводной сети происходит потеря некоторого давления при транспортировке по трубопроводу из-за трения воды о поверхность труб: атрибутом, на основе которого может быть вычислен вес, является длина трубы.

Теперь, когда мы определили основные понятия и характеристики геометрических сетей, рассмотрим задачи сетевого анализа. **Анализ сетей включает трассировку**. Термин “*трассировка*” используется здесь для описания построения набора элементов сети в соответствии с определенной процедурой. При работе с сетями результат трассировки определяется связанностью. Элемент может быть включен в результат трассировки, только если он как-либо связан с другими элементами результата трассировки. **Результат трассировки – это набор объектов сети, найденных операцией трассировки.**

Например, предположим, что вы хотите найти все объекты вверх по течению от определенной точки в речной сети. Используя прозрачный лист, наложенный на карту сети рек, вы можете нарисовать все участки реки, расположенные вверх по течению от заданной точки. То, что будет нарисовано на прозрачном листе после выполнения этой операции, и будет вашим результатом трассировки.

Аналогично, когда вы выполняете операцию трассировки в ArcMap, результатом будет набор элементов сети, включенных в трассу.

До начала трассировки вы должны определить **флаги и барьеры**.

Флаги определяют начальные точки трасс. Например, если вы выполняете трассировку вверх по течению, вы можете использовать флаг, чтобы определить, с какой точки должна начаться трасса вверх по течению. Флаги можно помещать в любую точку на ребрах и соединениях. Элементы сети, связанные с этими ребрами или соединениями, будут рассмотрены на предмет включения в результат трассировки.

Барьеры определяют точки в сети, за которые не может быть продолжена трасса. Если вам нужно провести трассировку только определенной части сети, вы можете использовать барьеры для изолирования этой части сети. При выполнении операции трассировки программа воспринимает объекты, на которых установлены барьеры, как отключенные, таким образом, трасса не может быть продолжена через эти объекты.

Например, в муниципальной сети водоснабжения магистральная труба была вскрыта и отключена в связи с реконструкцией улицы, вода не может проходить через эту секцию магистральную трубу. Отключение объекта сети, представляющего эту трубу, остановит трассировку на этом объекте.

Слайд 17. Функции сетевого анализа.

Существуют различные виды трассировки, реализуемые в ArcGIS через *Функции сетевого анализа*.

Функции сетевого анализа позволяют решать следующие задачи:

- определить направление потока

Слайд 18. Функции сетевого анализа: трассировка сетей

- найти все элементы, расположенные вниз по течению от заданной точки вашей сети (задача **Трассировка вниз по течению**).
- найти все элементы, расположенные вверх по течению от заданной точки вашей сети (задача **Трассировка вверх по течению**).
- найти все объекты сети, лежащие выше по течению от заданной точки, используйте (задача **Найти накопление вверх по течению**).

Слайд 19. Функции сетевого анализа (продолжение)

- найти путь от заданной точки в сети вверх к источнику (задача **Найти путь вверх по течению**).
- найти все общие объекты, находящиеся вверх по течению для заданного набора точек (задача **Найти общих предков**).
- найти путь между двумя заданными точками в сети (задача **Найти путь**). Этот путь может быть лишь одним из вариантов пути между этими двумя точками, если ваша сеть содержит петли.
- найти петли в сети (задача **Найти петли**). Петли могут влиять на наличие нескольких путей между точками в сети.

Тема №12. Поверхности.

Слайд 2. Непрерывные и дискретные поверхности.

Поверхности - это объекты, которые чаще всего представляются значениями Z , распределенными по области, определенной координатами X и Y . Параметр Z чаще всего ассоциируется с высотой рельефа местности, но не обязательно. На самом деле, любые измеримые величины, которые могут встречаться на определенной территории могут рассматриваться как образующие поверхность. Обычно используется термин «статистическая поверхность» (statistical surface), поскольку значения параметра Z часто можно трактовать как статистическое представление величины рассматриваемых явлений или объектов. Определение статистической поверхности как непрерывной означает, что имеется бесконечное число точек, в каждой из которых может быть свое значение. Однако, провести измерения в бесконечном числе точек физически невозможно, так же как невозможно хранить бесконечный объем данных. Поэтому определение непрерывной поверхности с помощью бесконечного числа точек должно быть заменено моделью, которая использует только некоторые точки (в наилучшем варианте, существенно важные отсчеты рассматриваемой величины). Говоря языком статистики, если выборка значений из всех возможных представительна, то модель будет адекватно отражать реальный объект или явление. В противном случае, мы можем получить модель очень далекую от реальности.

С другой стороны, сами значения имеют погрешность наблюдения или расчета, т.к. не существует средств и методик получения абсолютно точных измеряемых или вычисляемых величин. Этот немаловажный фактор также вынуждает нас рассматривать создаваемые поверхности как приближенные (статистические) модели окружающей нас действительности.

Выбор типа модели зависит, в общем случае, от двух факторов:

- вида пространственного распределения точек (X, Y, Z) – расположение точек может быть регулярным (упорядоченным) или нерегулярным (случайным или сгруппированным);
- и характера значений Z - измеренные величины могут быть непрерывными (если значение существует в любой точке пространства) либо дискретными (если значение существует только в определенных точках пространства). В последнем случае статистическая поверхность не будет являться непрерывной: в некоторых точках значение будет изменяться скачком. Говорят, что в таких точках поверхность претерпевает разрыв.

Наиболее распространенным способом представление поверхностей является способ расчета и представления поверхностей в виде изолиний.

Изолинии (они же изометрические линии) - линии равных значений непрерывной статистической величины. Существуют также изоплеты - линии равных значений дискретной статистической величины, они называются также псевдоизолиниями, так как между ними статистическая величина может изменяться скачком или быть не определена. Изоплетами обычно изображаются вычисляемые, а не измеряемые непосредственно величины.

На рисунке представлены поверхности в виде изолиний для двух различных картографируемых явлений: изменение высоты местности и изменение численности населения. Внешне те и другие выглядят одинаково, но суть их различна: в то время, как высота местности существует в каждой точке пространства (непрерывная величина), людям не свойственно селиться равномерно и непрерывно – они, как правило, размещаются локальными группами (поверхность с разрывами).

Слайд 3. Модели поверхностей.

Кроме изолиний в ГИС используются два типа моделей для представления поверхностей: модель GRID и модель TIN. Более верным будет высказывание, что основными моделями поверхностей как раз и являются гриды и TINы, а изолинии являются производными объектами от TIN или GRID.

GRIDы представляют поверхности по регулярно распределенным точкам. TINы представляют поверхности на основе нерегулярного пространственного распределения точек.

Слайд 4. TIN - нерегулярные сети треугольников.

Структура данных TIN позволяет вам точно воссоздать любой тип поверхности. В TIN можно не только интерполировать высоты для любого местоположения, но и хранить такие естественные перегибы в уклоне поверхности, как гребни и тальвеги.

Термин нерегулярная триангуляционная сеть (triangulated irregular network) точно описывает свойства TIN.

"Нерегулярная" определяет ключевое преимущество TINs в моделировании поверхности — точки могут быть взяты с переменной плотностью для моделирования поверхности: для описания ровных или плавно меняющихся участков поверхности достаточно небольшого числа точек, для описания резко изменяющего рельефа поверхности выбирается большее число точек.

"Триангуляционная" указывает на способ построения оптимизированного набора

треугольников по набору точек. Треугольники дают хорошее представление о локальной части поверхности, так как три точки со значениями z однозначно определяют плоскость в трехмерном пространстве.

TIN строятся по 3D точкам, то есть точкам, имеющим 3 координаты (X, Y, Z) . По этим исходным точкам выполняется триангуляция по набору точек. В TIN треугольники называют гранями (faces), точки становятся узлами (nodes) граней, а линии граней называют ребрами (edges). Каждая грань TIN является частью плоскости в трехмерном пространстве. Все грани в TIN точно смыкаются с соседними в каждом узле и вдоль каждой грани. Грани не могут пересекать друг друга.

"Сеть" отражает топологическую структуру, которая присуща TIN. Такая структура делает возможным сложный анализ поверхности и, кроме того, компактное представление поверхности.

Слайд 5. TIN: триангуляция Делоне.

Для заданного набора точек можно построить много вариантов разбиения на треугольники. В ArcGIS для оптимального моделирования поверхности треугольниками применяется алгоритм, известный как триангуляция Делоне (Delaunay triangulation). Основным принцип этого алгоритма заключается в том, чтобы построить треугольники, которые все вместе будут по возможности близкими к равносторонним фигурам. При такой интерполяции значения высот для новых точек будут более близки к известным исходным точкам.

Слайд 6. Топология в TIN.

TIN - топологическая структура данных, управляющая информацией об узлах, которые входят в каждый треугольник, и о соседях каждого треугольника.

В специальной таблице для каждого треугольника храниться список узлов, его образующих, и список смежных треугольников (треугольники всегда имеют 3 узла и обычно имеют три соседних треугольника, треугольники на внешней границе TIN могут иметь одного или двух соседей). Тем самым полностью описываются пространственные отношения элементов TIN, включающие информацию о связности и соседстве.

Интересное свойство представления данных TIN состоит в том, что точки отображаются в трехмерном пространстве, но топологическая сеть граней ограничивается

двумя измерениями.

По этой причине иногда говорят, что представление данных TIN является 2,5-мерной моделью пространства. Это описание не совсем «научно», но привлекает внимание к тому факту, что хотя поверхности существуют в трех измерениях, каждая точка может обладать только одним z-значением. Следовательно, TIN - пример однозначной функции, когда при заданном исходном местоположении может задаваться только одно z-значение. Некоторое ограничение TIN состоит в том, что они не могут моделировать достаточно редкие случаи с отрицательными уклонами, типа нависающих утесов или пещер и туннелей.

Слайд 7. Представление морфологии поверхности с помощью TIN.

Когда вы формируете TIN, то последовательно добавляете массовые точки, линии перегиба, и полигоны, чтобы развить и усовершенствовать модель поверхности. В любой момент вы можете прибавить дополнительные пространственные объекты поверхности, чтобы далее улучшить модель.

Точечные пространственные объекты поверхности. 3D точки (mass points) представляют точки, в которых намерено z-значение. После триангуляции они сохраняются как узлы с тем же самым местоположением и высотой. Эта начальная TIN отражает общую форму поверхности, но еще не достаточно точно изображает резкие изменения топографии типа потоков и гребней.

Линейные пространственные объекты поверхности. Когда добавляются линии перегиба, то в пересечениях с начальными ребрами создаются новые узлы. TIN обновляется, чтобы включить эти новые узлы.

Линии перегиба (breaklines) — линейные пространственные объекты, которые соответствуют естественным пространственным объектам — водотокам и гребням, или искусственным пространственным объектам типа автомагистрали. Есть два вида линий перегиба: резкие и нерезкие.

Резкие линии перегиба (hardbreaklines) отражают разрыв в уклоне, например, линию водотока. В то время как сама поверхность всегда непрерывна, то ее уклон может и не быть непрерывным. Резкие линии перегиба передают изрезанность поверхности и улучшают анализ и отображение TIN.

Нерезкие линии перегиба (soft breaklines) позволяют вам добавлять ребра, чтобы показать линейные пространственные объекты, не нарушающие уклон. Например, вам нужно добавить очертания дороги к вашей модели поверхности, но это может не

сказываться на локальном уклоне поверхности.

Площадные пространственные объекты. Аналогичным образом к элементам TIN добавляются полигоны. Полигоны могут представлять такие объекты как озера, морские акватории или строительные площадки. Когда добавляются полигоны, то в пересечениях границы полигона с начальными ребрами создаются новые узлы. TIN обновляется, чтобы включить эти новые узлы.

Полигоны, как элементы TIN, могут быть полигонами **замещения, полигонами стирания, полигонами отсечения, полигонами заполнения.**

Полигоны замещения (replace polygons) задают одно постоянное z-значение границе и всем высотам внутри нее.

Полигоны стирания (erase polygons) отмечают все области в пределах полигона как находящиеся вне зоны интерполяции для данного режима. При аналитических операциях типа вычислений объема, построения изолиний и интерполяции такие области игнорируются.

Полигоны отсечения (clip polygons) отмечают все области вне полигона как находящиеся вне зоны интерполяции для модели.

Полигоны заполнения (fill polygons) назначают целочисленное атрибутивное значение всем граням в пределах полигона. Никакой замены высоты, стирания или отсечения не происходит.

Слайд 8. Пространственные объекты поверхности в типовой TIN.

Эта иллюстрация изображает типовую TIN и демонстрирует некоторые ключевые типы пространственных объектов поверхности.

TIN может быть изображена условными знаками, показывающими высоту, экспозицию склона, крутизну, или другие характеристики поверхности. На этом изображении цветом показаны высоты.

Здесь же показаны ребра и узлы. Области с пологими склонами для отображения требуют немного массовых точек. Чтобы представить локальную форму крутых участков, с резкими изменениями топографии, необходимо большее количество массовых точек.

Озеро или пруд может быть смоделировано полигоном замещения, приводящим высоту водоема к постоянной высоте.

В общем случае TIN может моделировать прерывистую поверхность. Каждый отдельный массив земли изображается с помощью одной оболочки и интегрированного набора узлов, ребер и граней. Береговая линия острова или массива земли моделируется с

помощью добавления полигона отсечения. Этот полигон определяет оболочку или внешнее ребро набора граней. Оболочка обрисовывает берег или проектную границу.

Жесткие линии перегиба или полигоны (в данном случае - граница искусственного водоема) символизируют пространственные объекты, которые соответствуют существенному изменению крутизны.

Мягкие линии перегиба (в данном случае – дороги) или полигоны представляют пространственные объекты поверхности, которые не отмечают нарушение наклона, но прибавляют ребра к TIN, чтобы лучше моделировать поверхность.

В заключении этой части лекции перечислим основные преимущества и недостатки моделей TIN.

Преимущества: 1) разрешение TIN может быть переменным - более детальным в зонах, где поверхность сложная и менее детальным там, где поверхность достаточно простая. В этом типе модели можно использовать все исходные данные, в соответствии с которыми строится триангуляционная сеть. То есть информация не теряется, и нет избытка информации: в этом случае можно с большей эффективностью и полнотой применять методы интерполяции; 2) Линейные объекты, такие как дороги или потоки, можно представить с достаточной точностью в виде последовательности ребер треугольников, составляющих модель местности.

Помимо очевидных преимуществ, модели TIN имеют и свои недостатки. Их создание и пересчет достаточно трудоемки. Стоимость получения качественных исходных данных может быть высокой, а их обработка требует больше времени и ресурсов, чем в случае гридов.

Слайд 9. GRID – модель.

Структура GRID – моделей полностью соответствует структуре растровых данных. Грид состоит из массива регулярно распределенных прямоугольных, чаще квадратных, ячеек со значениями z . Размер ячейки определяет разрешение грида: чем меньше размер ячейки, тем точнее модель грида. Точно также как и в растровых наборах данных, значение ячейки грида может быть целочисленным или с плавающей запятой. Для гридов как и для растров можно строить пирамидные слои и применять те же методы отображения, что и для растров. Регистрация гридов в географическом пространстве выполняется теми же методами, что и регистрация растров.

Преимущества и недостатки GRID. Модель грид достаточно простая, ее обработка

обычно более эффективна, чем других моделей. Хорошо разработанные алгоритмы обработки гридов в массе своей проистекают из методов обработки растровых изображений. Данных по высотам рельефа в формате грида достаточно много и они относительно дешевы. С другой стороны, регулярная структура данных не в полной мере подходит для описания изменчивой поверхности (потеря информации в промежутках между точками регулярной сетки, или ее избыточность в случае плавно изменяющихся поверхностей). Кроме того, базовая структура грида не позволяет адекватно показать линейные объекты в приложениях, где требуются крупные масштабы.

Гриды обычно используются в приложениях, где применяются региональные и мелкомасштабные представления, в то время как TINы - для более детального крупномасштабного моделирования. Применяйте грид, если пространственная точность исходных данных невысока, или если Вам не нужно точное представление линейных объектов типа дорог или потоков.

Отличие гридов от растровых наборов данных состоит в том, что значения ячеек гридов- это величины, рассчитанные с помощью различных методов интерполяции (высота, величина и направление уклона, уровень шума от аэропорта, величина рН почвы, объем осадков, величина гравитационного поля и т.д.), в то время как растровые данные могут содержать непосредственно измеренные величины: в растрах, представляющих изображения, значения могут указывать цвет или спектральную отражающую способность. Другое отличие от растров состоит в том, что растровые наборы данных в общем случае могут иметь много каналов, грид- это всегда одноканальный набор данных.

Термин «грид» заимствован (транслитерация) от англоязычного grid, что в переводе означает решетка, сетка, сеть. Существует некоторая путаница в использовании этого термина, и многими он пока признается термином научного жаргона. Гридами также называют и растры, и цифровые модели рельефа. В общем случае гридом можно назвать любую структуру (растровое изображение, матрицу значений), являющуюся результатом пересчета нерегулярно распределенных данных в регулярные узлы. Этот пересчет выполняется путем интерполяции, аппроксимации, сглаживания и иных трансформаций исходных данных.

Алгоритмы создания и обработки гридов в настоящее время очень хорошо разработаны и используются многими программными продуктами (не только ГИС) для представления поверхностей. Исходные данные, как правило, распределены случайным образом, и переход к регулярной сетки осуществляется методами интерполяции входных значений. В ArcGIS алгоритмы интерполяции реализуются в дополнительных модулях **Spatial Analyst** и **3D Analyst**.

Слайд 10. Интерполяция.

Поскольку реальные поверхности непрерывны, невозможно зафиксировать все составляющие их местоположения. Поэтому описывающие их модели основываются на конечном числе точек, а для заполнения промежутков между этими местоположениями используются математические методы, называемые интерполяцией.

Интерполяция основана на предположении, что пространственно распределенные объекты связаны пространственной корреляцией; другими словами, близко расположенные объекты обладают сходными характеристиками. Например, если на одной стороне улицы идет снег, можно с большой степенью точности сказать, что на другой стороне улицы тоже будет идти снег. У вас, будет меньше уверенности в том, что снег идет во всем городе, и еще меньше уверенности в том, какова погода в соседней области. С помощью этой аналогии легко понять, что значения точек, расположенных ближе к ячейке, скорее будут похожи на рассчитываемое значение ячейки, чем значения дальних точек. Это основа интерполяции.

В нижней части слайда на рисунке слева представлены точки с известными значениями, справа - результат интерполяции. В данном примере точки замеров попадают в центры ячеек: на практике, как правило, такого не бывает. Одна из проблем при создании растров путем интерполяции - это то, что исходная информация в некоторой степени ухудшается - даже если точка данных попадает в пределы ячейки, нет гарантии, что ячейка получит именно такое значение.

Наиболее распространенными методами интерполяции являются: *Интерполяция значений с весом, обратно пропорциональным расстоянию, Сплайн, Тренд и Кригинг*. Каждый из них опирается на определенные предположения о том, как точнее вычислить значения ячеек. Для наилучшего соответствия расчетных значений реальным в разных случаях следует использовать разные способы интерполяции, в зависимости от того, какое явление отражают значения и как распределены точки замеров. Однако при любом методе интерполяции качество результата прямо пропорционально количеству исходных точек.

Таким образом, достоверность результатов интерполяции зависит от качества исходных данных и применяемого метода интерполяции.

Слайд 11. Метод обратно взвешенных расстояний (IDW).

Метод обратно взвешенных расстояний (ОВР или IDW). Этот метод предполагает, что влияние значения измеренной переменной убывает по мере увеличения расстояния от точки

замера. ОВР вычисляет значения ячеек, усредняя значения точек замеров, находящихся вблизи каждой ячейки. Чем ближе точка к центру оцениваемой ячейки, тем больший вес, или влияние, имеет ее значение в процессе вычисления среднего. Для определения значения в каждой выходной ячейке может быть использовано либо заданное число точек, либо все точки в пределах заданного радиуса. Этот метод следует применять, когда влияние переменной уменьшается с увеличением расстояния от точки измерения. Например, в большинстве случаев степень загрязнения окружающей среды быстро уменьшается при удалении от его источника. Или при интерполяции покупательной способности клиентов для анализа размещения магазинов, покупательная способность более удаленных клиентов меньше, поскольку люди обычно предпочитают делать покупки ближе к дому.

Средневзвешенное значение вычисляется как

$$\check{Z}(s_0) = \sum_{i=1}^N \lambda_i Z(s_i),$$

где $Z(s_i)$ - измеренное значение в ячейке i ; λ_i - вес измеренного значения ячейки i ; $\check{Z}(s_0)$ – значение в рассчитываемой ячейке; N - число измеренных ячеек, участвующих в интерполяции.

В ArcGIS при использовании метода ОВР вы можете контролировать несколько параметров интерполяции: степень, тип радиуса и барьеры.

Степень: Вы можете контролировать влияние точек замеров на интерполированное значение на основании их расстояния от ячейки. При задании большого показателя степени влияние ближних точек будет более значительным, поверхность получится более детальной и менее гладкой. Задание меньшего показателя степени увеличит влияние дальних точек, и поверхность получится более гладкой. Обычно используется степень 2, она же установлена по умолчанию.

На характеристики интерполируемой поверхности влияет также выбор радиуса (фиксированного или переменного), который ограничивает количество исходных точек участвующих в интерполяции значения ячейки.

Тип радиуса - фиксированный. При фиксированном радиусе величина радиуса постоянна, поэтому для всех интерполируемых ячеек круговая окрестность поиска точек одинакова.

Тип радиуса - переменный. При использовании переменного радиуса поиска задается количество точек, участвующих в вычислении значения интерполируемой ячейки,

поэтому радиус поиска для каждой ячейки индивидуален и зависит от того, как далеко от каждой ячейки набирается заданное число точек. Метод переменного радиуса позволяет создавать более точные поверхности, если плотность точек измерений различна в разных частях поверхности. Если вы знаете, что есть области, где точки измерений распределены редко, вы можете задать максимальное расстояние, которое не должен превышать поиск. Если радиус окрестности достигает максимального расстояния, не набрав заданного числа точек, вычисление значения этой ячейки будет выполнено на основе того количества точек, которое оказалось в пределах максимального радиуса.

Барьер: Набор данных линий или полигонов, используемый для установки границы поиска точек измерений. Линия может представлять обрыв, горный хребет или другой разрыв ландшафта. При вычислении значений ячеек будут учитываться только те точки, которые расположены по ту же сторону барьера, что и ячейка.

Слайд 12. Метод обратно взвешенных расстояний (продолжение).

Расчет гридов с использованием различных методов интерполяции реализуется в ArcGIS инструментами модулей **Spatial Analyst** и **3D Analyst**. На слайде показаны окна модуля **Spatial Analyst**, в которых для интерполяции выбирается метод ОВР. В окне «Обратно взвешенные расстояния» устанавливаются параметры интерполяции.

На рисунках справа показаны результаты интерполяции по исходному набору точек. Рисунок внизу демонстрирует результат интерполяции с учетом барьеров. Барьеры являются объектами линейного набора данных.

Слайд 13. Интерполяция: метод Сплайн (Spline).

Сплайн рассчитывает значения ячеек на основе математической функции, минимизирующей кривизну поверхности, в результате получается сглаженная поверхность, точно проходящая через все точки измерений. Идея аналогична растягиванию резиновой пленки, так чтобы она проходила через все точки, при минимизации кривизны поверхности. Этот метод наиболее удобен для медленно меняющихся поверхностей, таких, как высота земной поверхности, уровень грунтовых вод или концентрация вредных веществ.

Методы сплайна. Существует два метода сплайна: регуляризация и натяжение.

Регуляризация. Метод регуляризации создает гладкую, постепенно меняющуюся поверхность, значения в которой могут выходить за пределы диапазона значений замеров.

Натяжение. Метод натяжения меняет жесткость поверхности в зависимости от характера моделируемого явления. Он создает менее гладкую поверхность, значения в которой ближе к рамкам диапазонов значений замеров.

Дополнительные параметры

Вес. Для метода Регуляризации “вес” определяет вес третьей производной от поверхности в выражении минимизации кривизны.

Чем больше вес, тем более гладкой будет поверхность. Значения, заданные для этого параметра, должны быть больше или равны нулю, например: 0, .001, .01, .1 и .5.

Для метода Натяжения “вес” определяет вес натяжения. Чем больше вес, тем грубее поверхность. Значения должны быть больше или равны нулю, например: 0, 1, 5, и 10.

Число точек. Параметр “Число точек” определяет количество точек, на котором должны быть основаны вычисления. Чем больше точек вы зададите, тем большее влияние будут иметь удаленные точки, и тем более гладкой будет поверхность.

Слайд 14. Интерполяция: метод Тренд (Trend).

Этот метод подбирает математическую функцию- полином заданного порядка- ко всем исходным точкам. Он опирается на регрессию наименьших квадратов и создает поверхности с наименьшим отклонением от исходных значений.

На слайде показано уравнение полинома степени n для одномерного случая. В двумерном случае вид уравнений будет более сложный, но суть от этого не меняется. Метод наименьших квадратов отыскивает такие значения коэффициентов полинома (в одномерном случае - коэффициенты a_i), чтобы суммарная разница между фактическими значениями и вычисленными значениями (дисперсия) была минимальной:

$$\sum_{i=1}^N (\check{z}_i - z_i)^2 \rightarrow \min,$$

где \check{z}_i - рассчитанное (оценочное) значение параметра z ,

z_i - наблюдаемое значение параметра z .

Тренд описывает наиболее общие черты поверхности, или общую тенденцию в поведении изучаемой величины z . Полученная поверхность редко проходит через входные точки. Поверхность тренда 1-го порядка является плоскостью, характеризующий общий наклон поверхности, полиномы 2 и 3 порядка описывают более сложное поведение тренда.

В ArcGIS Метод Тренда доступен через ArcToolbox.

Слайд 15. Интерполяция: метод Кригинг.

Методы интерполяции ОВР, Сплайн и Тренд называют детерминистическими методами интерполяции, поскольку они непосредственно основаны на **измеренных** значениях в окрестности точек или на **заданных математических формулах**, определяющих сглаженность результирующей поверхности. Вторая группа методов интерполяции состоит из геостатистических методов (таких, как Кригинг), основанных на статистических моделях, включающих автокорреляцию (статистические зависимости между измеренными точками). Такая технология позволяет не только получить расчетную поверхность, но также определить значение точности или достоверности расчета.

Кригинг обрабатывает поверхности, считая их образованными из трех независимых величин. Первая, называемая дрейфом или структурой поверхности, представляет поверхность как общий тренд в определенном направлении. Далее, кригинг предполагает, что имеются небольшие отклонения от этой общей тенденции, вроде маленьких пиков и впадин, которые являются случайными, но все же связанными друг с другом пространственно (мы говорим, что они пространственно коррелированы). Наконец, мы имеем случайный шум (random noise), который не связан с общей тенденцией и не имеет пространственной автокорреляции. Кларк [Clarke, 1990] удачно иллюстрирует этот набор значений посредством аналогии: когда мы идем вверх по горе, рельеф местности изменяется в восходящем направлении между отправной точкой и вершиной; это - дрейф. По пути мы встречаем локальные снижения и повышения, сопровождаемые случайными, но все же коррелированными высотами, которые присущи этой поверхности. Также по пути нам встречаются камни, которые приходится переступить, их можно представлять как шум значения высоты, так как они не связаны непосредственно с основной поверхностной структурой, прежде всего создающей изменения высоты.

С каждой из трех переменных надо оперировать в отдельности. Дрейф оценивается с использованием математического уравнения, которое наиболее близко представляет общее изменение поверхности, подобно поверхности тренда. Локальное поведение поверхности и уровень шумов оценивается с использованием вариограммы (variogram, semivariogram).

Слайд 16. Метод Кригинг: вариограмма.

Вариограмма представляет собой график, на котором по горизонтальной оси откладывается расстояние между отсчетами, вертикальная ось несет так называемую полудисперсию (semivariance), которая определяется как половина дисперсии между каждым значением высоты и его соседями: Полудисперсия(расстояние h) = $0.5 * \text{среднее}[(\text{значение в точке } i - \text{значение в точке } j)^2]$ для всех пар точек, разделенных расстоянием h .

Формула включает вычисление квадрата разницы между значениями в паре точек. На рисунке слева показано образование пар заданной точки (красной) со всеми остальными точками замеров. Эта процедура выполняется для каждой точки замера. Часто расстояние для каждой пары точек уникально, и часто этих пар много. Быстрое построение графика для всех точек становится нереальным. Вместо работы с каждой точкой точки объединяются в *интервальные группы (лаги)*. Например, вычисление средней вариантности всех точек, расположенных на расстоянии больше 40 метров и меньше 50 метров.

Пространственная автокорреляция реализует основной принцип географии - близкие объекты более похожи, чем удаленные. Посмотрев внимательно на график полудисперсии, вы можете заметить, что когда расстояние между точками отсчета мало, полудисперсия тоже мала. Это значит, что измеренные значения близки и, следовательно, взаимосвязаны вследствие их пространственной близости. С ростом расстояния между точками растет и полудисперсия, показывая быстрый спад пространственной корреляции значений. Таким образом, полудисперсия является мерой взаимосвязи измеренных значений, зависящей от того, как близко друг к другу они находятся.

После отображения каждой пары точек (или интервальных групп), для них подбирается модель (аппроксимирующая функция показана на графике синей линией).

Если вы посмотрите на модель вариограммы, вы заметите, что на некотором расстоянии модель выравнивается. Расстояние, на котором модель начинает выравниваться, называется **предельный радиус корреляции (range)**. Точки, разделенные расстоянием меньше предельного радиуса корреляции, пространственно автокоррелированы, а точки на расстоянии больше радиуса влияния - нет. Это говорит нам о том, какая окрестность должна быть использована (например, в ОВР-интерполяции), чтобы охватить все точки, значения высоты которых будут взаимосвязаны.

Другим важным свойством графика является то, что аппроксимирующая кривая не проходит через начало координат. По идее, если между отсчетами нет

расстояния, то не должно быть и дисперсии, так как отсчеты являются по сути одной точкой. Но нужно помнить, что кривая является оценочной. Разница между нулевой дисперсией при нулевом лаге и предсказываемым положительным значением является остаточной, пространственно некоррелированной "шумовой" дисперсией, которая называется **остаточной дисперсией** (nugget variance). Эта остаточная дисперсия объединяет дисперсию ошибок: методических, аппаратурных и других неучтенных погрешностей.

Слайд 17. Моделирование вариограммы.

Подбор модели для эмпирической вариограммы аналогичен регрессионному анализу, в котором к облаку точек подбирается непрерывная линия или кривая. В качестве модели мы должны подобрать некоторую функцию (например, сферического типа), которая сначала поднимается, а затем выравнивается на больших расстояниях. На эмпирической вариограмме есть отклонения от модели; некоторые точки находятся выше кривой, а некоторые - ниже. Но, если мы измерим расстояние от каждой точки до кривой, то суммарное расстояние от точек, лежащих выше кривой должно быть близко суммарному расстоянию от точек, лежащих ниже кривой.

Существует множество разных моделей вариограмм: Круговая, Сферическая, Экспоненциальная, Гауссова и Линейная.

На слайде показаны некоторые из основных моделей и их отличия:

- Сферическая модель. Эта модель показывает постепенное снижение пространственной автокорреляции (и, соответственно, возрастание вариации) до определенного уровня, после которого автокорреляция равна нулю. Сферическая модель - одна из наиболее часто используемых.

- Экспоненциальная модель. Эта модель используется, когда при уменьшении расстояния автокорреляция возрастает по экспоненте. Здесь автокорреляция полностью исчезает только при бесконечном расстоянии. Экспоненциальная модель тоже часто используется.

- Линейная модель предполагает, что дисперсия возрастает прямо пропорционально расстоянию.

Выбор модели основывается на пространственной автокорреляции данных или на предварительном знании явления.

Выбор модели (моделирование вариограммы) - это ключевой шаг от пространственного описания к пространственному прогнозированию, так как основное назначение Кригинга (как и любого метода интерполяции) - это вычисление

предполагаемых значений в точках, не охваченных замерами. Вариограмма дает нам значение полудисперсии на любом расстоянии от точки замера. Выбор модели существенно влияет на вычисление неизвестных значений, особенно, вблизи точки начала отсчета. Чем круче кривая около точки начала отсчета, тем больше влияние ближайших соседей на вычисление. В результате полученная поверхность будет менее гладкой. Каждая модель лучше отражает свой тип явлений.

Слайд 18. Метод Кригинг: вычисление предполагаемых значений.

Говорят, что Кригинг использует данные дважды: первый раз для оценки пространственной автокорреляции данных и второй раз для вычислений значений. Другими словами Кригинг выполняет двухшаговый процесс: (1) создаются вариограммы для оценки значений статистических зависимостей, называемых пространственной автокорреляцией, и (2) определяются предполагаемые значения пустых ячеек.

Кригинг похож на ОВР в том, что он учитывает вес окружающих измеренных значений для того, чтобы определить расчетное значение для ячейки, в которой не было данных. Общая формула для обеих интерполяций представляет собой суммирование данных с учетом веса:

$$\hat{Z}(s_0) = \sum_{i=1}^N \lambda_i Z(s_i)$$

где

$Z(s_i)$ - измеренное значение в ячейке i ;

λ_i - неизвестный вес измеренного значения ячейки i ;

s_0 - расположение ячейки, для которой вычисляется прогноз;

N - число измеренных значений.

В ОВР вес, λ_i , зависит только от расстояния от оцениваемой ячейки.

Как и интерполяция с ОВР, Кригинг определяет вес окружающих измеренных точек, чтобы вычислить предполагаемое значение в неизмеренной ячейке. Как и в ОВР, точки, расположенные ближе к оцениваемой ячейке, имеют большее влияние. Однако, присвоение веса окружающим точкам в методе Кригинга более сложно, чем в ОВР. Так, в обычном кригинге вес, λ_i зависит от модели вариограммы, расстояния до

оцениваемой точки и пространственного распределения точек замеров вокруг оцениваемой точки.

При выполнении интерполяции очень важно правильно задать радиус поиска, или радиус интерполяции, который определяет число и конфигурацию точек замеров участвующих в формировании значения оцениваемой ячейки. При выборе радиуса интерполяции мы можем ориентироваться на предельный радиус корреляции и тем самым исключить точки с незначительным влиянием. Влияние удаленных точек не только мало, оно может быть даже негативным, если эти точки расположены на участке, характеристики которого сильно отличаются от участка, на котором находится оцениваемая ячейка. Другая причина для установки радиуса поиска - скорость вычислений.

Также как и в ОВР радиус поиска может быть **фиксированным и переменным.**

Слайд 19. Условия применения Кригинга.

Метод Кригинга назван по фамилии южно-африканского геолога D.G.Krige (неица по происхождению), который применял его для определения запасов золота в россыпях. Кригинг существует в двух основных формах. Общий, или универсальный, Кригинг применяется, когда поверхность оценивается по нерегулярно распределенным отсчетам при наличии тренда. Одинарный Кригинг по своей сути является лишь улучшением метода ОВР, в котором учитываются не только расстояния от интерполируемой точки, но и расстояния между самими исходными точками так, что веса более близких друг к другу исходных точек уменьшаются.

Универсальный Кригинг предполагает, что в данных имеется тенденция к доминированию определенных значений (например, господствующий ветер), и его можно смоделировать с помощью детерминистической или полиномиальной функции. Этот полином вычитается из исходных значений измерений, и автокорреляция моделируется по случайным ошибкам. Когда к случайным ошибкам подобрана модель, перед вычислениями полином добавляется обратно к предполагаемым значениям, чтобы получился осмысленный результат. Универсальный Кригинг используется в тех случаях, когда вы знаете, что в ваших данных есть определенные тенденции, и вы можете привести научное описание для их подтверждения.

Одинарный Кригинг- это наиболее общий и широко используемый из методов Кригинга. Он основан на предположении, что постоянное среднее значение неизвестно. Это разумное предположение, если нет никакой причины предполагать обратное.

Кригинг часто дает довольно точные оценки пропущенных значений, но эта точность обходится ценой времени и вычислительных ресурсов. Но даже при этом кригинг имеет еще одно преимущество перед другими методами интерполяции, - он не только дает интерполированные значения, но также и оценку возможной ошибки этих значений. Это может привести на мысль, что данный метод следует применять повсеместно, но увы. Когда мы имеем дело с большим уровнем локального шума из-за ошибок измерений или большие вариации значений между отсчетами, в данном методе становится трудным или даже невозможным построение кривой полудисперсии. А в таких условиях результаты кригинга будут не лучше, чем полученные другими методами.

К тому же Кригинг (впрочем как и другие методы интерполяции) очень чувствителен к форме пространственного распределения исходных точек: в случае крайне неравномерного или анизотропного (когда в одном направлении расстояние между точками существенно выше, чем в другом) распределения точек Кригинг может дать результат хуже чем другие методы.

На слайде показан пример неудачного применения Кригинга для моделирования рельефа дна водохранилища в случае анизотропного распределения исходных точек. Иллюстрации показывают, что простая триангуляция в этом случае дает более правдивую картину рельефа в сравнении с методом кригинга.

Слайд 20. Сравнение методов интерполяции.

Интерполяция для многих видов исследований является очень полезным инструментом. Допустим, что вы являетесь специалистом в области инженерной геологии, и некая строительная компания обратилась к вам с вопросом о том, не попадет ли планируемая застройка в зону возможного затопления. У вас нет карты, показывающей границы этой зоны. При этом вы знаете, что максимальный уровень наводнений за сто лет составил 60 метров над уровнем моря. У вас также имеются заметки о нескольких прежних участках строительства, и они включают данные высот для каждого построенного дома. Изобразив данные на карте местности, вы можете использовать интерполяцию для оценки высот для территории вашего участка. По этим данным вы сможете начертить изолинию, показывающую зону наводнений за 100 лет, и, просто сравнив ваше местоположение с ней, узнаете, нужно ли менять место.

Другой случай: нужно проложить шоссе по ненанесенной на карту территории и дорожные строители не могут начать строительство, не зная среднего уклона (градиента поверхности). Вы можете создать карту поверхности тренда, чтобы показать общий

характер уклона. Или, положим, вы являетесь горным инженером, пытающимся определить общий тренд рудного месторождения на основе информации из множества кернов, показывающих кровлю и подошву залежи. Интерполяция даст информацию о форме кровли и подошвы, а также толщине рудного слоя и его уклона под землей. Кроме того, метод кригинга окажется полезным в оценке качества рудного слоя.

Можно привести еще множество примеров, где интерполяция является очень полезным, или даже единственным источником информации об объекте или явлении. Но всегда следует помнить о том, что интерполяция, является по своей сути предсказательной моделью, а не точным решением.

В общем случае можно сказать, что чем больше исходных точек мы имеем, тем более точной будет интерполяция и тем с большей вероятностью интерполированная поверхность будет хорошей моделью. Чем сложнее поверхность, тем больше точек данных требуется. А для важных объектов, таких как впадины и долины рек, требуют дополнительные точки данных, чтобы гарантировать представление необходимой подробности. Однако, существует предел числу отсчетов, которые могут быть сделаны для любой поверхности. Постепенно достигается момент снижения отдачи: большее количество точек не улучшает существенно качество результата, но увеличивает время вычислений и объем данных. При очень большом числе точек точность на самом деле снижается.

Метод интерполяции также существенно влияет на результат. Это положение уже проиллюстрировал предыдущий слайд. На этом слайде представлены поверхности, полученные по одному и тому же набору данных (точки наблюдения распределены довольно равномерно), но с использованием различных методов интерполяции. Видно, что общие черты поверхности от рисунка к рисунку повторяются, но все же есть и различия. Также разнятся и диапазоны интерполированных значений.

Линейная интерполяция дает значения интерполируемых ячеек самые близкие к исходным значениям. Метод дает наилучший результат, когда точки съемки расположены не в регулярном порядке, а фиксируют «характерные» точки поверхности (например, локальные максимумы и минимумы). Но при малом числе точек поверхность будет грубой.

К достоинствам метода ОВР можно отнести то, что он является локальным, то есть на значения моделируемой функции и ее производных в любой точке практически не оказывают влияния значения в опорных точках, «далеко» отстоящих от нее. Этот метод хорош, когда опорные точки расположены достаточно плотно. Кроме того, изменяя весовую функцию и степень полинома, можно в интерактивном режиме

подбирать модельную поверхность, наиболее приемлемую с точки зрения специалиста.

В случае редко расположенных опорных точек метод сплайн-интерполяции или кригинга дает лучшие результаты. Считается, что ОВР лучше всего подходит для моделирования явлений, влияние которых быстро уменьшается с расстоянием (объем продаж магазина, концентрация вредного вещества). Сплаины лучше подходят для моделирования плавно изменяющихся поверхностей (наземный рельеф, уровень грунтовых вод).

Метод кригинга является универсальным методом, и получил широкое распространение в геологии при оценке качества запасов полезного ископаемого. В основном он дает неплохие результаты.

В любом случае, прежде чем обращаться к моделированию «невидимой» поверхности, следует поискать аналогию (на что она похожа, или какие известные явления проявляют себя схожим образом). Если же аналогии найти не удастся, следует выдвинуть гипотезу для привлечения косвенных данных. Если же гипотезы нет, то следует брать аппроксимирующую функцию возможно более простую (например, линейную).

Если мы вернемся к предыдущему слайду, то еще раз увидим, что методы интерполяции могут дать модель поверхности далекую от действительности. Поэтому некоторые исследователи продолжают строить изолинии вручную, повергая в изумление своих менее вдумчивых коллег; они хорошо знают природу явления и их не может удовлетворить «поведение» предложенных математических функций.

Тема 13. Классификации.

Слайд 2. Принципы классификации.

Классификация - одна из самых естественных вещей, которые делают люди. Даже самые заурядные вещи подвергаются классификации. Мы носим одежду различных типов, водим машины разных марок, имеем различные категории занятости. Мы мужчины или женщины, члены той или другой политической партии или независимые, взрослые, подростки или дети и т.д. Используемые нами классификации приводят людей, места и вещи некоторую систему, которая дает нам понимание того, как они функционируют: подобно другим членам своих групп и отлично от членов других групп. В этом - суть классификации.

Число возможных методов классификации и переклассификации бесконечно. Все они сильно зависят от потребностей пользователя, как и работа всей ГИС в целом. Благодаря сегодняшним ГИС мы можем хранить исходные данные, не подвергнутые классификации, и обрабатывать их в соответствии с нашими потребностями. Это наиболее ценная сторона классификации средствами ГИС: чем более точно ваша классификация настраивается на нужды пользователей, тем более полезной она становится.

Все множество методов классификации и переклассификации условно можно разделить на 2 больших группы: 1) классификации на основе атрибутивной информации; 2) классификации на основе информации о положении, размере и форме объектов. В большинстве случаев эти представления не выделяются как отдельные методы: все они часто комбинируют друг с другом для создания широкого разнообразия методов классификаций.

Мы рассмотрим некоторые из этих методов, начиная с простейшей переклассификации на основе атрибутивной информации.

Слайд 3. Простейшая переклассификация на растровом типе данных.

Очень часто (особенно если данные получены из внешнего источника) мы работаем с уже классифицированными кем-то данными. И нашу работу можно охарактеризовать как переклассификацию полученных данных. Точки, линии и полигоны могут быть переклассифицированы простым перекодированием

атрибутов в их таблицах или перекодированием значений ячеек растра для создания новых наборов данных. В этом простом процессе пользователь меняет сами атрибуты и не более того.

В случае растрового набора данных, вы просто меняете числа кодов или имена атрибутов для этих ячеек растра.

На слайде представлен простой пример перекодировки ячеек растра LANDUSE. Первоначально на растре имеется 7 категорий ячеек :

Атрибутивная таблица LANDUSE			
Value	Count	Landuse	Новый код
1	294	Переходная зона	5
2	62187	Вода	NO DATA
3	28	Пустыри	6
4	36034	Постройки	3
5	85054	Сельскохозяйственная зона	10
6	671722	Леса	4
7	12241	Болота	NO DATA

Известно, что некоторые типы землепользования лучше подходят для строительства нового здания, чем другие, с точки зрения стоимости строительства. Чтобы показать пригодность типов земель с точки зрения стоимости, переклассифицируем растр так, что меньшие значения будут показывать менее пригодное землепользование, большие- более выгодное. Вода и Болота получают значение “No Data”, поскольку они совсем непригодны, их нужно исключить.

В результате получаем растр с меньшим числом категорий, пригодный для дальнейшего анализа.

Таким образом, переклассификация на растре означает просто замену входных значений ячеек новыми выходными значениями.

Существует множество различных целей переклассификации данных. Вот наиболее распространенные из них:

- **Замена значений на основе новой информации.** Переклассификацию полезно использовать, когда вы хотите заменить значения во входном растре новыми значениями. Это может потребоваться, если вы выяснили, что значение одной или нескольких ячеек в действительности должно быть другим, например, если изменился тип землепользования.
- **Группировка значений.** Вам может потребоваться упростить информацию в растре. Например, вы можете объединить различные типы лесов в один класс лесов.

- **Переклассификация значений в наборе растров по единой шкале.** Другая причина для переклассификации - присвоение растру значений предпочтительности, чувствительности, приоритета или другого аналогичного критерия. Эту операцию можно выполнить с одним растром (растру типов землепользования можно присвоить значения 1–10, отражающие вероятность эрозии), или с несколькими растрами, чтобы привести их к общей шкале значений. Например, при поиске склонов, на которых наиболее велика опасность появления лавины, исходными данными должны быть уклон, тип почв и растительность. Каждый из этих растров можно переклассифицировать по шкале 1-10 в зависимости подверженности каждого атрибута каждого растра образованию лавин (например, самым крутым склонам в растре уклона должно быть присвоено значение 10, поскольку они наиболее подвержены возникновению лавин).
- **Замена определенных значений на “нет данных” или замена “Нет данных” на значение.** Иногда вам нужно исключить определенные данные из анализа, например, если на некоторые типы почв (например, на болота) наложены ограничения, не позволяющие вести на них строительство. В таких случаях вы можете заменить эти значения на “Нет данных”, чтобы исключить их из дальнейшего анализа. В других случаях может потребоваться заменить “Нет данных” на значение, например, если появились данные для ячеек, ранее помеченных, как “Нет данных”.

Слайд 4. Переклассификация с использованием операторов отношений.

В ArcGIS переклассификацию растров также можно выполнить, используя Калькулятор растров. Вы можете использовать язык *Алгебры карт* для создания запросов выбора. Условия выбора ячеек растра описываются с помощью операторов отношений: == - равно,

> - больше чем,

< - меньше чем,

≠ - не равно,

>= - больше или равно,

<= - меньше или равно.

Операторы отношений оценивают определенные условия отношений. Если условие TRUE (истинно), на выходе присваивается 1, если условие FALSE (ложно), на выходе присваивается 0. Например, результатом “Inlayer1 <> 3” (значения слоя Inlayer1 не равны 3) может создать выходной растр всех территорий кроме лесов, если лесам соответствует значение 3.

Слайд 5. Простейшая переклассификация на векторном типе данных.

В векторных системах процесс переклассификации требует как изменения атрибутов, так и геометрии объектов. Во-первых, надо удалить все линии, которые **разделяют** классы, которые должны быть объединены. Эта операция **называется растворением границ**. Затем в атрибутивной таблице нужно записать новые значения атрибутов вновь образованных объектов.

На практике эту операцию проще сделать наоборот: создать новое поле, на основе которого будет проходить объединение векторных объектов, и задать векторным объектам одного класса один и тот же атрибут. Возможно, такое поле уже в таблице существует. Затем воспользоваться функцией Dissolve (Слияние), которая создаст новый векторный набор данных, представляющий векторные объекты с одинаковым значением выбранного атрибута как один объект. Например, у вас есть полигональный шейп-файл, представляющий административные единицы РФ на уровне районов. В атрибутивной таблице наверняка есть поле (условно назовем его «great_unit»), указывающее на принадлежность района области, или округу, или республике: словом более крупной территориальной единице. Используя функцию Dissolve, вы легко можете создать новый класс объектов, представляющий более крупные административные единицы РФ, выбрав в качестве поля для слияния поле «great_unit».

И в растровой, и в векторной переклассификации полигонов имеется интересная особенность. В обоих случаях по окончании мы имеем меньшее число категорий, чем имели вначале. Этот результат называется **агрегированием данных**.

Слайд 6. Ранжированные классификации.

До сих пор мы рассматривали переклассификацию данных для номинальной шкалы измерений. Нетрудно представить, как можно было бы переклассифицировать данные других шкал. В картографических методах это делается созданием диапазонов категорий данных, что часто называется **ранжированными** (range graded) классификациями. Такой сценарий требует от нас всего лишь перекодировать эти данные на основе классовых интервалов, в которые они попадают. И, как и в случае данных номинальной шкалы, мы просто перекодируем ячейки растра или выполняем замену атрибутов и растворение границ.

Ранжированные классификации, как правило, применяются для числовых атрибутов. Методов подобных классификаций множество, рассмотрим некоторые из них. Наиболее распространенным методом является **метод естественных интервалов**. В ArcGIS этот метод классификации устанавливается по умолчанию. Метод естественных интервалов создает классы таким образом, чтобы сгруппировать близкие по значениям данные. Другими словами в пределы одного класса попадают (по возможности) объекты, значения атрибутов которых близки.

На рисунке в верхней части слайда показан результат классификации штатов США по численности их населения методом естественных границ. Штаты разбиты на 5 классов: в первый класс попадают всего 2 штата (Калифорния и Нью-Йорк), численность населения которых (23 и 18 млн. человек) существенно отличаются от численности населения других штатов. Второй класс объединяет 6 штатов, численность которых составляет от 9 до 14 млн. человек, остальные классы объединяют большее число штатов с существенно меньшим числом жителей. Таким образом, метод естественных интервалов дает «правильное» представление о численности населения США.

Метод равноинтервальной классификации разбивает значения атрибута на равные по размеру диапазоны значений. В приведенном примере показан результат классификации по методу равных интервалов. Данные по численности населения 48 штатов разбиты на 5 классов. В результате создан отдельный класс для штата Калифорния с населением 23 миллиона в сравнении со следующим на втором месте по уровню населения штатом Нью-Йорк с населением около 18 миллионов. Прерывания между другими штатами в данном случае не так очевидны.

Слайд 7. Ранжированные классификации (продолжение).

При **равновеликой классификации** (или метод квантиль) значения подразделяются таким образом, что каждый класс содержит одинаковое количество объектов. Возможно, количественные классы наиболее доступны для понимания, однако они могут ввести в заблуждение. Например, значения по уровню населения не годятся для их классификации таким методом, поскольку места с меньшей численностью населения будут необоснованно включены в один класс с местами с более высокой численностью населения. На приведенной карте, где 48 штатов подразделяются на пять классов, Мичиган, с численностью населения около 9

миллионов человек, попадает в один класс с Калифорнией, имеющей население около 23 миллионов. Вы можете скорректировать данное искажение, увеличив число классов. Например, используйте 10 классов вместо пяти.

Есть еще один интересный метод классификации, который основан на статистических характеристиках атрибута. Когда вы используете классификационный метод среднеквадратического отклонения, программа находит среднее значение, затем устанавливает верхние и нижние границы класса с интервалом по значению среднеквадратического отклонения, равному 1/4, 1/2 или 1. Затем программа сгруппирует в два класса все значения, удаленные от среднего значения на величину, равную трем среднеквадратическим отклонениям: ">3 Std Dev." (превосходит среднее значение больше чем на 3 среднеквадратических отклонения) и "<3 Std Dev." (недостает до среднего значения больше чем на 3 среднеквадратических отклонения). В ArcGIS вы всегда можете просмотреть статистические характеристики (такие, как среднее, минимальное, максимальное значения, медиану и среднеквадратическое отклонение) для любого числового поля.

Классификации с использованием статистики позволяют увидеть объекты атрибутивные значения которых существенно отличаются от средних по набору данных.

Слайд 8. Ранжированные классификации: растровый тип данных.

Ранжированные классификации на растровом типе данных выполняются совершенно аналогичным способом. Используя какой – либо метод классификации (на слайде показан метод естественных интервалов), весь диапазон значений ячеек (целочисленных или с плавающей запятой) разделяются на интервалы. Число интервалов определяется числом классов. Каждому интервалу присваивается новый код. Далее выполняется операция переклассификации.

Слайд 9. Переклассификации основанные на окрестности.

Переклассификация на основе атрибутов во всех перечисленных случаях дает полезную информацию. Но во многих других случаях этого бывает недостаточно, т.к. она ограничивает нас атрибутами в пределах каждого объекта. Было бы замечательно, если бы могли классифицировать объекты с учетом их окружения. Другими словами, если бы мы могли взглянуть на объекты исследования с высоты птичьего полета, то могли бы узнать не

только о существовании определенного объекта, но и о том, как он расположен по отношению к другим объектам. Такие процедуры переклассификации основаны на идее характеризования каждого объекта как части большей **окрестности** объектов.

Окрестность может быть рассмотрена как вся область в целом или как некоторая часть области. С этой точки зрения различают целевой анализ и анализ общего соседа.

Целевой анализ, также называемый непосредственной окрестностью, включает только места, непосредственно прилегающие к целевой области или точке. Анализ общего соседства, называемый также расширенной окрестностью, включает местоположения, которые находятся в непосредственной близости, а также и удаленные на некоторое расстояние.

Функции анализа обоих этих типов разделяются на статические, в которых анализ проводится сразу по всей выбранной целевой области, и функции скользящего окна, где анализ проводится только в рамках окна, которое перемещается по покрытию.

Рассмотрим некоторые виды анализа по окрестности. Все эти виды анализа можно отнести к комбинированной переклассификации, в которой учитываются и пространственное положение объектов, и их атрибутивные характеристики.

Функция Статистика по окрестности применяется для растровых типов данных. Эта функция на основе входного растра (или грида) создает новый растр, у которого значение каждой ячейки является функцией от значений входных ячеек в окрестности этой ячейки.

Для каждой ячейки входного растра функция Статистики по окрестности вычисляет статистическую величину на основании значения этой ячейки и ячеек в пределах заданной окрестности, и затем записывает полученное значение в соответствующую ячейку выходного растра.

Формы окрестностей. Окрестность может представлять собой прямоугольник любой площади, круг любого радиуса, кольцо любого радиуса или сектор круга (клин) в любом направлении. Параметры окрестностей можно задать либо в ячейках, либо в единицах измерения карты.

По окрестности каждой ячейки входного растра могут быть вычислены следующие виды статистических величин:

- **Большинство:** определяет для каждого местоположения ячейки наиболее часто встречающееся значение в окрестности
- **Максимум:** определяет для каждого местоположения максимальное значение в окрестности

- **Среднее:** определяет для каждого местоположения среднее значение в окрестности
- **Медиана:** определяет для каждого местоположения медиану значений в окрестности
- **Минимум:** определяет для каждого местоположения минимальное значение в окрестности
- **Меньшинство:** определяет для каждого местоположения наиболее редкое значение в окрестности
- **Диапазон:** определяет для каждого местоположения диапазон значений в окрестности
- **Стандартное отклонение:** определяет для каждого местоположения стандартное отклонение значений в окрестности
- **Сумма:** определяет для каждого местоположения сумму значений в окрестности
- **Разнообразие:** определяет для каждого местоположения количество уникальных значений в окрестности

Функция окрестности для отдельной ячейки. Рассмотрим ячейку со значением 5 на схеме в центральной части слайда. При работе с прямоугольной окрестностью 3x3 ячейки, сумма значений ячеек соседей плюс значение самой обрабатываемой ячейки составит 24. Поэтому ячейке выходного растра, позиция которой совпадает с позицией обрабатываемой ячейки входного растра, присваивается значение 24.

Функция окрестности для всего набора данных. Каждая ячейка представленного ниже выходного растра получила значение суммы по окрестности 3x3 ячейки. Выделенные красным ячейки указывают окрестности обрабатываемых ячеек со значением 5, 0, «нет данных» и выходных ячеек со значением 16,14, «нет данных». Этот процесс повторяется для каждой ячейки, пока не будут вычислены значений всех ячеек.

Слайд 10. Функция «Статистика по окрестности».

На этом слайде показан результат применения функции «Статистика по окрестности» модуля Spatial Analyst. Для исходного растра были рассчитаны среднее и стандартное отклонение в окрестности размером 20 ячеек, заданной в виде круга.

Растр среднего значения по окрестности позволяет увидеть наиболее общие черты исходного растра, Растр стандартного отклонения показывает области, где имеются наибольшие и наименьшие вариации значений исходного растра.

Слайд 11. Переклассификации основанные на окрестности: фильтры.

В предыдущей теме курса мы рассматривали такую меру формы полигона как развитость границ. Развитость границ рассчитывается с помощью оконной функции, которая представляет собой матрицу чисел. Оконные функции также используют для различных видов фильтрации данных, поэтому их называют просто «фильтрами». Итак фильтр представляет собой матрицу чисел (размером $n \times m$ ячеек), которые служат операндами в выражениях со значениями ячеек растра.

Довольно часто этот метод, используется в обработке изображений дистанционного зондирования, но имеет такую же применимость и в растровых ГИС. В частности, фильтры используются для выделения краев областей или линейных объектов (**фильтры высоких частот (ФВЧ)**), усиления общих градиентов и устранения мелких флуктуации и шумов (**фильтры низких частот (ФНЧ)**), или даже для подчеркивания ориентации объектов (**анизотропные фильтры**).

Фильтр высоких частот предназначен для выделения деталей в растровом наборе данных содержащих относительно близкие значения. Допустим, что мы заинтересованы в обнаружении мелких гребней в растровом топографическом покрытии. Каждая ячейка растра содержит отсчет высоты, и мы хотим подчеркнуть контраст между несколько более высокими значениями для гребня и несколько более низкими значениями, окружающими этот объект. Типичный метод для выполнения такой фильтрации высоких пространственных частот состоит в том, чтобы создать матрицу фильтра 3×3 с весовым коэффициентом 9 в центральной ячейке и минус 1 — во всех остальных. Этот фильтр помещается поверх каждой группы ячеек растра 3×3 в нашем покрытии, и члены каждой пары соответствующих ячеек растра и матрицы фильтра перемножаются: то есть, значение высоты в центральной ячейке растра топографического покрытия умножается на коэффициент 9, а все остальные значения умножаются на минус 1. Затем, эти девять только что созданных произведений суммируются для получения результирующего значения центральной ячейки растра. Другими словами, эта одиночная операция с двумя матрицами по девять чисел производит единственное значение, которое помещается в центр нового покрытия.

Следующий шаг должен переместить фильтр на одну ячейка растра вправо, так,

чтобы центральная ячейка в топографическом покрытии была теперь (3,2). Вычисления выполняются также, как и прежде, приводя к новому значению ячейки (3,2). Процедура повторяется для всего покрытия.

Заметьте, что некоторые значения стали на много больше, а другие намного меньше, так что в итоге в результирующем покрытии гребни видны гораздо лучше.

Если вы хотите подавить более высокие пространственные частоты, чтобы удалить не имеющие ценности топографические флуктуации, получив таким образом покрытие, которое показывает топографию в более общем (упрощенном, сглаженном) виде, вы можете использовать фильтр с меньшим различием коэффициентов. Наиболее распространен ФНЧ с матрицей 3x3, все коэффициенты которой равны 1. В результате значения каждой ячейки будут усреднены с соседними.

Слайд 12. Высокочастотные (ФВЧ) и низкочастотные (ФНЧ) фильтры.

На этом слайде показаны результаты фильтрации реального топографического растра фильтрами низкой и высокой частоты. Применение фильтра низкой частоты позволяет удалить мелкие детали рельефа, и представить его в более обобщенном (упрощенном) виде.

Применение ФВЧ позволяет более отчетливо отобразить и изучить на растре мелкие детали рельефа.

Хотя обычные методы фильтрации растра используют стандартные матрицы 3x3 с "прошитыми" коэффициентами, нет необходимости к этому привязываться. Большинство программ, выполняющих фильтрацию, позволяют менять размер матрицы и значения коэффициентов. Чем больше размер матрицы, тем большее пространственное усреднение можно получить в случае ФНЧ, поскольку большее число ячеек принимают участие в усреднении. В случае ФВЧ с ростом размера матрицы может быть получено более качественное подчеркивание мелких деталей поверхности. Решение использовать нестандартные размеры и коэффициенты часто возникает в результате экспериментирования.

Слайд 13. Анизотропные фильтры.

Поскольку края областей и линейные объекты имеют различные ориентации (например, вы можете обнаружить какое-то число гребней, связанных с

некоторым геологическим циклом образования складок и эрозии), иногда полезно согласовать фильтр с определенной ориентацией. Например, если вы хотите выделить гребни, ориентированные с востока на запад, вы должны использовать фильтр, у которого ячейки средней строки содержат положительные числа, а остальные - отрицательные. Для ориентации с севера на юг нужно сделать числа среднего столбца положительными, а остальные - отрицательными. Для ориентации с северо-запада на юго-восток нужно сделать положительными числа главной диагонали матрицы фильтра, либо использовать более контрастный фильтр (подобный тому, что показан справа).

Слайд 14. Переклассификация поверхностей: уклон.

Поверхности также могут быть переклассифицированы на основе идеи окрестности. Наиболее широко используются следующие преобразования поверхностей: **расчет уклона поверхности, азимута (или экспозиции склона), отмывка рельефа, взаимная видимость, расчет объемов и площадей.**

Наиболее просто переклассификация поверхностей выполняется на моделях TIN, так как для треугольников TIN их местоположение и ориентация в трехмерном пространстве определяется однозначно. И более того, элементы ориентирования каждой грани записаны в структуре TIN, и, следовательно, не требуют расчета. Для модели TIN уклон (θ) будет представлять собой угол между гранью и горизонтальной плоскостью. Угол ϕ - азимут уклона или экспозицию склона.

Для модели GRID (или другого растра) расчет уклона и экспозиции требует дополнительных вычислительных затрат. В этом случае обычный метод вычисления уклона состоит в том, чтобы провести наиболее подходящую плоскую поверхность через соседние точки и измерить отношение изменения высоты на единицу расстояния. Вернее, ГИС просчитает это отношение по всему покрытию, создавая набор категорий величины уклона. Простейший способ состоит в оценке восьми непосредственных соседей каждой ячейки растрового покрытия. Программа строит плоскость по восьми ближайшим соседним ячейкам поиском либо наибольшей величины уклона для окружения ячеек растра, либо среднего уклона (обычно выбор – за пользователем). Для каждой группы ячеек программа использует разрешение растра, чтобы оценить отношение разности высот к расстоянию, на котором это превышение наблюдается. Построив подходящую плоскость, также можно оценить и ее экспозицию.

Слайд 15. Уклон (продолжение).

Функция Уклон вычисляет максимальную скорость изменения значения между соседними ячейками—например, максимальный угол наклона земной поверхности или градиент гравитационного поля. Каждой ячейке выходного раstra присваивается значение уклона. Чем меньше значения уклона, тем ровнее территория; чем больше значение уклона, тем круче склоны. Выходной набор данных уклона можно вычислить в форме градуса уклона и процента уклона.

Когда угол уклона равен 45 градусам, на единицу подъема приходится единица по горизонтали. При выражении в форме процента такой угол уклона равен 100 процентам. Обратите внимание, что по мере приближения уклона к вертикальному (90°), процент уклона приближается к бесконечности.

Функция Уклон наиболее часто применяется к наборам данных высот, как показано на рисунке внизу. В выходном наборе данных уклона наиболее крутые склоны закрашены красным.

Эту функцию можно применять также к другим непрерывным наборам данных, чтобы определить участки резкого изменения значения.

Слайд 16. Переклассификация поверхностей: экспозиция.

Экспозиция указывает направление наиболее крутого уклона от каждой ячейки к соседним. Ее можно считать направлением уклона или направлением по компасу, куда обращен склон холма.

Экспозиция измеряется в градусах против часовой стрелки от 0 (направление на север) до 360 (опять на север, сделав полный круг).

Значение каждой ячейки в наборе данных экспозиции указывает направление склона в данной ячейке. Плоские участки не имеют направления, и им присваивается значение «-1» (или FLAT).

С помощью функции Экспозиции вы можете:

- Найти на горе все склоны, направленные на север, при поиске наилучших склонов для катания на горных лыжах.
- Выяснить освещенность солнцем каждой точки территории при изучении биоразнообразия в какой-либо области.
- Найти все южные склоны в гористом районе, чтобы определить, где в первую очередь будет таять снег, при изучении опасности затопления жилых районов паводком
- Определить плоские участки при поиске места для аварийной посадки самолета.

Слайд 17. Переклассификация поверхностей: отмывка рельефа

Функция Отмывки рельефа создает эффект искусственного освещения поверхности, определяя значения освещенности в каждой ячейке растра. Для этого она устанавливает позицию гипотетического источника света и вычисляет значения освещенности для каждой ячейки относительно других ячеек. Таким образом, можно значительно улучшить вид данных при визуализации или создать новую информацию для дальнейшего анализа.

По умолчанию оттенки шкалы серого цвета связываются с числами от 0 до 255 (по возрастанию от черного к белому).

Параметры отмывки рельефа:

Азимут - это угловое направление расположения солнца, измеряемое от севера против часовой стрелки, в градусах от 0 до 360. По умолчанию установлено значение 315 (северо-запад).

Высота - это угол высоты источника освещения над горизонтом: от 0 (горизонт) до 90 градусов (зенит). По умолчанию установлено значение 45 градусов.

Помещая растр высот поверх созданной отмывки рельефа, а затем, сделав растр высот прозрачным, вы можете создать реалистичную картину ландшафта. Добавьте другие слои, например, реки или дороги, чтобы еще повысить информативность изображения.

При моделировании теней (по умолчанию) вы вычисляете локальную освещенность - находится ячейка в тени или нет.

При моделировании теней вы можете вычислить, какие ячейки попадают в тень других ячеек в определенное время дня. Ячейки, попадающие в тень других ячеек, получают значение 0; другим ячейкам присваиваются значения от 1 до 255. Вы можете переклассифицировать полученный растр в растр, показывающий участки, полностью погруженные в тень.

Слайд 18. Переклассификация поверхностей: видимость.

В то время как рассмотренные выше способы классификации поверхностей характеризуют всю поверхность или ее точки по отношению к соседним, рельеф может анализироваться и более сложным образом. Взаимная видимость показывает, что если

вы расположены в определенной точке топографической поверхности, то одни области рельефа будут вам видны (области видимости), а другие нет.

В векторной системе (модели TIN) простейший метод состоит в соединении точки наблюдателя с каждой возможной целевой точкой покрытия. Затем выполняется трассировка лучей, т.е. вы следуете вдоль линии (луча), ища отметки высоты, которые выше этой линии. Более высокие точки будут закрывать для наблюдателя то, что за ними.

Растровые методы определения действуют почти также, но имеют большую вычислительную стоимость, так как высота ячейки наблюдателя, должна быть так и иначе сравнена с высотами всех остальных ячеек.

Видимость полезна, когда вам нужно определить, насколько хорошо видны разные объекты на ландшафте- например, для поиска мест для коммуникационных вышек.

В приведенном ниже примере определен обзор из точки наблюдения. Растр высот показывает высоту земной поверхности, а точка наблюдения помечена голубым кругом со звездочкой. Зеленые ячейки видны из точки наблюдения, красные ячейки не видны.

Функция Видимости может учитывать кривизну земной поверхности и распределение лучей света при прохождении сквозь атмосферу, если для входной поверхности определена проекция.

Слайд 19. Переклассификация поверхностей: расчет объемов.

Функция Насыпей/Выемок показывает площади и объемы изменений между двумя поверхностями. Она указывает площадь и объем поверхности, в которой произошли изменения, связанные с добавлением или удалением вещества поверхности.

Имея на входе две поверхности для данной области для двух различных периодов времени, функция Насыпей/Выемок создает растр, содержащий регионы, где добавлено вещество, регионы, где удалено вещество, и области, где за период времени не произошло изменений. Отрицательные значения объема указывают области, где количество вещества выросло, а положительные значения указывают области, где оно убыло.

Взяв морфологию речной сети в качестве примера, чтобы проследить величину и местоположение участков эрозии и отложений в речной долине, необходимо наметить серии профилей через долину и регулярно проводить наблюдение для определения областей эрозии и отложений.

Функция Насыпей/Выемок выполняет эти наблюдения для вас, определяя области, изменившиеся за счет размывания или, наоборот, отложений, а также области без изменений. При этом также вычисляется объем вещества, который добавлен или изъят из каждой области.

На рисунках ниже показано как функция Насыпей/Выемок использует поверхности “До” и “После” для определения областей с насыпями, выемками и неизменившихся в результате извержения вулкана Св. Елены в Национальном парке Пиночета. Области с насыпями отображены зеленым, области с выемками- красным, а неизменившиеся участки показаны на диаграмме желтым цветом.

При помощи функции Насыпей/Выемок:

- Определить области эрозии и отложений в речной долине.
- Вычислить объемы и площади вещества поверхности, которые необходимо изъять или насыпать до определенного уровня строительного участка.
- Определить области, которые часто засыпает во время оползней при поиске безопасных участков для строительства домов.

Слайд 20. Буферы.

Еще одним распространенным методом переклассификации данных является процесс построения буферов. **Буфер (buffer) - это полигон, с границей на определенном удалении от точки, линии или границы области.** Мы можем легко построить буфер вокруг объектов разного геометрического типа. Для создания такого буфера требуется всего лишь отсчет заданного расстояния по всем направлениям от каждой точки границы выбранного объекта. Нам нужно знать только расстояние по границы буфера.

Поскольку буфер связан с положением, формой и ориентацией объекта, мы можем легко отнести буферизацию к методам переклассификации на основе положения.

Буфер, обычно создается как отдельный объект и хранится в отдельном классе векторных объектов. Мы знаем, как ГИС выполняет измерение расстояний в растре и в векторах; в действительности, создание буфера — всего лишь расширение этой процедуры. Но поскольку эта процедура весьма полезна и часто применяется, большинство ГИС имеют специальные команды для построения буферов.

Слайд 21. Буферы (продолжение)

При анализе данных вам может быть понадобится построить второй буфер вокруг первого, третий - вокруг второго и т.д., которые вместе называются **многослойным буфером**. Процедура его построения относительно проста, так как каждый новый слой буфера - всего лишь новый буфер вокруг предыдущего слоя.

Так какого же размера должен быть буфер? Однозначного ответа на этот вопрос нет. Некоторые буферы показывают, что вокруг объекта, на неизвестное, или даже не могущее быть известным, расстояние простирается регион, который требует защиты, исследования, охраны или иного особого обращения. Такой сценарий не так уж необычен, как можно подумать. Многие буферные зоны в реальном мире так же произвольны, как и те, что мы устанавливаем в наших ГИС. Строители обычно сами создают буфер вокруг стройплощадки, чтобы защитить прохожих от тяжелых машин и падающего строительного мусора. Границы областей, загрязненных ядовитыми газами, радиоактивными материалами, разливами опасных жидкостей обычно устанавливаются правительственными агентствами или правоохранительными органами. Но довольно часто эти зоны устанавливаются лишь предположительно, это **произвольные буферы**. Чаще всего предположения строятся на интуиции или дилетантской информации из неизвестных источников. Размеры этих буферов, как правило, больше, чем это необходимо. Лишняя площадь буфера, часто так мешающая населению, обычно добавляется к произвольному буферу для безопасности.

Размеры буфера могут также основываться на любой процедуре измерения или переклассификации. Например, мы могли бы создать буфер, основанный на функциональном, а не евклидовом расстоянии от объекта. Это был бы **мотивированный буфер**, основанный на априорном знании площади буфера. Допустим, что мы создаем буфер вдоль реки, чтобы показать возможность загрязнения почвы по обеим ее сторонам. И мы знаем, что с одной стороны реки почва глинистая, в то время как на другой - песчаная. Поскольку загрязняющие вещества проникают через песок быстрее, чем через глину, буфер должен строиться на основе фрикционных свойств глинистой почвы. В результате буфер будет менее широким со стороны глины, нежели со стороны песка, отражая различия в проницаемости почв разных типов. Использование фрикционных поверхностей и барьеров — обычная практика при построении буферов, так как они дают некоторое основание для выбора размера буфера. Однако, поскольку точное определение величины фрикционного или барьерного импеданса часто затруднительно, буфер, построенный на основе этой величины, может оказаться не

более полезным, чем произвольный буфер, построенный из простых соображений.

Размер буфера может быть выбран по результатам изменения какого-либо явления, или, например, на мерах взаимной видимости. В этом случае мы получим **измеримый буфер**.

Существует еще четвертый вид буферов — **нормативный, когда буферизация определяется нормативными актами**. Например, нам говорят, насколько близко к пожарному гидранту мы можем парковать машину, и какая часть палисадника в действительности принадлежит местному сообществу. Строительные нормы указывают расстояния вокруг объектов коммунальных служб и между зданиями; природоохранные организации создают защитные полосы; вдоль железных дорог и линий электропередачи по закону устанавливаются буферы отчуждения, и т.д. В каждом случае есть нормативное основание для создания буфера определенного размера.

Независимо от типа буфера (произвольный, мотивированный, измеримый или нормативный) всегда есть вероятность того, что буфер не будет иметь одинаковую ширину вдоль всего линейного объекта или со всех сторон полигона. Такие различия, проиллюстрированные нашим примером буфера разной ширины вдоль реки в зависимости от типа почвы, создают класс буферов, называемых **варьируемыми**. Варьируемый буфер может определяться импедансом, барьерами или любой другой функцией окрестности.

Тема №14. Пространственный анализ: операции наложения.

Слайд 2. Некомпьютерный метод наложения.

К наиболее мощным возможностям современных ГИС относится их способность комбинировать картографическое представление одного выбранного набора данных с другим (или другими). Этот процесс называется **наложением** (overlay), или оверлейными операциями. Интуитивно он очевиден, и его применяли задолго до появления современных компьютерных геоинформационных систем.

Среди **наиболее влиятельных** разработчиков этого направления был Ян **МакХарг** (Ian McHarg), работа которого была связана с окружающей средой. Он использовал некомпьютерный метод наложения. Для изготовления своих карты он брал прозрачную пластиковую пленку [McHarg, 1971], на отдельных листах которой отображал разные параметры окружающей среды. Чем темнее участок пленки, тем выше чувствительность среды к воздействию негативного фактора. При наложении листов друг на друга чувствительность по разным параметрам складывалась. В результате создавалось новое суммарное покрытие, которое характеризовало чувствительность среды в целом, и могло быть использовано для принятия решений по ее рациональному использованию.

Мы далеко не всё еще знаем о причинных связях пространственных феноменов. Эта тема является значительной частью того, чем занимаются географы и другие исследователи. Геоинформационные системы дают теперь возможность легко выполнять процедуры наложения, благодаря чему могут возникнуть новые гипотезы, теории и даже законы об этих корреляциях.

Многие специалисты из разных областей могут получить ценные сведения о пространственных корреляциях, которые ранее не наблюдались или даже не могли наблюдаться без применения компьютеров. По мере развития этого процесса мы будем получать все более детализированный набор правил о допустимости наложения покрытий и выводе утверждения о причинно-следственных связях. А пока нам следует быть внимательными, ибо современные ГИС позволяют использовать не один единственный простой метод наложения, а десятки. Эти развитые возможности наложения легко могут привести к большому числу ошибок при сравнении наборов данных в действительности не связанных друг с другом, т.к. пространственные корреляции могут возникнуть чисто случайно.

Слайд 3. Растровое наложение

Процесс наложения на растровом типе данных подразумевает сравнение (сопоставление) ячеек двух или более растров, пространственное положение которых совпадает. Поэтому, пользователь должен помнить о том, что экстенд и пространственное разрешение всех входных растров должны быть одинаковыми. Если растры охватывают разные по размеру географические области, имеют разную географическую привязку или разное пространственное разрешение, то для их корректного совмещения необходимо предварительно привести их пространственные характеристики к одной и той же величине.

Традиционно наложение рассматривается как метод сравнения полигональных покрытий. Но существуют и другие типы наложений, использующие точечные и линейные данные. Различают следующие типы наложений: «точка в полигоне», «линия в полигоне», «полигон в полигоне».

Ячейки растра, представляющие точечную информацию (точки отбора проб, колодцы, скважины и др.) могут сравниваться с полигонами с помощью присваивания этим точкам, как и полигонам, легко отличимых чисел или категорий. При сопоставлении ячеек двух растров становится сразу же очевидно, какие ячейки точечного набора данных расположены в пределах интересующих нас полигонов.

Алгоритмы сопоставления ячеек, используемые компьютером, основаны на матричной алгебре. Обычно эти операции называют картографической алгеброй или алгеброй карт (*map algebra*). Используя язык *Алгебры карт* для выполнения математических вычислений над растрами с помощью *операторов* и *функций*, мы можем создавать различные варианты наложений. *Математические операторы* применяют математическую операцию к значениям двух или более входных растров. В **Калькуляторе растров модуля Spatial Analyst** есть три группы математических операторов: *Арифметические*, *Булевы* и *операторы отношений*. Рассматривая материал предыдущей темы, мы уже видели, как можно использовать операторы отношений для решения задач переклассификации растров.

Рассмотрим теперь, как *Арифметические* и *Булевы операторы* выполняют аналитические операции наложения.

Например, у нас есть одно покрытие с полигонами, представленным числовыми значениями 1, 2 и 3. Окружающий фон имеет значение 0. В другом покрытии ячейки со значением 10 и 20 представляют положения точечных объектов, остальные ячейки содержат значение 0. При наложении этих двух покрытий с помощью простой операции сложения

мы получим ячейки со значением 11,12,13,21,22,23, которые дадут нам точную информацию о положении точек со значения 10 и 20 относительно исследуемых полигонов.

Процесс растрового наложения «линия в полигоне» и «полигон в полигоне» по сути так же прост, как и наложение точек, поскольку в растре линии и полигоны являются группами ячеек с одинаковыми наборами числовых значений.

Слайд 4. Растровое наложение (продолжение).

Булевы операторы, в отличие от Арифметических, можно применить к данным, измеренным и в количественной, и номинальной шкале. Булевы операторы применяют Булеву логику попарно к входным растрам, возвращая значения TRUE (ИСТИННО) или FALSE (ЛОЖНО). На выходе значению TRUE соответствует 1, FALSE - 0.

Булевы операторы: And – «И»,

Or - «ИЛИ»,

Xor - «Исключающее ИЛИ»,

Not - «ДОПОЛНЕНИЕ выражения».

На слайде показан пример применения оператора **And** для двух входных наборов данных - *[Inlayer1]* и *[Inlayer2]*. Данная операция находит ячейки, в которых в обоих входных растрах стоит не 0. Другие операторы могли бы работать так:

Or (|)-находит ячейки, в которых в одном из входных растрах, или в обоих, стоит не 0;

Xor (!)-находит ячейки, в которых в одном из входных растрах, но не в обоих, стоит не 0;

Not (^)- находит в одном входном растре ячейки, в которых не присутствует ненулевое значение.

Мы можем применить Булевы операторы и для данных, измеренных в номинальных шкалах. Например, имея растр литологических разностей горных пород, подстилающих почвенный горизонт и растр, показывающий степень увлажненности почв, можем найти территории с сухими песчанистыми почвами.

Слайд 5. Растровое наложение: Статистика по ячейкам

Функция *Статистика по ячейкам* - это локальная функция, где значение каждой ячейки выходного растра является функцией от значений ячеек с тем же местоположением из входных растров.

При вычислении статистики по ячейкам вы можете вычислить для каждой ячейки выходного растра статистическую величину на основании значений ячеек с таким же

местоположением во всех входных растрах.

Следующие функции могут быть вычислены для каждого положения ячейки во входных растрах и записаны в соответствующую ячейку выходного растра.

Большинство: определяет для каждого положения ячейки наиболее часто встречающееся значение во входных растрах.

Максимум: определяет для каждого положения максимальное значение из входных растров.

Среднее: определяет для каждого положения среднее значение из входных растров.

Медиана: определяет для каждого положения медиану значений из входных растров.

Минимум: определяет для каждого положения минимальное значение из входных растров.

Меньшинство: определяет для каждого положения наиболее редкое значение во входных растрах.

Диапазон: определяет для каждого положения диапазон значений во входных растрах.

Стандартное (ср.кв.) отклонение: определяет для каждого положения стандартное отклонение значений из входных растров.

Сумма: определяет для каждого положения сумму значений из входных растров.

Разнообразие: определяет для каждого положения количество уникальных значений во входных растрах.

Используйте эту функцию, когда вам нужно вычислить статистические данные по нескольким растрам. Например, проанализировать развитие определенного явления во времени, например, среднюю урожайность за 10-летний период или разницу температур в разные годы.

На рисунке внизу вычислено разнообразие типов землепользования для каждой ячейки за несколько лет, что позволяет найти участки, где разнообразие больше единицы (участки, закрашенные серым). Это показывает районы, где тип землепользования изменился за указанный период времени, в данном случае отражая рост застройки прилегающих к городу территорий (показанной красным).

Слайд 6. Растровое наложение: Зональная статистика.

С помощью функции *Зональной статистики* можно вычислить статистические данные по каждой зоне в наборе данных по зонам на основании значений в другом наборе данных.

Зона - это все ячейки растра с одинаковым значением, независимо от их

пространственной связанности. Однако, в качестве “набора данных зон” можно использовать как растровые, так и векторные данные. Так, например, участки с жилой застройкой - это зона растрового набора данных землепользования, а векторный набор данных дорог может представлять зоны для статистики ДТП.

Функции зональной статистики работают позонно; одно выходное значение вычисляется для всей зоны в растровом наборе данных.

По каждой зоне можно вычислить следующие статистические значения.

Большинство: определяет для каждого местоположения ячейки наиболее часто встречающееся значение в зоне.

Максимум: определяет для каждого местоположения максимальное значение в зоне.

Среднее: определяет для каждого местоположения среднее значение в зоне.

Медиана: определяет для каждого местоположения медиану значений в зоне.

Минимум: определяет для каждого местоположения минимальное значение в зоне.

Меньшинство: определяет для каждого местоположения наиболее редкое значение в зоне.

Диапазон: определяет для каждого местоположения диапазон значений в зоне.

Стандартное (ср.кв.) отклонение: определяет для каждого местоположения стандартное отклонение значений в зоне.

Сумма: определяет для каждого местоположения сумму значений в зоне.

Разнообразие: определяет для каждого местоположения количество уникальных значений в зоне.

Слайд 7. Зональная статистика (продолжение).

Используя функцию Зональной статистики, вы можете вычислить среднюю высоту над уровнем моря для каждой лесной зоны или количество дорожно-транспортных происшествий на каждой улице города. На этом слайде представлен пример использования функции Зональной статистики для определения числа типов растительности, характерных для каждой высотной зоны. Здесь в качестве входных данных взяты переклассифицированный растр высот и растр «Виды растительности». Рассчитаем сколько видов растительности размещается в пределах каждого интервала высот. В качестве расчетного статистического параметра более всего подходит «разнообразие». Выходными данными функции Зональной статистики является таблица, в которой записывается номер или код зоны (поле **Value**) и значение статистического параметра (в данном случае поле **Variety** – «разнообразие»). Для наглядности представления результатов табличные данные можно показать на диаграмме. На

выходной диаграмме и в таблице показано разнообразие видов растительности для каждой высотной зоны. Наибольшее многообразие растительности наблюдается в зонах в районе 2500 метров.

Завершая рассмотрение способов наложения на растровом типе данных, следует отметить следующее. Растровым наложениям свойствен недостаток пространственной точности, но при этом они обладают высокой гибкостью и скоростью выполнения операций. Общеизвестно, что в общем случае растровое наложение более предпочтительно вследствие его вычислительной легкости: поскольку каждая ячейка раstra одного покрытия обязательно совмещена с такой же ячейкой в других покрытиях, компьютеру не приходится тратить ресурсы на вычисление координатных взаимоотношений объектов разных покрытий. Вместо этого все ресурсы тратятся на сравнение атрибутивных данных, что, конечно, повышает скорость выполнения данной операции.

Слайд 8. Векторное наложение: типы наложений.

Операции векторного наложения дают те же преимущества, что и основанная на векторах компьютерная картография - они создают картографические продукты, напоминающие традиционные, рисованные от руки, карты гораздо больше, чем созданные из растровых карт.

Наложение САПР. Первый и простейший метод компьютерного векторного наложения очень похож на ручной метод в том, что мы просто располагаем символы отображения данных на одной поверхности. Результатом этой операции является изображение на экране, но не класс пространственных объектов в файле. Программа не отвечает за объединение атрибутов объектов. Нет также и топологии. В данном случае наложение всего лишь графический прием последовательной отрисовки отдельных слоев для создания объединенной картинки. Оно может быть напечатано и сохранено как файл на диске компьютера, но как объект аналитических функций ГИС использоваться быть не может. Этот момент довольно важен, и даже опытные пользователи ГИС иногда в нем путаются.

Давайте рассмотрим пример с использованием системы автоматизированного проектирования (САПР). Допустим, мы оцифровали карту дорог, создав два слоя: собственно дороги и гидрография. Далее мы отобразили их на экране, чтобы определить, как проехать к некоторому озеру. Это вполне нам удастся визуально, и мы сохраняем карту для других случаев. Позднее другой человек пытается использовать эту карту для поиска дорог, идущих вдоль береговой линии на расстоянии до ста метров от нее. Хотя визуально

это довольно просто сделать, сама программа не сможет нам ответить на такой простой вопрос, так как она не знает, как пространственно объекты карты связаны друг с другом. У нее нет БД, а надписи являются такими же графическими элементами, помещаемыми рядом с объектами, как и сами эти объекты. Она не может также определять меру близости или наличие соединения.

Наложение САПР очень полезно в своих аналитических пределах. Не всегда нужно выполнять сложные в вычислительном отношении запросы для определения соседства или пространственного сходства между объектами. Это легко можно сделать визуально. Преимущества наложения данного типа - простота и быстрота выполнения. Многие системы управления инженерными сетями и промышленными объектами действуют на такой же основе, и обеспечивают пользователей тем, что им нужно. То есть, они достаточно быстро создают карты нужных объектов, которые можно взять с собой на место работ, где графика может быть прочитана и интерпретирована вручную.

Слайд 9. Топологическое векторное наложение.

Топологическое векторное наложение подразумевает, что программа должна уметь определять положения точек или линий относительно границ полигонов, с которыми мы их сравниваем. При этом мы также будем подразумевать, что наборы векторных данных имеют общую систему координат.

Наложение «точка в полигоне» заключается в соотнесении координат точек с границей полигона с целью определения принадлежности этих точек внутренней области полигона. Рисунок иллюстрирует определение принадлежности точки с координатами (2.3, 3.4) изображенному полигону.

Такое наложение может выполняться с целью создания нового покрытия, состоящего только из тех полигонов, которые содержат указанные точки. Или другая задача – выбрать только те точки, что принадлежат указанным полигонам. Или- рассчитать сколько точек расположено в пределах каждого полигона.

Наложение "линия в полигоне" заключается в соотнесении координат концевых и промежуточных точек линии с границей полигона с целью определения принадлежности этих точек полигону, то есть, по сути, оно сводится к выполнению точечного наложения. Дополнительным моментом является то, что линия может пересекать границу полигона. В простейшем случае можно считать, что если хотя бы одна точка линии принадлежит полигону, то и вся линия принадлежит ему. Но более корректным подходом является определение точек пересечения линии с границей

полигона и создание в них узлов, что позволит разделить атрибуты внутренних и внешних по отношению к полигону частей линии. Например, если лесозащитная полоса как линейный объект пересекает границу между пашней и лугом, мы сможем сказать, какая часть полосы принадлежит каждой из областей, а также создать таблицу, показывающую это отношение.

Как видите, векторные наложения "точка в полигоне" и "линия в полигоне" являются вопросом не только графического отношения объектов, но и отношения атрибутов. В конце концов, объекты карты представляют часть реального мира.

Для векторного наложения полигонов, подобно случаю «линия в полигоне», программа должна определить точки пересечения границ полигонов одного покрытия с границами полигонов другого покрытия. Эти точки пересечения становятся узлами. Далее программа разделяет или объединяет полигоны, в зависимости от поставленной задачи. Затем программа отслеживает передачу атрибутов в новое покрытие.

Слайд 10. Векторное наложение «полигон в полигоне».

Поскольку векторные ГИС связаны с СУБД, не удивительно, что булева логика, используемая в запросах к БД, используется также и для пространственных запросов. Булево наложения является широко распространенным подходом. Его легко понять, особенно если вы уже знакомы с операциями над множествами. Эти операции часто иллюстрируются так называемыми диаграммами Венна (рисунок слева).

В случае Булева наложения мы сравниваем не сами атрибуты, а пространство, занимаемое каждым из наборов атрибутов. Допустим у нас есть простое покрытие, имеющее только два типа полигонов землепользования - сельское и городское, плюс к этому другое покрытие, показывающее, находится ли земля в собственности или арендуется. Когда нам нужно найти находящиеся в собственности сельские участки, мы выполняем операцию наложения с пересечением, которая позволяет нам найти такие участки земли. Программа геометрически определяет пересечения полигонов, далее передает атрибуты входных полигонов вновь созданным полигонам. После чего выбираются только те полигоны, которые удовлетворяют заданному критерию по атрибутам. Другими словами, выполнение операции пересечения над двумя покрытиями равносильно созданию диаграммы Венна для двух областей с разными наборами атрибутов. В результате мы можем определить собственников сельских участков.

Слайд 11. Пример векторного наложения полигонов.

На слайде представлен пример векторного наложения «полигон в полигоне», выполненный для нескольких наборов векторных данных. В иллюстрируемом примере планировщик регионального уровня должен подготовить план контролируемого роста городского населения. При разработке плана рост может допускаться в те области, где имеются почвы, годные для строительства домов с фундаментом. При этом следует везде сохранять почвы высшего качества для ведения сельского хозяйства. Кроме того, следует предотвратить застройку земель, находящихся в собственности федерального правительства, земель, которые уже используются в качестве сельхозугодий, земель, на которых производятся археологические раскопки, а также земель, где находятся области обитания охраняемых видов животных. В результате наложения остается незаполненной та часть карты, на которой можно вести строительство.

Слайд 12. Методы наложения полигонов.

Как уже упоминалось, исторически сложилось так, что сравнение полигональных покрытий является наиболее распространенным подходом к выполнению наложения, в следствие чего разработчики геоинформационных систем изначально развивали именно этот тип наложения. Поэтому существуют различные методы к выполнению наложения полигонов, ориентированные на определенные потребности пользователей. Рассмотрим способы выполнения наложения полигонов, реализуемые системой ArcGIS.

Разбиение (Split). Разбиение создает новые классы пространственных объектов из объектов входного класса. При этом операция Разбиение выполняет серию действий вырезания геометрии объектов входного класса по форме каждого объекта класса разбиения. В результате выходные классы объектов содержит только те части входных объектов, которые перекрыты определенными полигонами объекта класса разбиения. Входные Характеристики могут быть полигонами, линиями, или точками. Класс разбиения может быть только полигональным. Выходные классы объектов имеют тот же набор атрибутов, что и входные. Для них так же создается топология.

Обновление. Наложение с обновлением действует следующим образом. В наложении участвуют 2 класса: входные данные и обновляющие данные. Если объект или часть объекта входного класса не перекрывается объектами класса обновления, то они передаются в выходной класс без изменений. Если объект или часть объекта входного

класса перекрывается объектами класса обновления, то их геометрии и атрибуты обновляются, становясь идентичными объектам класса обновления.

Идентичность. Операция Идентичность вычисляет геометрическое пересечение входных объектов и объектов класса коррекции. Атрибуты и геометрия объектов входного класса будут скорректированы объектами класса коррекции. Результаты записываются в выходной класс объектов.

Слайд 13. Методы наложения полигонов (продолжение).

Стирание. Наложение со Стиранием создает класс объектов, состоящих из объектов и/или их частей, находящихся за пределами объектов стирающего класса.

Симметричная разность. Симметричная разность аналогична Булеву дополнению: рассчитывается геометрическое пересечение входных данных, в выходной класс сохраняются непересекающиеся объекты и части объектов входных наборов данных.

Объединение. Объединение рассчитывает геометрическое пересечение объектов входных классов, все объекты сохраняются в выходном наборе данных с атрибутами всех входных классов.

В случае пересекающихся объектов, область перекрытия будет продублирована (рисунок на слайде внизу слева), чтобы поддерживать все атрибуты и области. В результате наложения создается 2 полигона с идентичной геометрией. У одного атрибуты- A=2, B="X", у другого - A=1, B="X".

Если между объектами есть промежутки (рисунок на слайде внизу справа), то по выбору пользователя, объекты на месте промежутка могут создаваться либо не создаваться. В случае заполнения промежутка, полигон, заполняющий промежуток получает специфический атрибут FID=-1.

Операция Объединение может работать с единственным входным классом. В этом случае, вместо обнаружения перекрытия между объектами нескольких классов, Объединение обнаружит перекрытие между объектами в пределах единственного класса. Области, где будет обнаружено перекрытие, получают атрибуты перекрывающихся объектов.

Слайд 14. Методы наложения полигонов: вырезание.

Наложение с вырезанием используется для создания нового класса, который представляет собой географическое подмножество объектов в другом, большем классе

объектов. Операция Clip вырезает объекты одного класса по форме объектов другого класса, подобно тому, как форма для печенья режет слой теста.

Входные данные могут содержать точки, линии, или полигоны. Режущий класс должен быть полигональным классом.

Атрибуты объектов в выходном классе те же самые, что и во входном (вырезаемом) классе объектов. В отличие от операций Пересечение (Intersect) или Объединение (Union), объединения атрибутов входных наборов данных не происходит.

Слайд 15. Методы наложения полигонов: Пересечение (Intersect).

Операция Пересечение аналогична Булеву пересечению. **Пересечение** рассчитывает геометрическое пересечение объектов входных наборов данных. В выходной класс записываются общие объекты или общие части объектов всех входных наборов данных. Объекты выходного класса получают полный набор атрибутов входных классов.

Выходной тип геометрии может быть тем же самым типом геометрии или геометрией более низкого уровня. Так пересечением полигонов может быть полигон, может быть линии или точка.

Слайд 16. и Слайд 17. Операция Пересечение (продолжение).

В качестве входных наборов данных могут быть использованы не только полигональные классы, но и любая комбинация типов геометрии (точка, мультиточка, линия, полигон). Соответственно, на выходе мы можем получить пересечения различных типов геометрии.

Функция Пересекать может также работать с единственным набором данных. В этом случае, вместо обнаружения пересечений между наборами данных, Пересечение обнаружит пересечения между объектами в пределах единственного класса. Это может быть полезным, для того чтобы обнаружить перекрытия полигонов или пересечения линии.

Слайд 18. О проблемах векторного наложения.

Результат векторного наложения может отличаться от растрового, и выглядеть несколько иначе. В векторной системе могут неожиданно оказаться десятки и даже сотни мелких полигонов, особенно вдоль границ пересекающихся полигонов. Эти визуально незначительные расхождения могут существенно влиять на результаты анализа.

..... В качестве примера рассмотрим сравнение двух равномасштабных покрытия

землепользования, созданных на основе двух аэрофотоснимков одного участка земной поверхности в моменты времени T_1 и T_2 . При сравнении двух внешне одинаковых покрытий оказывается, что они совпадают не точно, что приводит к образованию осколочных полигонов вдоль границ накладываемых полигонов. Перед нами встает вопрос: вызвано ли это реальными изменениями в промежуток между двумя моментами съемки или причина кроется в погрешностях съемки и/или оцифровки? Интуиция подсказывает, что столь малые изменения землепользования маловероятны, особенно учитывая то, что сами полигоны выглядят столь похожими, а даже небольшие изменения в самом землепользовании дали бы гораздо более значительные изменения картины.

..... Игнорируя микроскопические изменения в землепользовании, пытаемся разобраться в причинах несовпадения полигонов. Сами полигоны могут быть одинаковой формы, но при этом чуть-чуть повернутыми. Мы можем попытаться устранить этот поворот, однако даже при самом тщательном исполнении редко удается достичь полного совпадения. К тому же, такой поворот зачастую может привести к рассогласованию других полигонов. Короче говоря, здесь нет простого решения, и некоторые пользователи даже прибегают к растровым ГИС, жертвуя точностью, чтобы избежать частых утомительных подгонов изображений. Как вы можете догадаться, данная проблема обостряется, если мы пытаемся сравнивать разнородные покрытия. Например, наложение покрытий почв и растительности вообще не позволяет предполагать совпадение каких-либо двух полигонов. Целью наложения может быть именно обнаружение расхождений, и мы сталкиваемся с необходимостью отделения реальных различий полигонов от ошибок.

..... К сожалению, несмотря на продолжающиеся исследования по определению величины погрешности образуемых в результате наложения покрытий, есть очень мало общих принципов и еще меньше ответов на этот вопрос, особенно если покрытия приходят из разных источников. Чаще всего вам придется полагаться на собственный опыт и интуицию. Ваши решения должны основываться на вашем собственном знании данных: их качества, качества их ввода и даже о полевой работе, давшей их. Ваши действия в значительной степени зависят от того, насколько точны должны быть результаты для решаемой задачи. Логично предположить, что если вы имеете несколько покрытий, то результат наложения будет иметь погрешность наихудшего из них, хотя на самом деле это не обязательно так, о чем тоже следует помнить. Поэтому, перед выполнением наложения следует хорошенько изучить данные, чтобы понять, какая величина погрешности может быть приемлема.

Для того, чтобы избежать проблемы возникновения осколочных полигонов в системе ArcGIS при определении параметров операций наложения задается параметр «Кластерный допуск».

Кластерный допуск- это расстояния, в пределах которого вершины и ребра объектов считаются идентичными (или совпадающими).

Если вершины находятся на расстоянии, которое меньше заданного, то они будут считаться совпадающими. И при построении топологии будут совмещаться в одну. Обычно значение величины кластерного допуска очень мало, чтобы минимизировать перемещение вершин правильно размещенных объектов.

Величина кластерного допуска оценивается исходя из точности входных данных. Для того, чтобы минимизировать ошибку, кластерный допуск, которые вы выбираете, должен быть как можно меньше. Например, если точность ваших данных находится в пределах 10 метров, кластерный допуск должен быть не больший 10 метров, а по возможности меньше. Обычно в качестве величины кластерного допуска выбирают величину на порядок ниже точности входных данных. Таким образом минимизируется сумма перемещений в данных при создании топологии, которые порождены кластерным допуском.

Следует обратить внимание, что результат наложения при различной величине кластерного допуска может получиться разным. Увеличение кластерного допуска может привести как улучшению результата, так и к нежелательным результатам.

Тема 15. Вывод в ГИС.

Слайд 2. Вывод: отображение результатов анализа.

Вывод результатов — конечный продукт любого анализа. И если он не понятен пользователю, то мы не справились с нашей задачей. В конце концов, целью является не просто анализ сам по себе, но и представление результатов. Здесь мы рассмотрим как способы вывода ГИС, так и некоторые требования дизайна для создания качественного, понятного выходного продукта. Обе стороны важны, так как техника физически ограничивает наши возможности создания изображений, а пользователи имеют физические и психологические ограничения и особенности восприятия, влияющие на интерпретацию результатов.

Вопросам дизайна вывода ГИС уделено относительно мало места в литературе, хотя в области традиционной картографии — предостаточно. Именно в подсистеме вывода ГИС наиболее близка традиционной картографии. Образно говоря, корни ГИС уходят в картографию - обе включают в себя карты и атрибуты, обе используют географические данные, включая масштаб, проекции, координатные системы.

При разработке макета карты вам также может помочь литература по графическому дизайну, которой тоже немало. Ценность этих знаний обусловлена тем, что качественный вывод делает пользователей довольными, а довольный пользователь это клиент, который к вам еще вернется.

Вывод результатов анализа может быть постоянным или временным, в зависимости от типа выходного устройства. К первой категории мы относим вывод на бумагу, пленку, компакт-диски или другие носители - все они могут хранить результат долгое время. Вторая категория – временный вывод, обычно выводит данные на экран монитора или проекционный экран, с целью демонстрации результатов анализа или предварительного просмотра файлов при решении об использовании их в анализе или о постоянном выводе.

Вывод может быть также разделен на человеко- и машинно ориентированный. Машинно-ориентированный вывод чаще всего используется для сохранения материала на компьютерных носителях информации. Человеко-ориентированный вывод предназначен для восприятия людьми. Машинно-ориентированные формы требуют решений о структуре данных, носителе информации и совместимости с другими программами и компьютерными системами, принять которые проще, чем решения о человеко-ориентированном выводе, поскольку люди больше различаются по опыту и уровню понимания графических средств коммуникации. По этой причине мы уделим больше внимания именно человеко-ориентированному выводу.

Слайд 3. Картографический вывод: назначение карт.

Традиционным является картографический вывод. Карты всё еще остаются наиболее компактным способом представления географической информации. Традиционно карты создавались для осуществления двух основных функций. Первая функция - это хранение информации. Создание карты - это способ зафиксировать информацию о местоположении пространственного объекта. Вторая функция состоит в том, чтобы через изображение донести пространственную информацию до пользователя.

Соответственно, различают карты двух типов. Первый - это основные справочные карты, например Геологическая карта Российской Федерации или карта города. Карты этого типа содержат информацию о взаимном расположении объектов. Второй тип - тематические карты, используемые для передачи информации по определенной одной или нескольким тематикам, например землепользование, уровень добычи полезного ископаемого или статистика эффективности нового способа бурения. Большинство карт на выходе ГИС относятся к тематическим картам, выделяющим структурные отношения в рамках выбранной темы. В случае ГИС слово "тема" можно заменить на "решение", так как такие карты чаще всего являются результатом решения определенной задачи или ключом к завершению процесса принятия решения.

Существует множество карт, каждая из которых может нести как базовую, так и специализированную информацию. Если для создания карты, отвечающей каким-либо специальным требованиям, обычно требуется опытный картограф, то любой желающий может сделать хорошую базовую и информативную карту, учитывая следующие простые принципы. Эти принципы организованы в 6 разделов, и вы можете руководствоваться ими для создания и совершенствования ваших карт.

Слайд 4. Картографический вывод: назначение карты.

1 раздел. Назначение. Как правило, у карты какое-либо одно назначение. Попытки соединить очень много информации в одной карте (то есть сделать многоцелевую карту) приводят к нечеткости восприятия и сбивают с толку пользователей карт. Лучше использовать несколько карт, каждая из которых выполняет определенную задачу.

Назначение карты принципиально для ее дизайна. При создании карты картограф должен знать ответы на некоторые фундаментальные вопросы, например: Что будет нанесено

на карту? Для кого она предназначена? Как она будет представлена: как самостоятельное издание, или как часть отчета? Какие технические средства будут использованы для отображения карты?

Каждый помещаемый на карту объект должен что-то говорить на тему этой карты. На карте не должно быть ничего лишнего. Следовательно, первое правило при составлении карты гласит, что она должна быть читаема, анализируема и интерпретируема.

2 раздел. Размер, масштаб и средства передачи также имеют большое значение. Физический размер карты, зависящий от географического экстенда, показанного на карте, будет предписывать масштаб карты, и определять, каким образом будет передаваться фактический размер и количество объектов на карте. Этот вопрос является актуальным в том случае, когда масштаб входных данных и масштаб выходного изображения существенно различаются. Например, если слой дорог был создан в масштабе 1:24 000, то при нанесении его на карту с масштабом 1:2 000 000 получится, что дороги на ней будут значительно более детализованы, чем требуется. Значит, вам понадобится сократить количество объектов-дорог на вашей карте.

При уменьшении масштаба вывода, когда проблема избыточности информации не может быть решена простым удалением объектов (все объекты должны быть отображены), возможно, вам потребуется выполнить генерализацию входных данных. Генерализация – это упрощение (сглаживание) формы линейных и полигональных объектов. Генерализация может быть выполнена путем сглаживания линий и границ полигонов (в различной степени и различными методами), либо путем изменения мерности объектов. Например, узкий вытянутый полигон при уменьшении масштаба может быть представлен линией, а изометричный- точкой.

Дизайн карты может быть очень сложным, особенно при широких возможностях современных компьютерных программ. Однако логические (интеллектуальные) и графические (визуальные) цели картографического дизайна часто конфликтуют. Разрешение этих проблем почти всегда происходит путем компромисса. Например, если вы располагаете символы домов на карте, то и логика, и эстетика диктуют согласованный подход, скажем, расположение символов в точных координатах домов. Но если там же имеется линейный символ, представляющий дорогу, проходящую очень близко от этих домов, то либо символ дороги, либо символы домов должны быть смещены. При некоторых обстоятельствах вам может потребоваться смещение как символов домов, так и символа дороги в сторону от их точных положений. Выбор,

показанный на рисунке в нижней части слайда отражает важнейшее положение о том, что графические объекты не абсолютно точно указывают реальное положение физических объектов, которые они представляют. Картографам приходилось решать огромное множество подобных проблем, поэтому они создали подробный набор соглашений и традиций, которые являются результатом проб и ошибок, а главное, — проверки среди пользователей карт. Эти результаты содержат самые эффективные методы достижения компромисса, они должны быть для вас руководством для ваших собственных картографических решений. Отклонений от установленных норм почти всегда приводит к созданию карты, менее эффективной в сообщении результатов вашего анализа, чем она могла бы быть.

Средства передачи также играют важную роль. Карты, напечатанные в газетах, не могут отображать детали с такой же четкостью, как карты, напечатанные на принтере с высоким качеством печати.

В свою очередь, детали на цифровой карте зависят от программы просмотра. Например, статистические карты, опубликованные в Интернете, несут значительно меньше информации, чем карты, созданные для программы типа ArcGIS.

Слайд 5. Принципы графического дизайна.

3 раздел. Следующим фактором графического дизайна является **целевая аудитория**. Будете ли вы делать карту для нескольких пользователей или для большой аудитории в сотни или даже миллионы человек. Учитывайте уровень подготовки аудитории к восприятию информации на вашей карте.

Многие пользователи выходных документов ГИС не имеют знаний и опыта в географии и картографии. В таких условиях карта должна восприниматься как можно легче, благодаря сохранению лишь наиболее важных объектов и названий и применению общеизвестных символов. В то же время, более опытные пользователи смогут извлечь дополнительную информацию из более плотного представления с более абстрактными символами. Значение может иметь также и возраст: например, пожилым людям и детям труднее воспринимать мелкие символы и текст.

4 раздел. Фокус. Другой аспект назначения карты, эргономическая задача имеет в виду не столько то, *что* отображается, сколько то, *как* это делается. Фокус подчеркивает то, что создатель карты считает нужным донести для пользователя в первую очередь.

Решив, что вы изобразите на карте, вам нужно также выбрать форму представления, которая сможет адекватно передать ваше сообщение. Если карта показывает опасные зоны, то ее дизайн должен недвусмысленно демонстрировать их важность: обычно холодные тона (голубой, зеленый и светло-серый) используются для фона, в то время как теплые тона (красный, желтый, черный) используются для того, чтобы акцентировать внимание на чем-либо. В противном случае приятная наружность карты может иметь некоторую художественную ценность, но посылать при этом зрителю не то сообщение, которое вы предполагаете.

5 раздел. Баланс карты. Компонуя карту, вам придется размещать ее элементы на ограниченной площади. Собственно карта должна быть доминирующим элементом на листе. Многие карты выглядят неинтересными вследствие того, на них слишком много или слишком мало белого фона. Карта, которая полностью состоит из основного изображения, менее желательна, чем карта с некоторым аморфным фоном, обособляющим основное изображение. Фон поднимает также контраст и визуальную привлекательность. Карта, на которой слишком много фона, уменьшает значение основного изображения и может вызвать у зрителя подозрение о том, что она не полна.

Используйте различные возможности размещения элементов на карте (например, не помещайте все заголовки наверху).

Слайд 6. Принципы графического дизайна: иерархическая организация

6 раздел. Последним принципом графического дизайна является **иерархическая организация**. Все графические элементы, присутствующие на карте, должны быть организованы таким образом, чтобы подчеркнуть то, что наиболее важно. Этот принцип слабо применяется на общегеографических картах, так как их цель - дать возможность различным по интересам пользователям фокусироваться на важных для них элементах, поэтому все элементы такой карты должны иметь равное значение. А тематические карты, наиболее распространенные в среде ГИС, должны подчеркивать конкретные объекты или результаты анализа. Это может быть сделано посредством иерархической организации, или разделения элементов по уровням визуальной значимости.

Существуют три основных метода достижения иерархической организованности. **Стереограммный** метод требует выбора и модификации графических приемов с тем, чтобы позволить наиболее значимым элементам выглядеть расположенными выше, чем менее важные элементы. Ключами к восприятию глубины могут быть использование трехмерных объектов, различия в толщине линий, цвете, яркости или размере.

Второй метод иерархической организации называется **расширительным**, он чаще всего используется для ранжирования линейных или точечных объектов. Например, главные дороги должны выглядеть более заметными, чем второстепенные. Здесь мы можем варьировать толщину линий, их яркость или внутреннюю структуру.

Последний метод создания иерархической графической организации, называемый **методом подразделительной иерархии**, применяется главным образом для показа различий во внутреннем устройстве областей. Так, например, удельный вес русского населения в этническом составе Южного Федерального округа может быть подразделен на высокий, средний и низкий. Различие между расширительным методом и методом подразделительной иерархии состоит в том, что первая больше соответствует порядковой шкале измерений, а вторая – номинальной.

Слайд 7. Процесс дизайна.

Составление карты представляет собой процесс, включающий ряд технологий, а также научный и художественный подходы, необходимые для создания географических карт. Картография развивалась и совершенствовалась задолго до появления компьютеров и ГИС и имеет устоявшиеся традиции. Например, теорию визуального разнообразия условных знаков Джека Бертина. В основе его теории лежит принцип визуального отличия одного символа от другого, что помогает в восприятии изображения. Совершенствуются также другие возможности картографии, учитывающие и человеческую психологию, и визуальное восприятие.

Первый шаг состоит в выборе типа карты, которую вы собираетесь создать, размещаемых на ней объектов и общего ее вида. Эта стадия интуитивна, результат ее – общий план вашей карты. Основные элементы карты представлены здесь на слайде:

- **Собственно карта** – основная отображаемая область. Вы можете показать более одного изображения на области отображения в вашем документе. Например, вы можете показать динамику какого-либо процесса, демонстрируя несколько карт, например карты численности населения за разные года. Ваша карта может также

содержать обзорную карту (мелкомасштабную карту, которая поможет пользователю понять, где расположена основная область интересов), карту-врезку (для передачи более детального изображения области, трудной для восприятия на основной карте) или индексную карту (как правило, показывающую последовательность карт, входящих в основную карту). Все это помогает донести информацию до пользователей. В ArcMap каждая из этих карт передается как отдельный фрейм данных.

- **Название** сообщает пользователю, что представлено на карте. Как правило, вводится в компоновку карты как текст.
- **Легенда** – список условных знаков, использованных на карте с описанием того, что они изображают. Создается или редактируется с использованием Мастера легенды в компоновке
- **Масштаб** можно отобразить числом (1:10 000), графическим элементом (масштабная линейка), или описанием (в одном сантиметре один километр).
- **Проекция.** На карте обязательно должна быть информация о проекции.
- **Направление** обычно отображается Стрелкой севера. На карте может показываться географический полюс или магнитный полюс. Этот элемент помещается в компоновку карты.
- **Источник данных** – библиографический список источников, использованных при создании карты.

Можно включить также множество других компонентов - даты, рисунки, координатные и градусные сетки, отчеты, таблицы, дополнительный текст, рамки и фамилии авторов.

Далее вы выбираете символы для отображения объектов, интервалы классов, цвета, типы линий и другие графические элементы. Это тоже можно сделать на эскизе.

При создании карты вы должны варьировать вид графических примитивов, представляющих точечные, линейные и площадные объекты, чтобы эти объекты были различимы.

Основные параметры, которые могут изменяться, это форма, размер, ориентация и цвет. Кроме этого, для заполнения площадных объектов могут использоваться **штриховки**, которые характеризуются **организацией** — регулярной или случайной, **частотой следования** элементов, позволяющей делать их светлее или темнее, и **ориентацией** этих элементов.

Все эти параметры могут изменяться для улучшения графического представления объектов и их групп. Поскольку карты воспринимаются как единое целое (в отличие текста, воспринимаемого

последовательно, слово за словом), необходимо также уделять внимание таким характеристикам дизайна, как **разборчивость, визуальный контраст, отношение основного изображения и фона и иерархическая структура.**

Графические символы должны быть прежде всего разборчивыми: отдельные линии - разделимыми, цвета - различимыми, формы - узнаваемыми. Классическим примером являются фигуры, используемые на дорожных знаках, - они позволяют передать сообщение без использования текста. Размеры символов должны учитывать расстояние, с которого карта рассматривается и которое может меняться от десятков сантиметров при индивидуальной работе до нескольких метров при демонстрации в коллективе. При этом нужно помнить и о физических ограничениях оборудования и человеческого глаза.

Слайд 8. Процесс дизайна (продолжение).

Заключительная стадия процесса дизайна состоит в точной настройке того, что было сделано на предыдущей стадии. Здесь в общий план карты могут вноситься только небольшие изменения. Прежде чем опубликовать карту, всегда имеет смысл дать кому-нибудь проверить ее, особенно грамматику и общий вид.

Главное - создание прототипа на экране монитора перед выводом на печать, так как принтеры и плоттеры намного медленнее экрана монитора. При этом следует учитывать возможные различия изображений на экране монитора и на бумаге, обусловленные использованием специальных языков управления периферийными устройствами. При выводе текста нередко получается так, что шрифты в компьютере и в принтере немного различаются, что приводит к наложению или сдвигу надписей, чего не наблюдалось на экране, или замене некоторых символов (особенно символов национального алфавита и специальных), коды которых различаются в компьютере и принтере. Эта проблема радикально решается установкой в драйвере периферийного устройства (например, принтера) флажка требования выводить всё, включая текст, в виде графики. При этом, возможно, некоторое замедление вывода, зато соответствие всегда будет полным.

Другая часто встречающаяся проблема – несоответствие цветов: заметно различающиеся на экране компьютера цвета могут оказаться практически одинаковыми на бумаге. Здесь можно порекомендовать вывести уменьшенную копию карты или ее часть, чтобы оценить соответствие и, если нужно, внести коррективы перед печатью всей карты. Более надежный путь - калибровка оборудования с использованием специальных калибровочных программ (обычно

используемых в издательском деле) для достижения максимального соответствия изображений на экране и бумаге. После калибровки учет всех, различий и их корректировка производится программой автоматически в процессе вывода.

Слайд 9. Нетрадиционный картографический вывод.

Традиционно, тематические карты создаются главным образом в их ортографической форме, когда зритель как бы смотрит на землю сверху. Такие карты рассчитаны на близкое к реальности воспроизведение форм и соотношений. Но существуют и другие картографические формы, имеющиеся во многих коммерческих ГИС. Хотя такие карты малоценны с точки зрения анализа, они очень эффективны в представлении результатов анализа, особенно представляемых в виде поверхностей. Поскольку современные программы позволяют визуально накладывать на эти трехмерные поверхности данные других покрытий, пользователь легко может идентифицировать связи между топографией и другими факторами.

Сегодня некоторые системы позволяют наблюдать карты этого и других типов не только статически, но и с эффектами **анимации**. Для двухмерных карт анимация позволяет наблюдать динамические системы так, как они действуют в реальном мире, или демонстрировать процессы, которые не воспринимались бы из-за их слишком медленного развития в реальности. При этом движение может быть замедлено, ускорено, приостановлено или обращено, что дает возможность лучшего понимания этих явлений. Сейчас существуют и системы с трехмерной анимацией, создающие впечатления пролета над местностью. Не следует преуменьшать мощь этих приемов, ибо зрение человека гораздо лучше замечает движущиеся объекты, чем неподвижные, что, безусловно, помогает выявлению взаимодействий объектов и их распределений.

В последние годы популярными стали изображения и анимации на глобусе, дающие полное и наглядное представление процессов и явлений планетарного масштаба или для того, чтобы показать, как выглядят объекты интереса на глобусе в разных масштабах.

Слайд 10 и 11. Нетрадиционный картографический вывод: картограммы.

Наконец, аналитик должен рассмотреть набор нетрадиционных картографических форм, которые несколько недооцениваются сообществом пользователей. Они называются **картограммами** и имеют внешний вид карт, но расположение объектов на них соответствует не реальному положению в пространстве, а значению некоторого показателя. Картограммы можно встретить повсюду, что говорит об их полезности.

Например, в автобусах и поездах, метро можно встретить схемы маршрутов. На них остановочные пункты располагаются на прямой линии, причем расстояние между ними также обычно не соответствует реальному. Эти условности упрощают карту до уровня схемы, оставляя только наиболее важные пространственные характеристики - состав и последовательность остановок и возможности пересадок. В дорожных атласах можно встретить **линейные маршрутные картограммы**, на которых показаны еще и расстояния между населенными пунктами и примерное время в пути.

Картограммы изменяют географическое пространство, преобразуя его в легко понимаемые модели реальности. Среди наиболее используемых видов картограмм находятся **площадные картограммы**, которые варьируют размер каждой нанесенной на карту области в зависимости от некоторого параметра этой области.

В самом начале 20 века немецкий картограф Вихель поразил воображение соотечественников необычными картами, которые он сам называл «картограммами». Это были карты результатов голосования в рейхстаг. А необычность заключалась в том, что площади административных единиц (земель) были пропорциональны количеству голосов, поданных, «за» ту или иную политическую фракцию.

Автор метода объяснял свою методику примерно так: предположим, что административные единицы выполнены из пластичного материала, толщина которого пропорциональна показателям картографирования. Если по такой «карте» пройти катком, чтобы толщина слоя стала бы равна среднему показателю, то толстые районы расширились бы, а за счет сжатия соседей, тонкие уменьшились по площади и стали толще, достигнув также средней величины. Таким образом, общая площадь картографируемой территории осталась бы прежней, а внутренние границы причудливым образом перераспределились бы за счет неоднородности картографируемого показателя.

Позднее появился термин «анаморфоза». Это может показаться странным и неправдоподобным, но в толковом словаре В.Даля такое слово есть. В.Даль давал такое толкование слову: «Безобразная, но правильно искаженная картина, принимающая свой вид в граненом или гнутом зеркале».

На слайде показана карта плотности населения Алтайского края и ее анаморфоза. Очень интересно (хотя бы поверхностно) сравнить эти два изображения. При визуальном анализе обращается внимание на экстремальные характеристики: на левой карте самая большая плотность в районе Барнаула (черная заливка), самая маленькая в Горном Алтае (белый фон). Глядя на правое изображение, мы можем оценить, что в первом регионе населения, по крайней мере, в 5-6 раз больше, чем во всем Горном Алтае.

Следующий слайд показывает распределение населения в мире на 2000 г.

Слайд 12. Пример интерактивного вывода: ГИС в службах спасения

Несмотря на преобладание карт и их производных среди выходных документов ГИС, некоторые данные лучше представляются в иных формах, когда карта не создается вовсе, или как дополнение к карте (когда, например, карта не может быть непосредственно воспринята целевой аудиторией).

Возможно, классическим примером замены картографического вывода альтернативной формой является использование ГИС в службе спасения. Если вызов требует тушения пожара, то ГИС позволяет одновременно с принятием вызова оператором определить ближайшую к указанному адресу пожарную часть и направить в нее электронный сигнал тревоги. То есть, в данном случае выводом ГИС будет не карта, а электронный сигнал. Кроме того, та же ГИС может определить кратчайший маршрут для пожарной машины и нанести его на карту, а саму карту дополнить маршрутным листом, в котором перечислены улицы, перекрестки и иные отметки, позволяющие пройти маршрут без помощи карты. Хотя эксплуатируемые сегодня системы не так интеллектуальны, технические возможности вполне позволяют реализовать описанный сценарий, в котором альтернативный вывод может как дополнять карту, так и заменять ее.

Слайд 13. Примеры некартографического вывода.

Существуют многие альтернативные виды вывода из ГИС, но более других распространены таблицы, текстовые отчеты, графики и диаграммы, которые, помимо прочего, не требуют для вывода дорогостоящего оборудования.

Построение текстовых отчетов с развитой графикой благодаря компьютеру стало почти что тривиальной задачей. Многие современные программы используют специальное программное обеспечение, именуемое генератором отчетов. Генераторы отчетов помогают максимально автоматизировать составление и печать отчетов на основе атрибутивной информации, а также дать им должную наглядность.

Многие современные ГИС интенсивно развивают свое программное обеспечение в области 3-х мерной графики и отображении объемных фигур, зачастую представляя результат обработки материала не в виде плоской карты, а в виде 3D- модели.

Существует еще несколько интересных возможностей в способах вывода данных. Например, модуль ArcGIS Schematics выполняет генерирование схем и геосхематических графических диаграмм, построенных на основе сетей, хранящихся в базе геоданных. Независимо от вида сети (электрическая, газовая, телекоммуникационная или табличное представление), ArcGIS Schematics генерирует соответствующий сетевой граф или схему. Под схемой ГИС-сети любого вида понимается упрощенное представление объекта или группы объектов, помогающее понять структуру и функционирование сети. Она также отличается от картографического представления объектов, как внешний вид компьютерной платы от ее принципиальной электрической схемы. Модуль ArcGIS Schematics позволяет нарисовать различные графические изображения структуры сети, поместить их в документ или на карту.

1. Что такое центроид ?
2. Даны 4 точки с координатами (x_1, y_1) , (x_2, y_2) , (x_3, y_3) , (x_4, y_4) и с весами s_1 , s_2, s_3, s_4 , соответственно. Рассчитайте центроид данного точечного распределения.
3. Назовите линейные объекты высокого уровня.
4. Назовите элементы сети ?
5. Как определяется направление потока в сети ?
6. Какова роль барьера в сети ?
7. Какие меры извилистости существуют ?
8. Назовите площадные объекты высокого уровня?
9. Рассчитайте число Эйлера для заданной конфигурации полигонов
10. Как выявить наиболее вытянутые полигоны в заданном полигональном наборе данных ?
11. Как определяются периметры и площади полигональных объектов на векторном типе данных ?
12. Как определяются периметры и площади полигональных объектов на растровом типе данных ?
13. Что такое аллокация ?
14. Что такое функциональное расстояние ?
15. С какой целью рассчитываются функциональные расстояния?
16. Что такое пространственное распределение ?
17. Перечислите виды пространственных распределений ?
18. Перечислите методы изучения пространственных распределений точечных объектов.
19. Для чего изучают пространственных распределения ?
20. Какую меру можно использовать для определения направленности линейных и площадных объектов? Какую информацию о распределении дает эта мера?
21. Рассчитайте гамма-индекс для заданной сети
22. Рассчитайте альфа -индекс для заданной сети
23. Назовите примеры непрерывной поверхности.
24. Назовите примеры дискретной поверхности.
25. Какие модели поверхностей используются в ГИС ?
26. Опишите алгоритм триангуляции Делоне.
27. Почему TIN считают 2.5 мерной моделью ?
28. Что такое резкая линия перегиба ?
29. Для чего в модели TIN используют полигоны отсечения ?
30. Чему равно разрешение грида ?
31. Назовите отличия молей GRID от других растровых моделей.
32. Определение интерполяции
33. Перечислите известные вам методы интерполяции
34. На основании чего определяется вес измеренного значения в методе IDW
35. На основании чего определяется вес измеренного значения в методе Кригинг.
36. Чему равен предельный радиус корреляции и остаточная дисперсия на данной вариограмме ?
37. Требуется построить структурную карту по кровле горизонта «А» с учетом разрывных нарушений . Какой метод интерполяции лучше выбрать ?
38. Требуется определить продуктивность литохимического ореола на основе карты распределения содержания металла (полезного ископаемого). Какой метод интерполяции лучше выбрать ?

39. Требуется рассчитать приблизительное положение уровня грунтовых вод на заданной территории. Какой метод интерполяции лучше выбрать ?
40. По результатам глубокого бурения в 10 скважинах были вскрыты породы кристаллического фундамента и определены абсолютные отметки кровли фундамента. Требуется оценить общее поведения поверхности фундамента на территории исследования. Какой метод интерполяции лучше выбрать ?
41. Нарисуйте полигоны Тиссена для данного распределения точек.
42. Объясните сущность и назначение процедуры классификации (переклассификации) пространственных объектов.
43. Для целочисленного тематического растра RASTER1 с помощью операторов отношения запишите условие, с помощью которого можно выполнить его переклассификацию. В выходном растровом наборе данных должно быть 2 класса. Первый класс должен содержать все ячейки, значения которых меньше 100, кроме ячеек со значением 50. Второй класс должен содержать все ячейки, значения которых меньше либо равно 100 и ячейки со значением 50.
44. В чем состоит принципиальное отличие процедур переклассификации на векторном и растровом типе данных.
45. Перечислите известные вам методы ранжированных классификаций.
46. Как работают переклассификации основанные на окрестности.
47. Что такое фильтр ?
48. Составьте матрицу низкочастотного фильтра размером 4x4.
49. Составьте какой либо вариант матрицы высокочастотного фильтра размером 3x3
50. Составьте какой либо вариант матрицы анизотропного фильтра размером 5x5, позволяющий выявить объекты, ориентированные в направлении север-юг.
51. Что такое буфер ? Назовите виды буферов.
52. Назовите виды переклассификации поверхностей.
53. Какую модель поверхности лучше для определения уклона и экспозиции?
54. В чем преимущества операций наложения на растровом типе данных.
55. Даны 3 целочисленных тематических растра: RASTER1, RASTER2, RASTER3. С помощью логических операторов запишите условие, с помощью которого можно обнаружить местоположения, где значения RASTER1 не превышают значения RASTER2 и значения RASTER2 не превышают значения RASTER3.
56. Даны 3 целочисленных тематических растра: RASTER1, RASTER2, RASTER3. С помощью логических операторов запишите условие, с помощью которого можно обнаружить местоположения, где значения RASTER2 не превышают значения RASTER3 или значения RASTER1 больше значений RASTER3.
57. Как работает функция «Статистика по ячейкам» ?
58. Как работает функция «Зональная статистика» ?
59. Для каждого месяца в течение года для урбанизированной территории были рассчитаны растровые поверхности, показывающие концентрацию углекислого газа. Как, комбинируя операции переклассификации и наложения на растровом типе данных, вы можете выполнить районирование урбанизированной территории по степени изменчивости содержания углекислого газа в течение года?
60. Какие преимущества дает топологическое векторное наложение.

61. Приведите пример задачи, для решения которой можно было бы использовать векторное наложение «точка в полигоне».
62. Приведите пример задачи, для решения которой можно было бы использовать векторное наложение «линия в полигоне».
63. Приведите пример задачи, для решения которой можно было бы использовать векторное наложение «полигон в полигоне».
64. Что такое кластерный допуск ?
65. Перечислите операции векторного наложения, которые вы можете реализовать в системе ArcGIS.
66. В чем сходство и различие функций CLIP и INTERSECT?
67. В чем сходство и различие функций CLIP и SPLIT?
68. Может ли функция UNION быть применена только к одному набору пространственных данных.