

Геоданные: задачи, источники, методы, ПО для работы с ними

Максим Дубинин

План

- Определение
- Особенности
- Типы
- Задачи
- Основные наборы данных
- Программное обеспечение
- Что делать?

Определение

Геоданные — данные о пространственных объектах.

Пространственный объект - цифровая модель материального или абстрактного объекта реального или виртуального мира включающая его идентификаторы, координатные и атрибутивные данные.

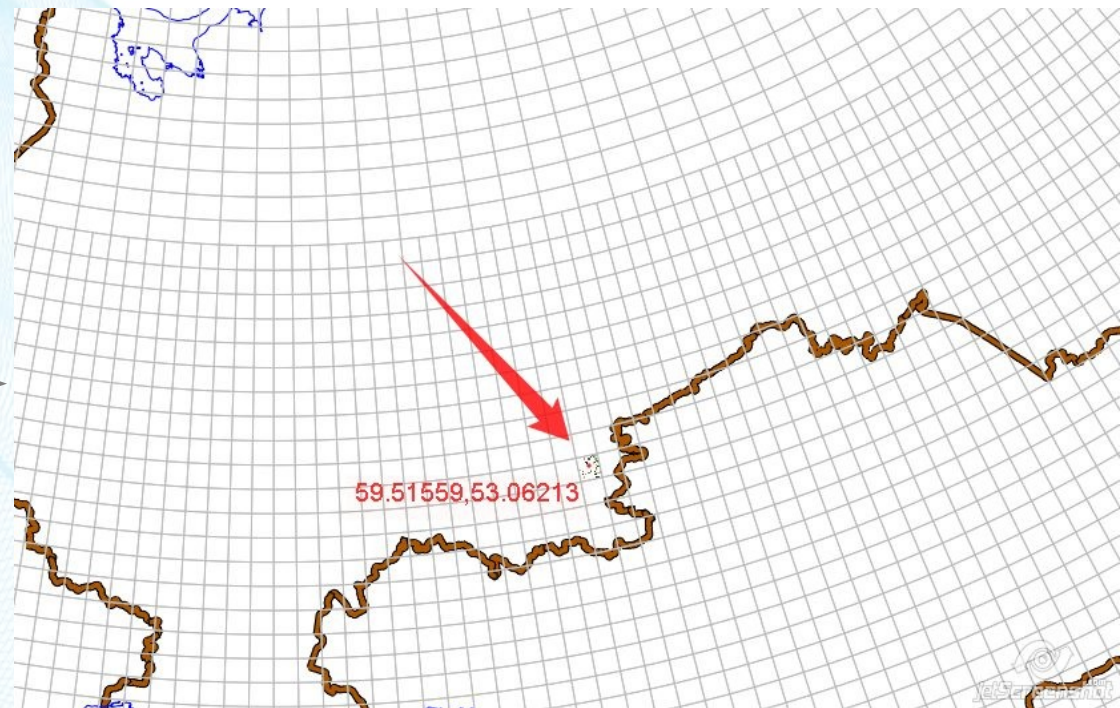
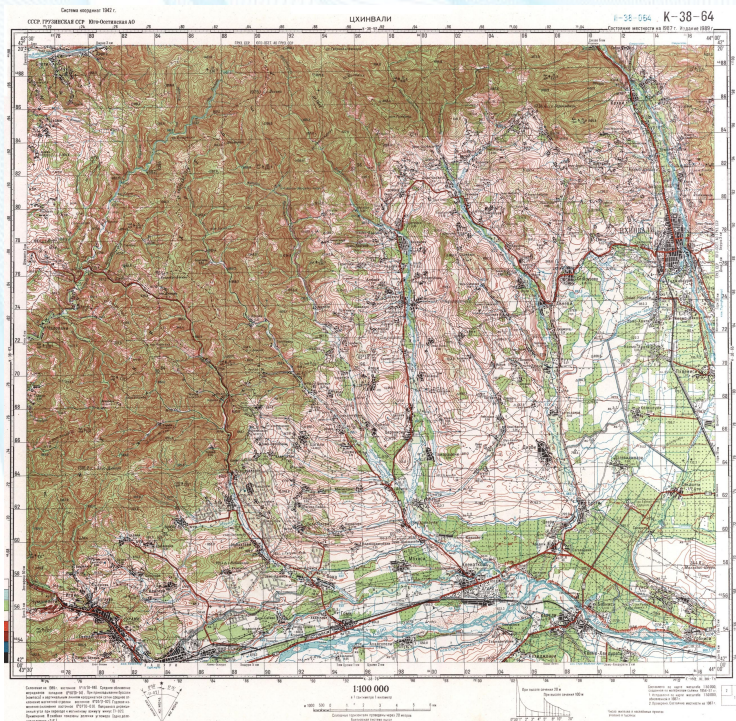
Особенности

Is geospatial special ...

- геопространственные системы координат
- высокая информационная насыщенность
- объединение геометрической и атрибутивной информации
- послойная организация

Особенности

- явное нахождение в одной из пространственных систем координат



Отсканированное изображение, каждый элемент имеет X,Y,Z

Географически привязанное изображение, каждый элемент имеет широту, долготу, Z

Особенности

- высокая информационная насыщенность

MODIS

Передача 10.6 миллионов бит данных в секунду ~ 53 Тб данных в день

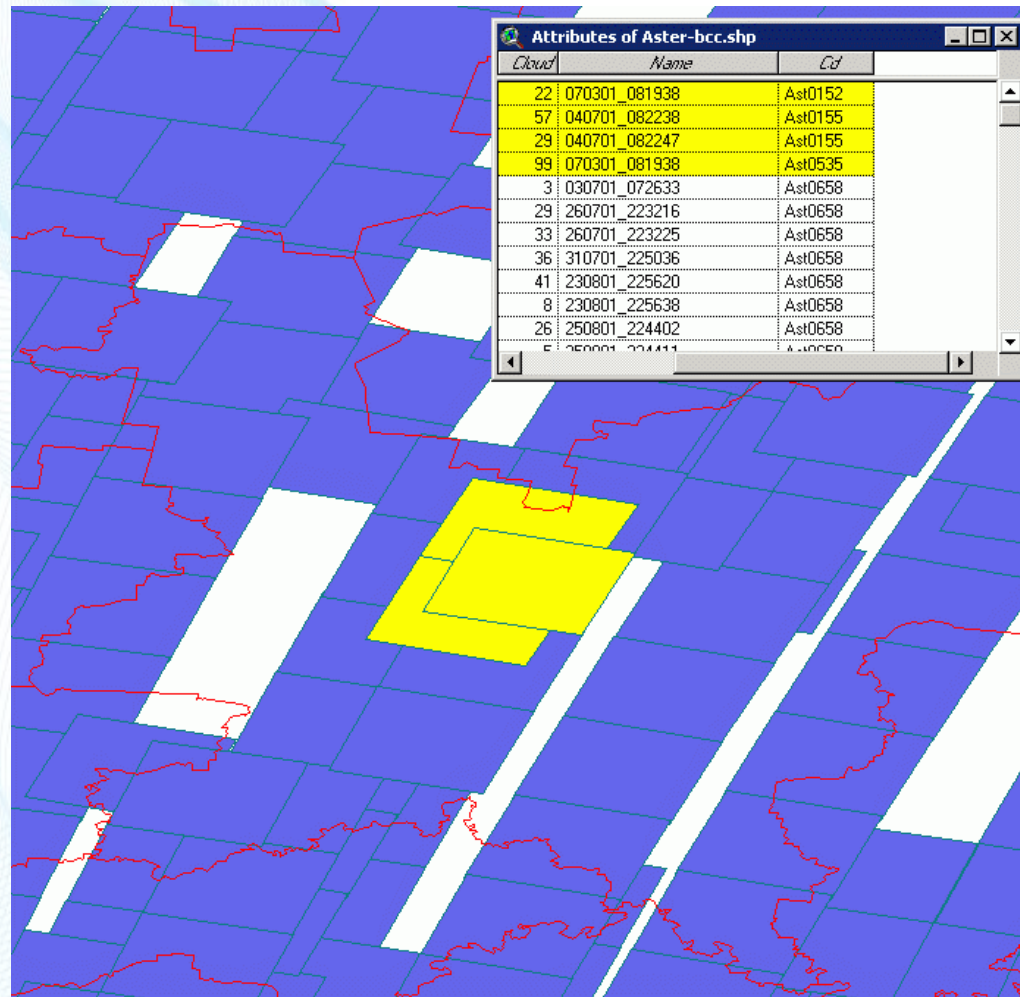
OpenStreetMap

Число пользователей: 0.5 миллиона

Число узлов: 1396 миллионов

Особенности

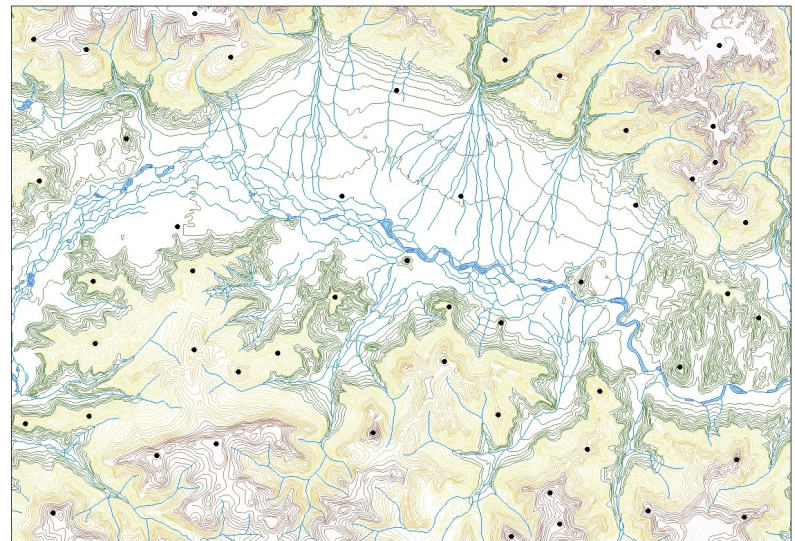
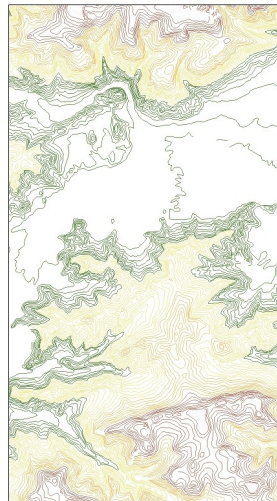
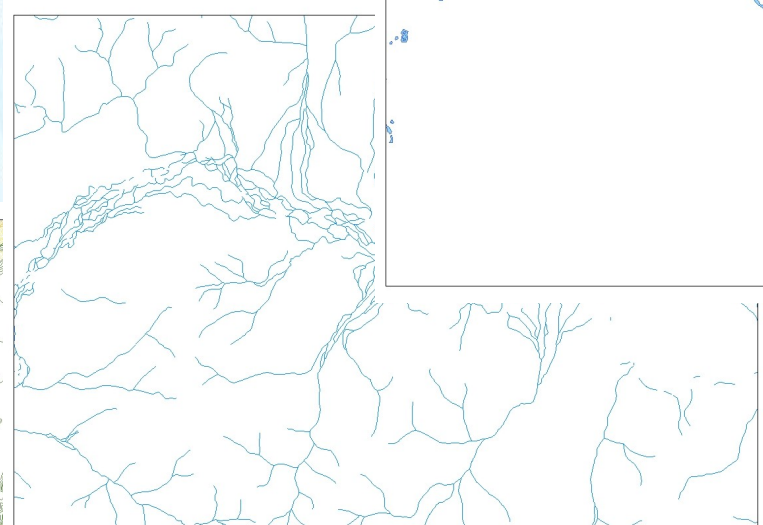
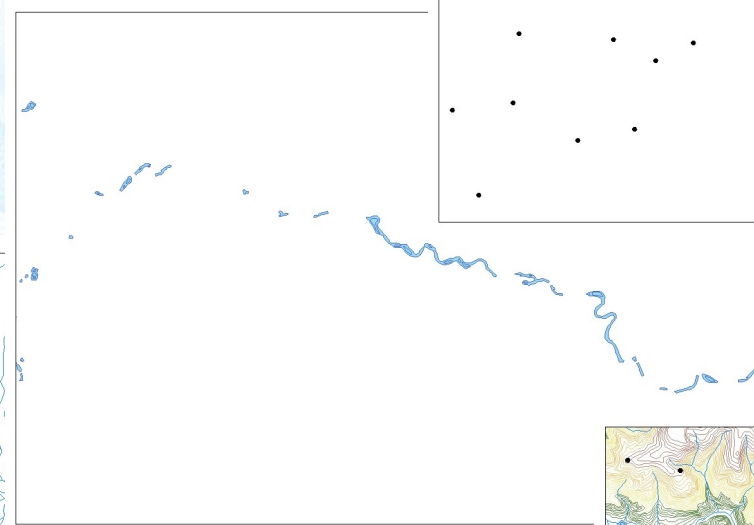
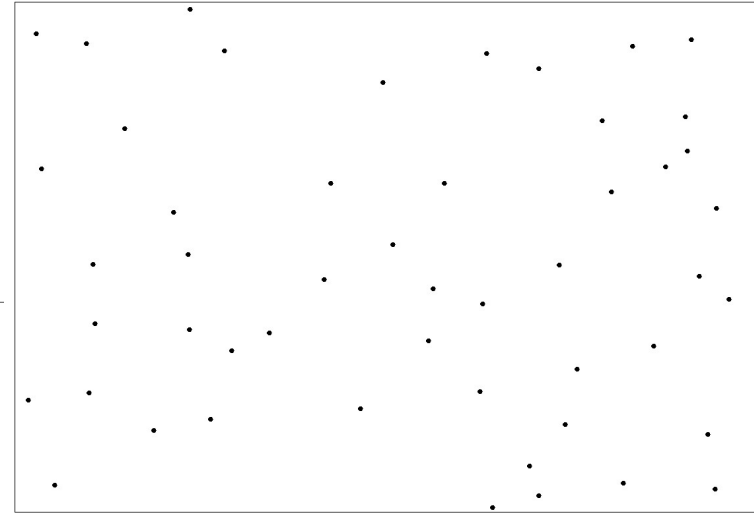
- единство геометрической и атрибутивной информации



Каждый элемент геометрии связан с набором атрибутивных полей (shape — flat table, sqlite/spatialite - БД)

Особенности

- послойная организация

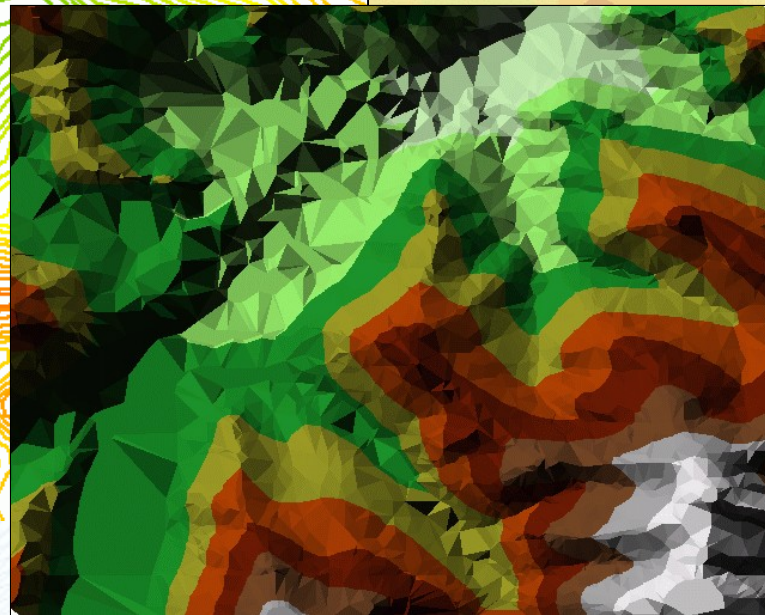
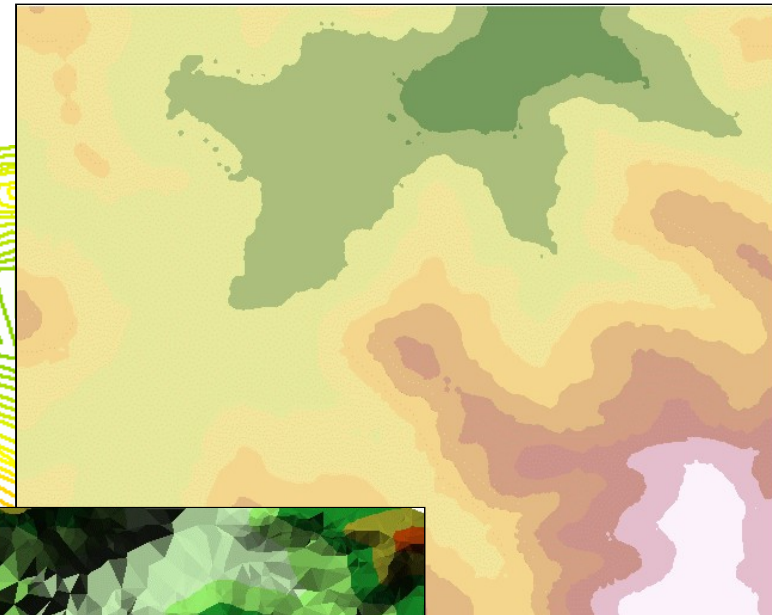
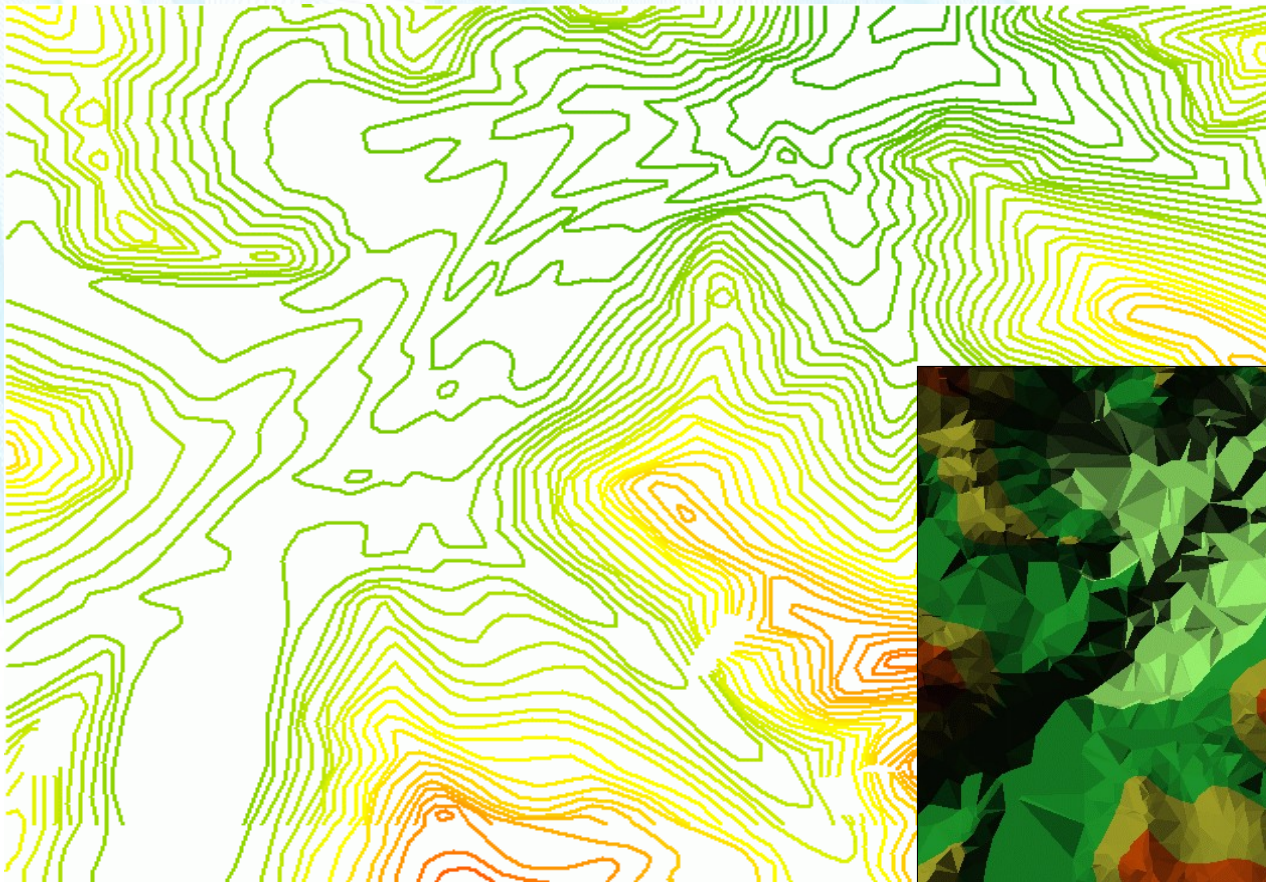


Типы

- Форма организации элементов
- Способ получения данных
- Масштаб, линейное разрешение на местности, охват
- Тематике
- Доступ

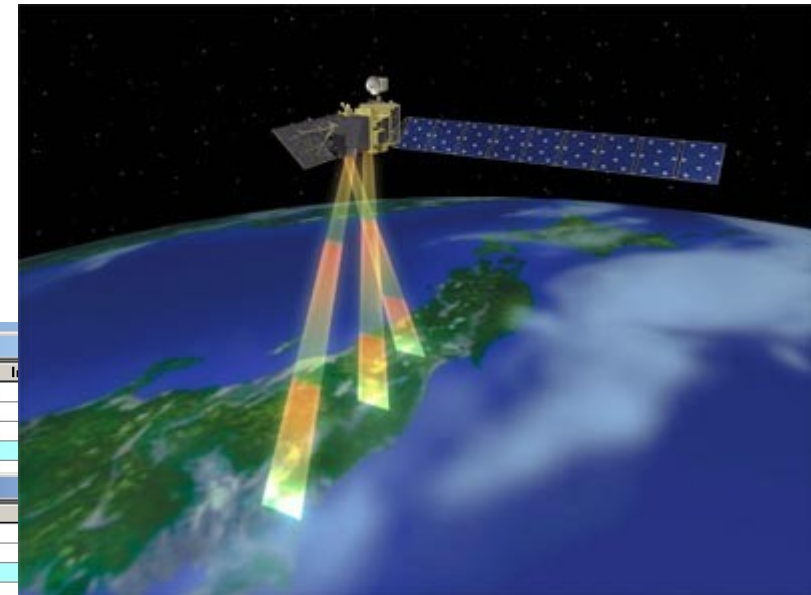
Типы

- По форме представления:
дискретная (вектор), регулярно сетевая (растр)
нерегулярно сетевая (TIN)



Типы

По способу получения: полевой, камеральный, дистанционный



The screenshot displays a GIS application with several data tables and a map. The map shows a geographical area with various features and data points. The tables provide detailed information about the data layers.

Source	Add_Info	Loc_Name
ков А., Лукаревский В	South_West Тува Octob	
ин А. по словам паста	Kuksin2002-03.doc	
как Байлан-оол, чабан	Expre1.doc	Барлык-Бажы
как Кызыл-оол, чабан	Expre2.doc	Ой-Бажы
как Орлан, чабан	Expre3.doc	

Cod	Descriptio
62	самка с котятами
67	самка с котятами
37	самка с котятами
68	самка с котятами?
69	нет данных
84	самка
166	самка с котятами
175	нет данных

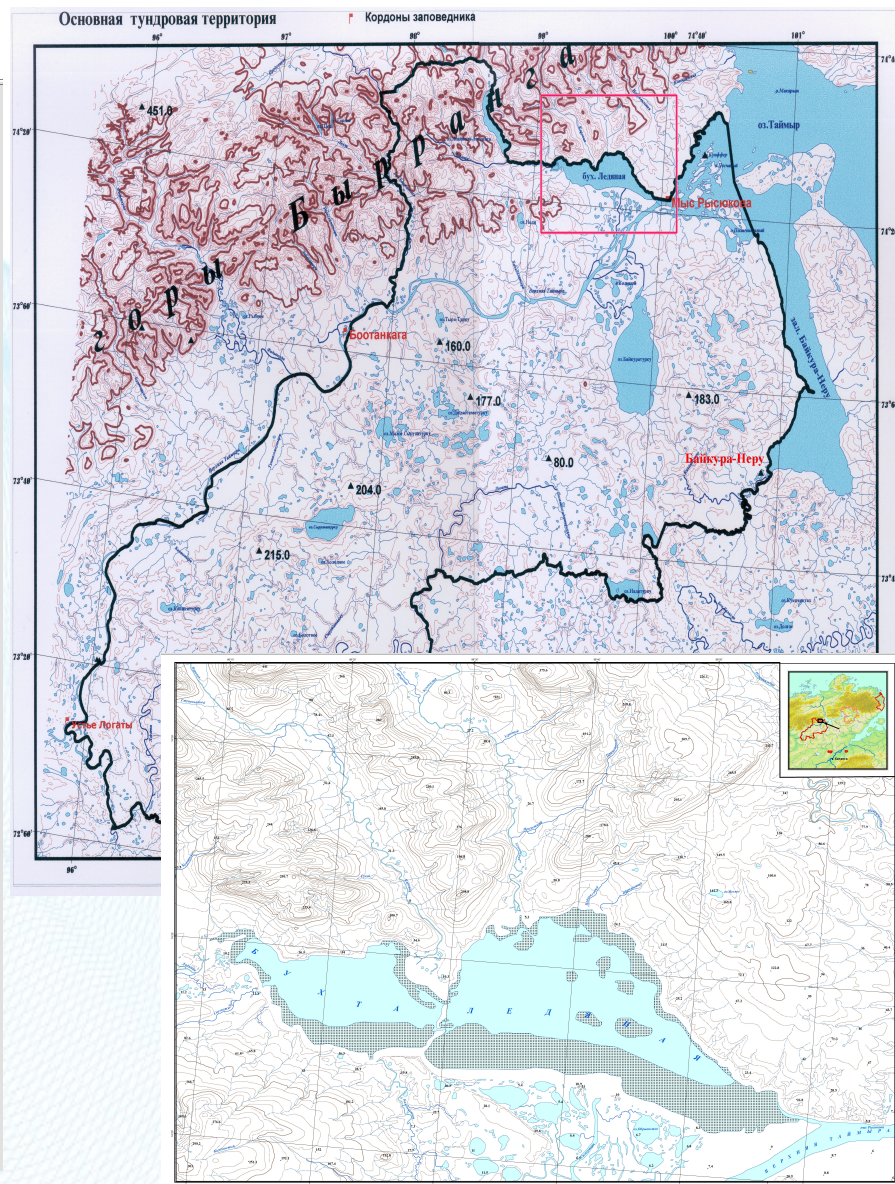
Code	Killed	livestock	Confirmati
36	4	яки	да
37	2	яки	да
38	3	лошади	да
39	4	МРС	нет
40	10	яки, МРС	нет
41	71	МРС	нет
42	32	МРС	нет
43	8	яки, МРС	нет
44	1	МРС	да

Code	What_found	Cause_deat
18	череп	нет данных
19	череп	нет данных
22	труп	фтороацетат бария
44	труп	застрелен при нападении на скот
63	шкура и череп	застрелен?
84	самка	застрелен
137	нет данных	застрелен
158	труп	пойман в петлю
189	труп	застрелен при нападении на скот

Code	Species	Sex_Age
58	сибирский козерог	сеголеток

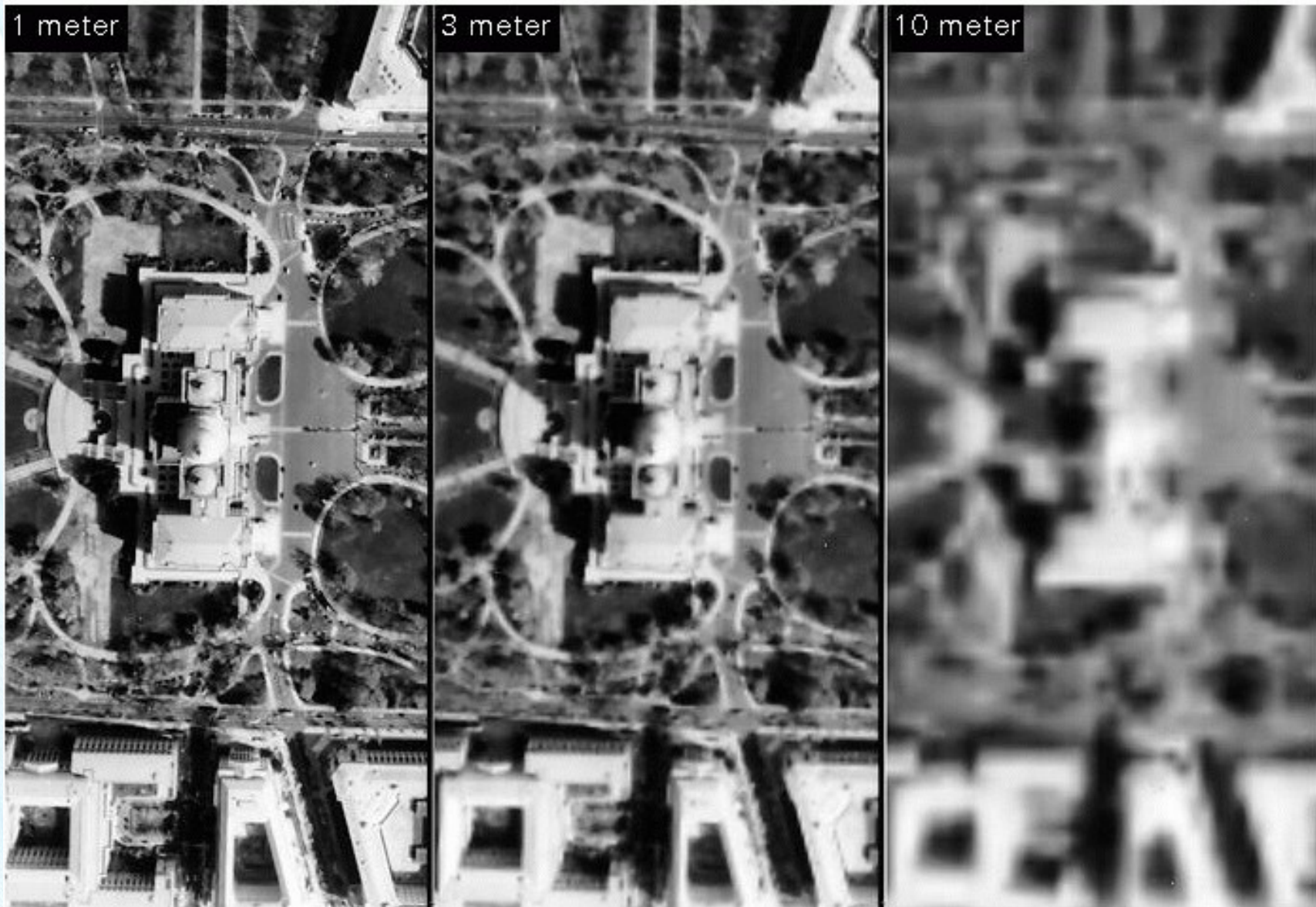
Типы

- По исходному масштабу: мелкомасштабные, среднемасштабные, крупномасштабные



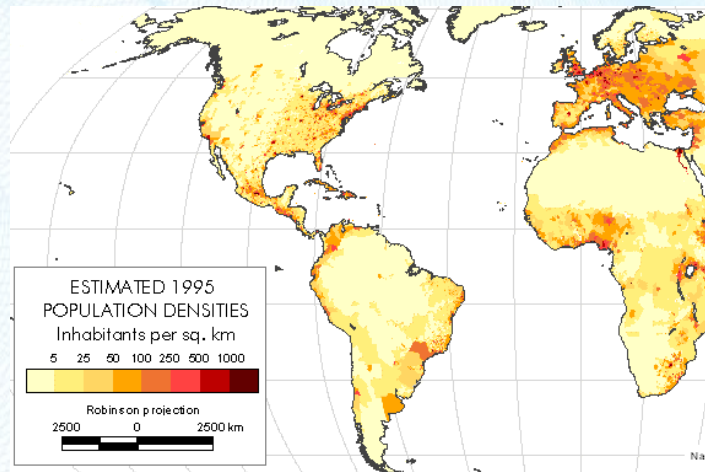
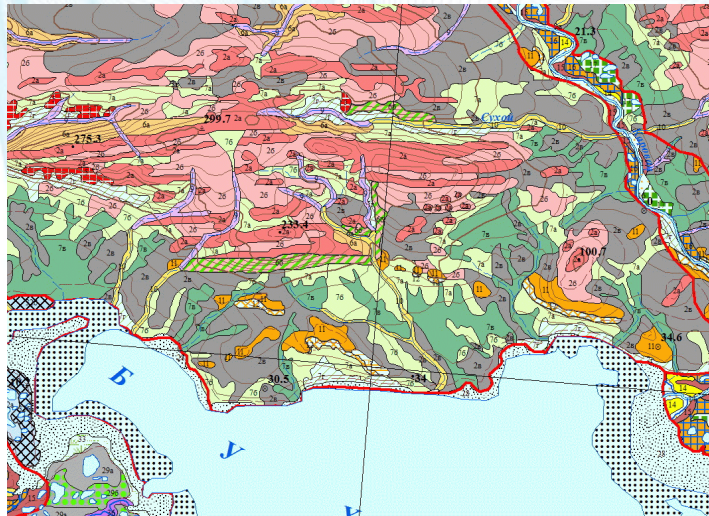
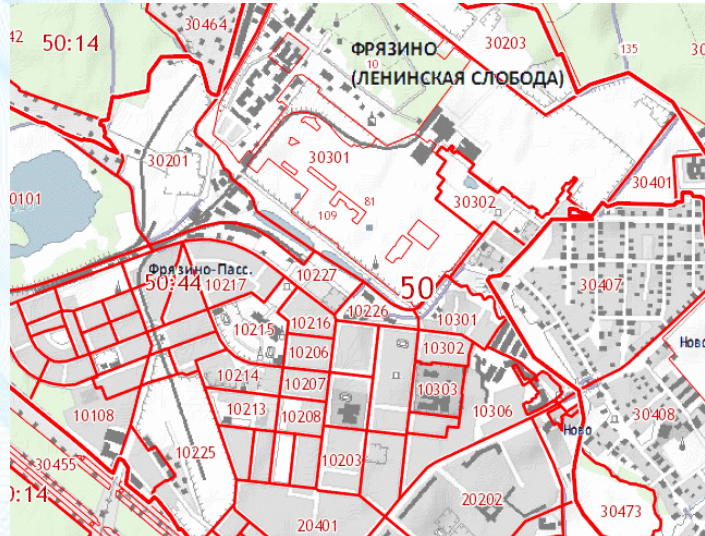
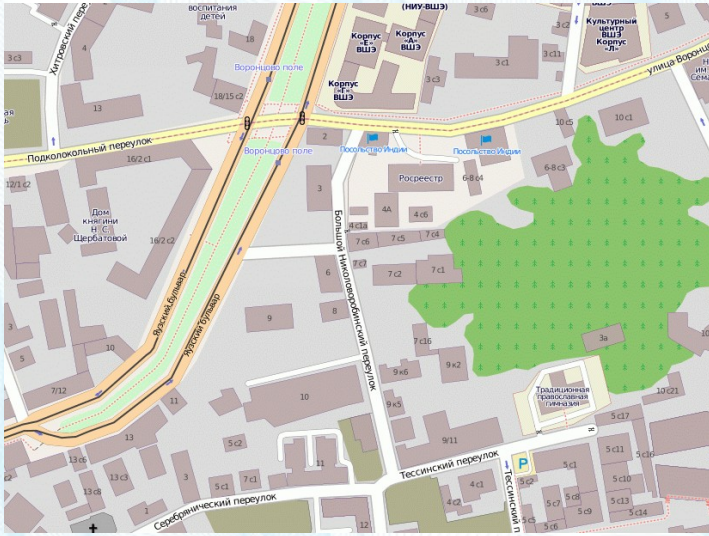
Типы

- По разрешению: сверхвысокое, высокое, среднее, низкое.



Типы

- По тематике: общая топография, кадастр, природные ресурсы, демография и многое другое



Типы

- По лицензии и режиму доступа: открытые, закрытые



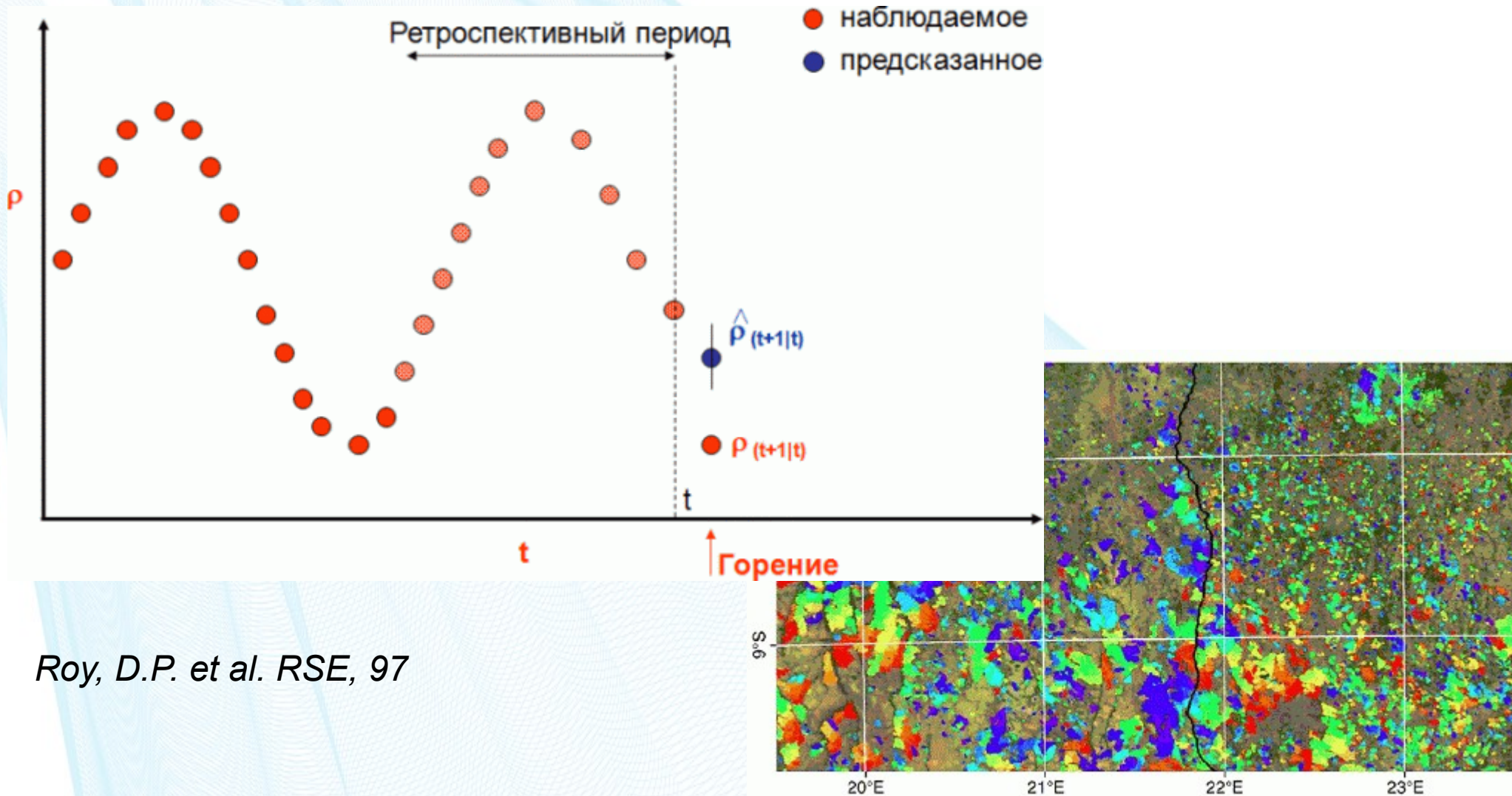
Задачи

- Картирование природных и антропогенных объектов и явлений: идентификация объектов и их состояния, количества, формы, свойств.
- Анализ пространственного распределения объектов и явлений: плотности, частоты, связность
- Выявление причинно-следственных связей между пространственным распределением объектов и явлений и другими данными.
- Моделирование и прогнозирование пространственно распределенных процессов.

Задачи

- Картирование природных и антропогенных объектов и явлений

Пример: Детектирование гарей по временным сериям

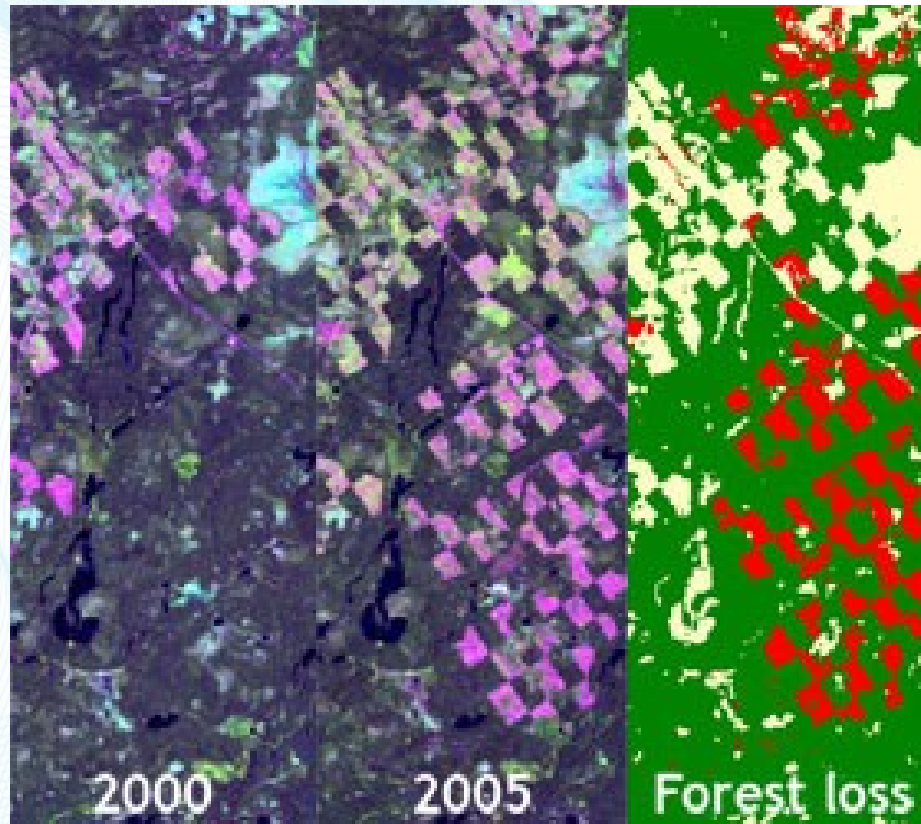


Roy, D.P. et al. RSE, 97

Задачи

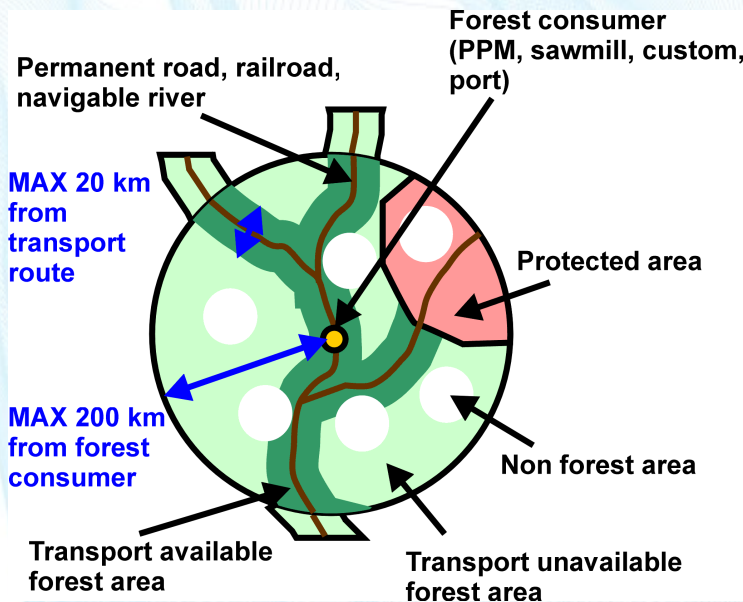
- Картирование природных и антропогенных объектов и явлений

Пример: Детектирование рубок по временным сериям



Задачи

- Анализ пространственного распределения объектов и явлений
Пример: Транспортная доступность лесного фонда



Potapov et.al

Задачи

- Анализ пространственного распределения объектов и явлений
Пример: Анализ горимости

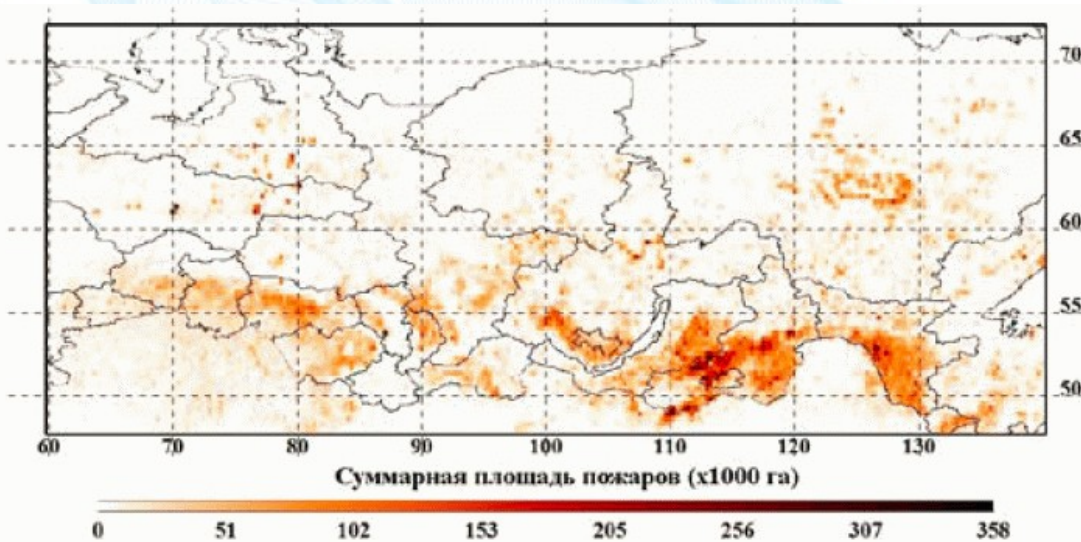


Рис. 1. Карта суммарной площади пожаров (в тысячах гектаров), по данным спутников NOAA за период с 1996 по 2008 гг.

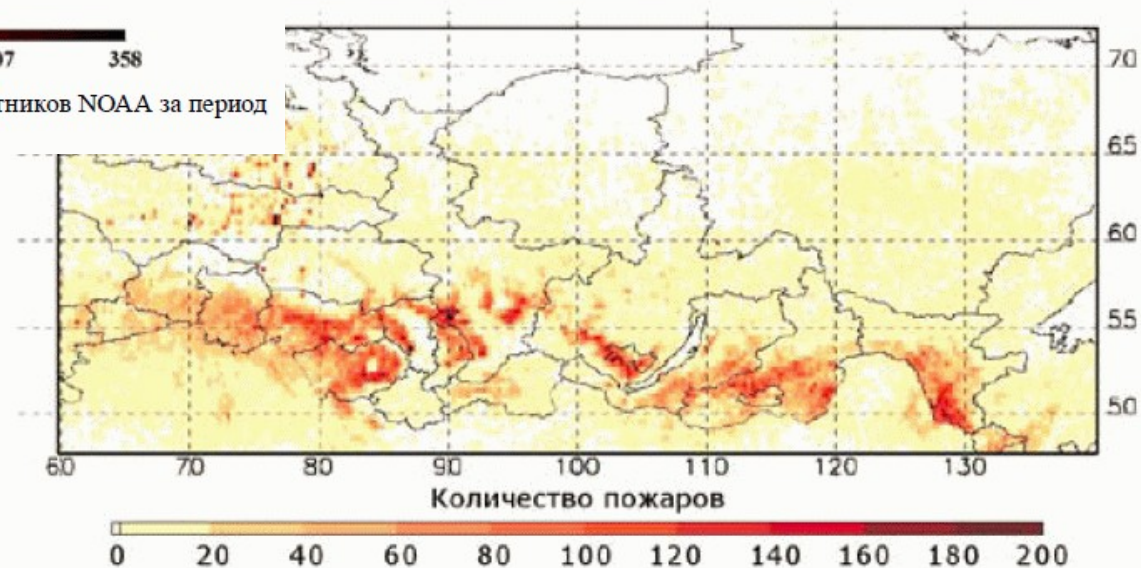
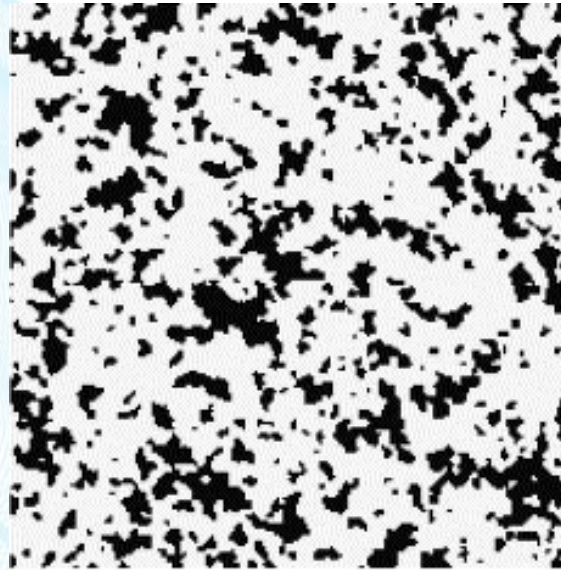
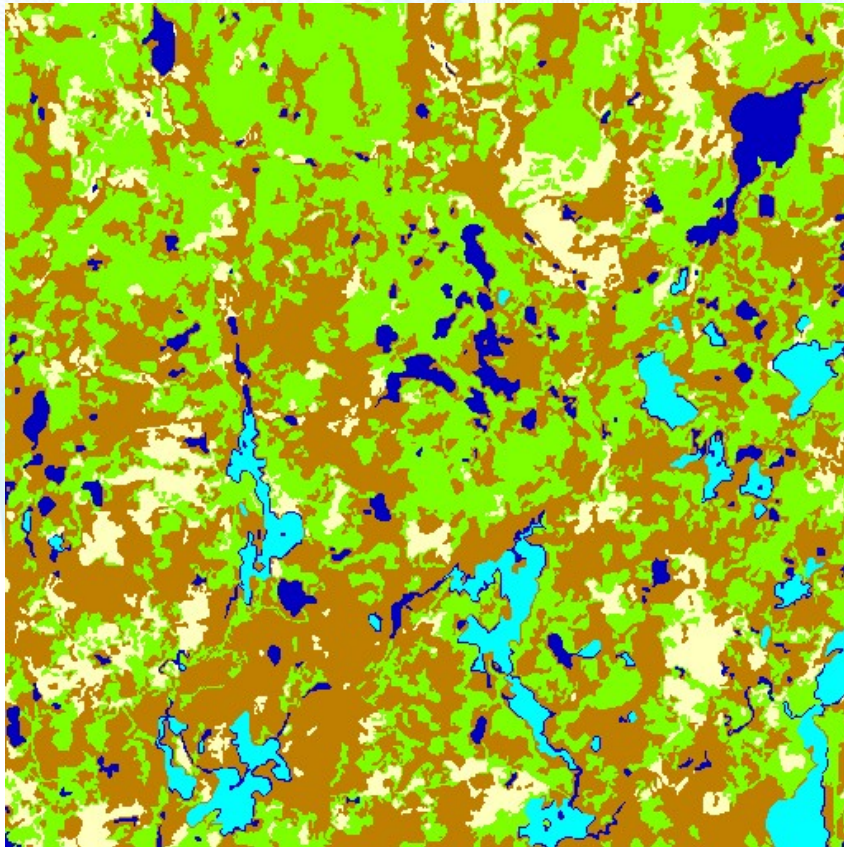


Рис. 2. Карта суммарного количества пожаров, по данным спутников NOAA с 1996 по 2008 гг.

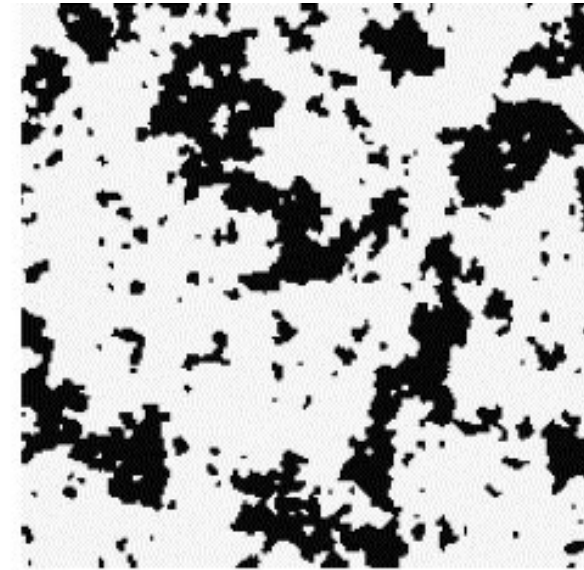
Рубцов А.В и др., 2010

Задачи

- Анализ пространственного распределения объектов и явлений
Пример: Ландшафтные индексы



(a)

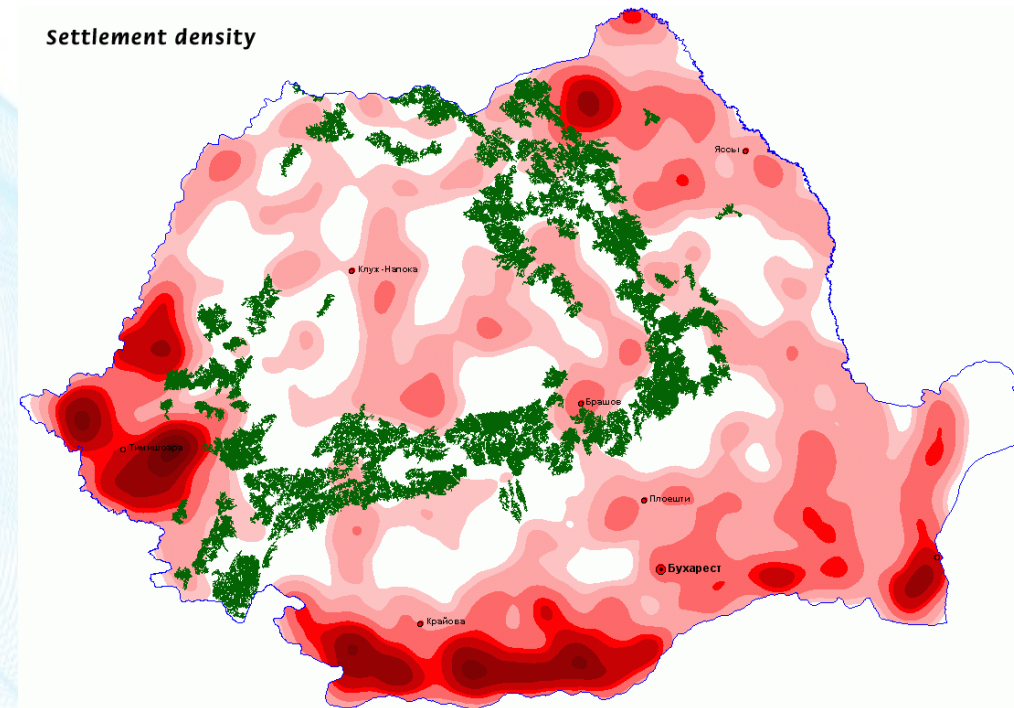
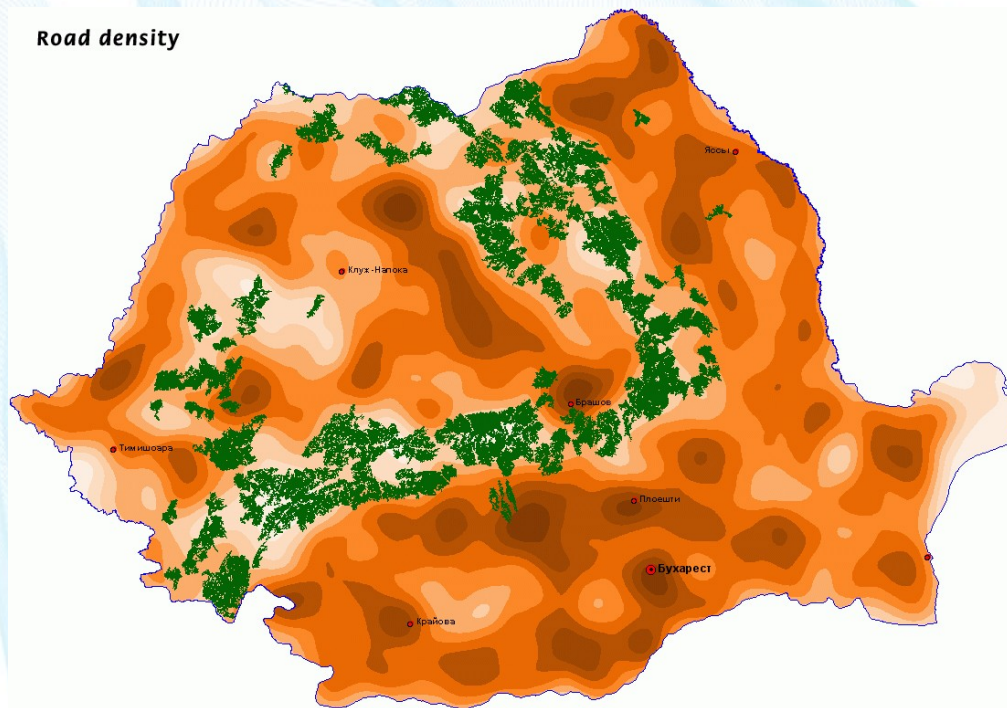


(b)

	# patches	Patch density	Mean patch size	Area weight. mean patch size	Edge length	Mean shape index
a	318	0.80	125.8	19955.4	9393	1.37
b	165	0.41	262.4	19907.2	5404	1.32

Задачи

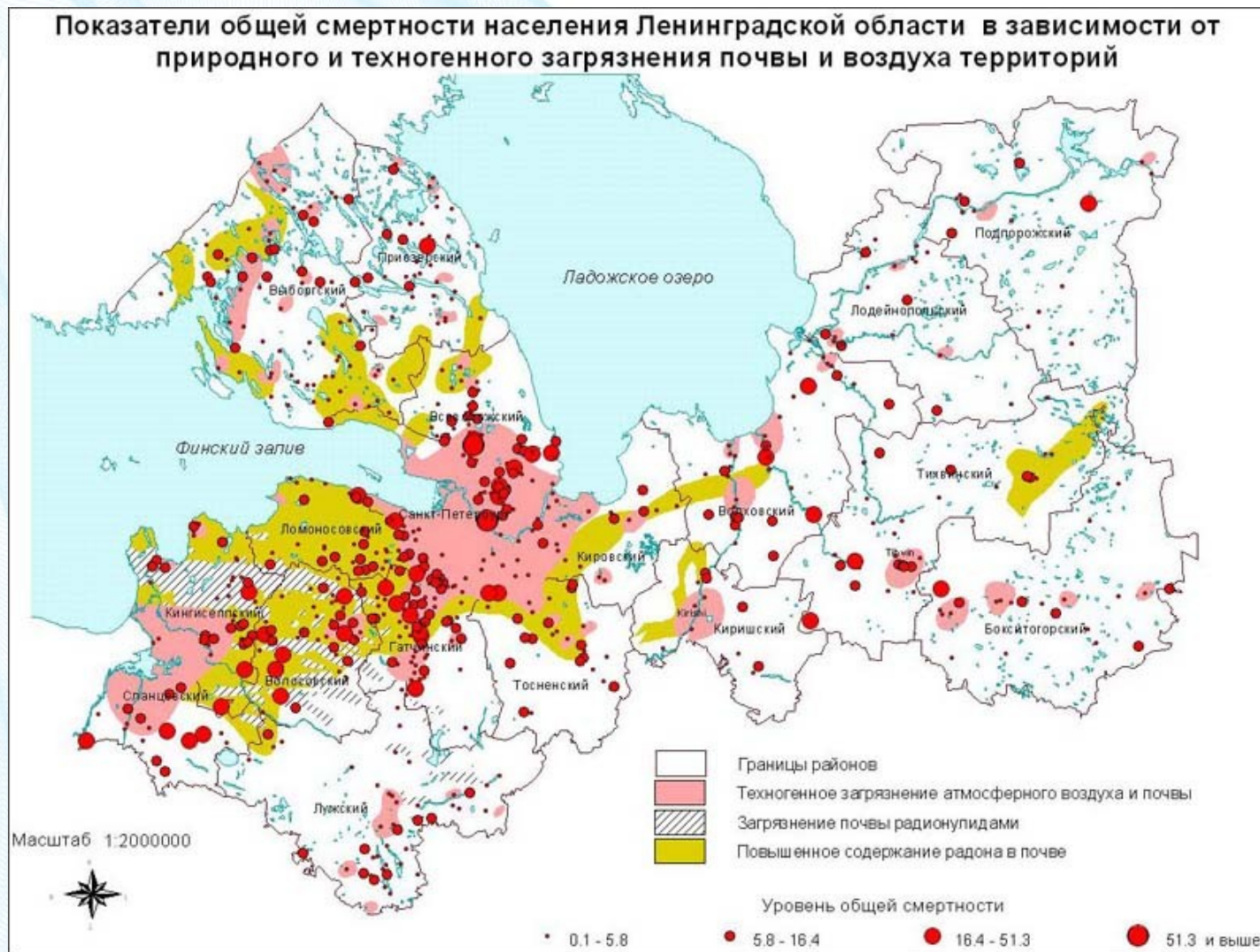
- Выявление причинно-следственных связей



Задачи

- Выявление причинно-следственных связей

Пример: Пространственная эпидемиология

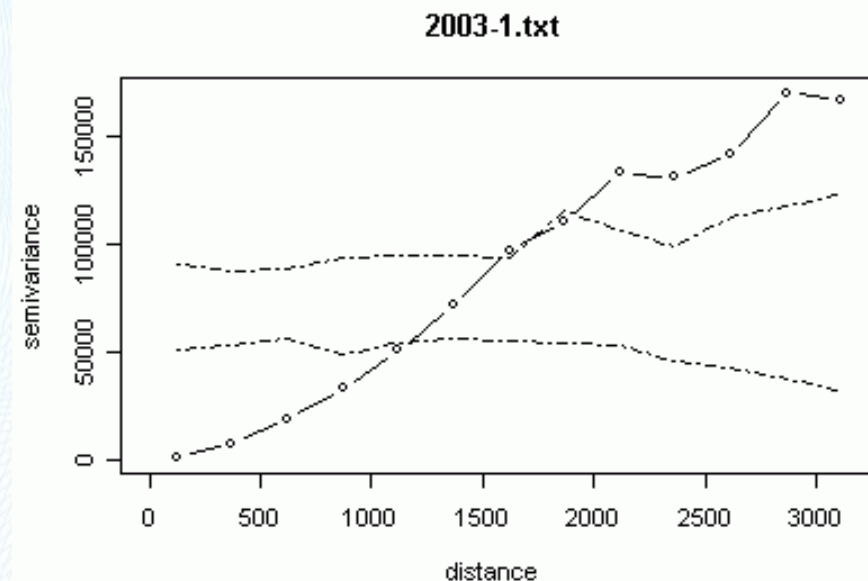
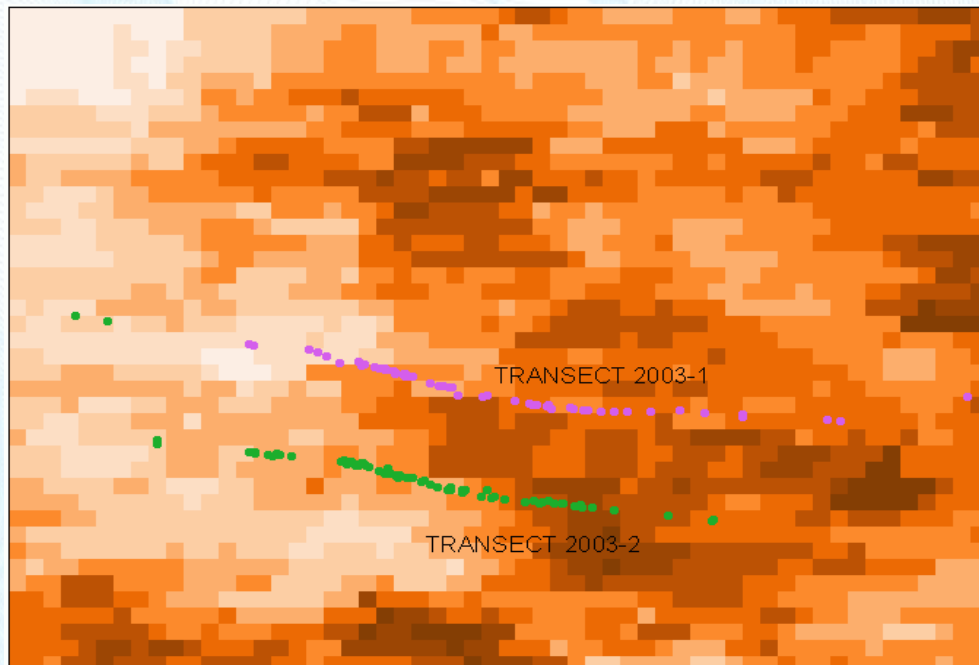


Подводные камни

- Пространственные автокорреляции

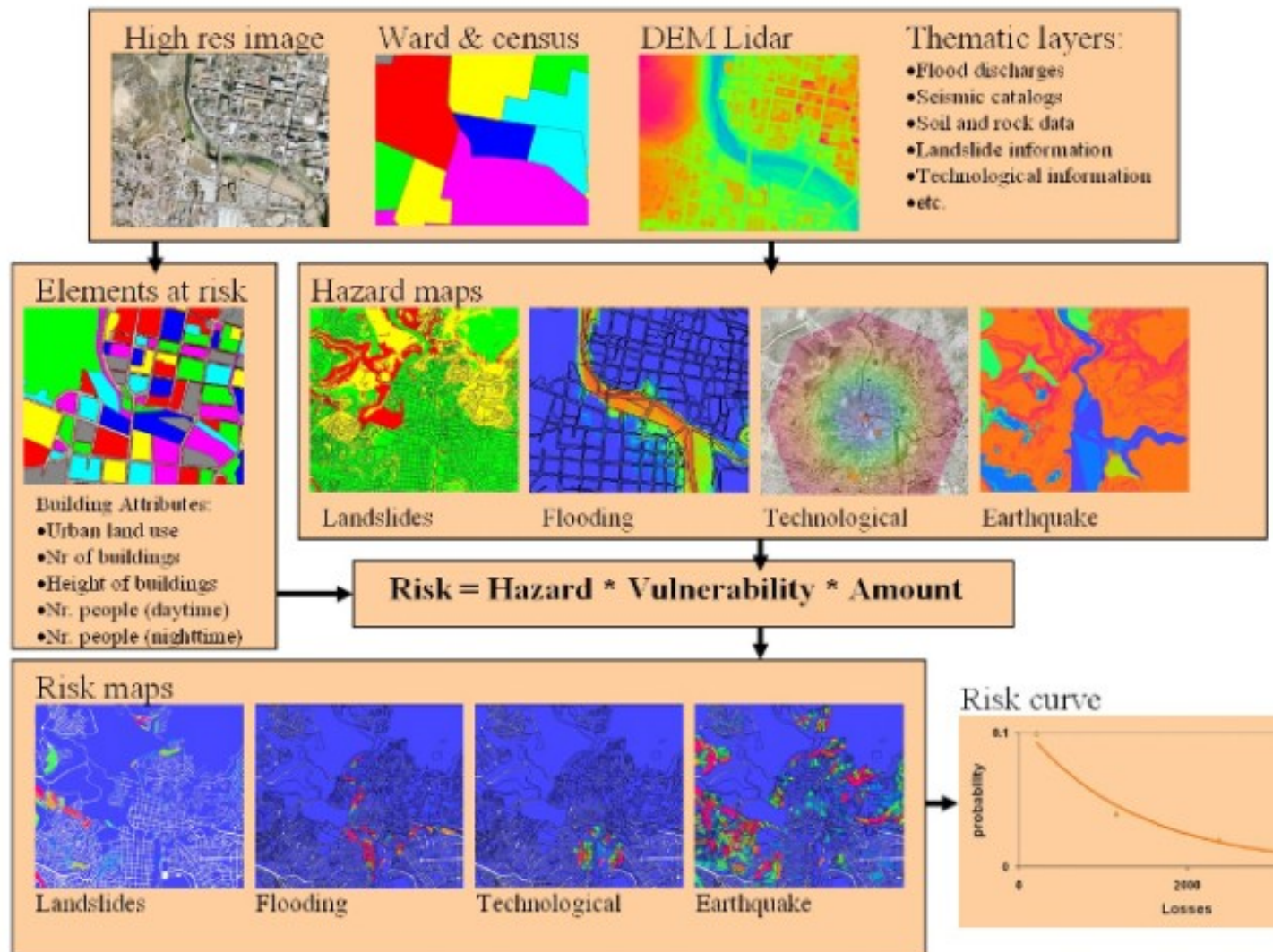
Первый закон географии Тоблера: все связано со всем, но то что ближе связано сильнее.

1. Показывает наличие пространственного процесса
2. Присутствие АК нарушает условие независимости наблюдений



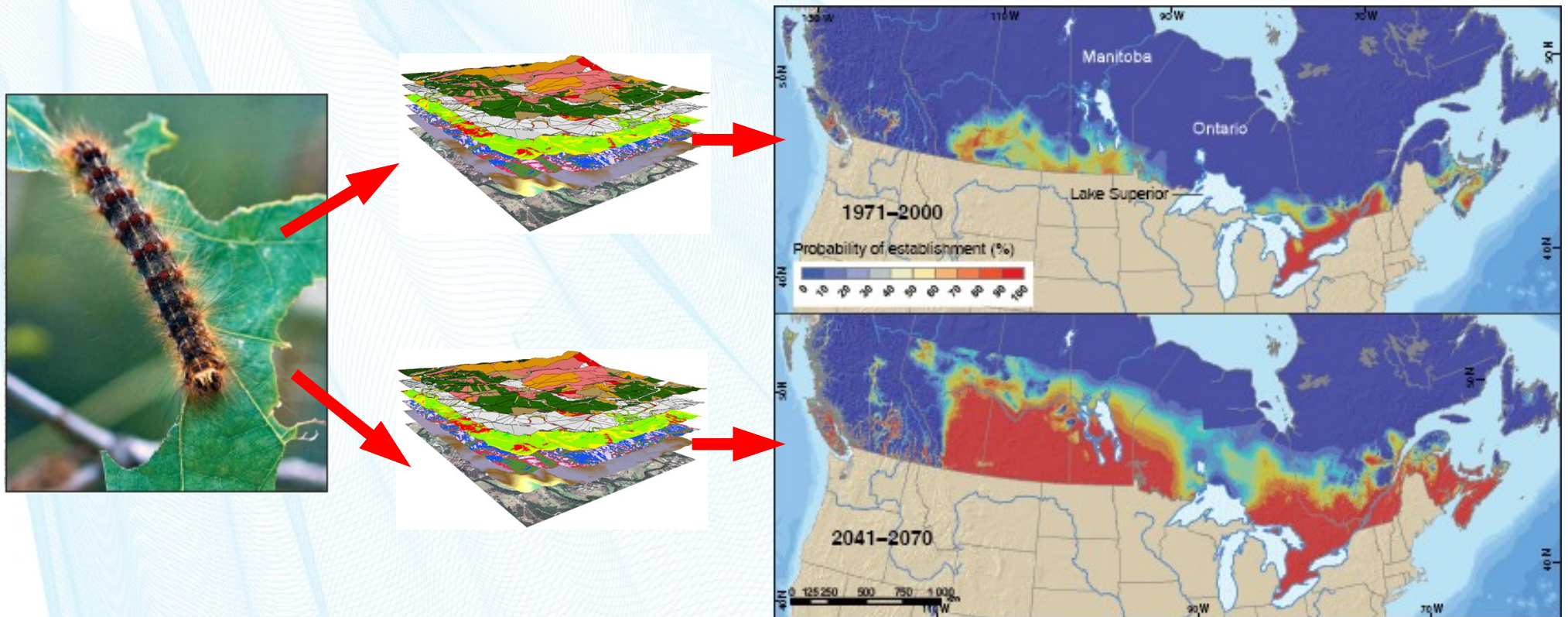
Задачи

- Моделирование и прогнозирование
 - Пример: Экспертные системы пространственной оценки рисков



Задачи

- Моделирование и прогнозирование
 - Пример: Моделирование экониш и прогнозирование будущих распределений объектов и явлений



Основные наборы данных

- Базовая топография
- Наземная съемка
- Рельеф
- Дистанционное зондирование

Основные наборы данных

- Базовая топография:

OpenStreetMap

- 1:2000, 1:5000
- Дешифрирование и треки



VMap0

- 1:1000000
- Оцифровка бумажных карт

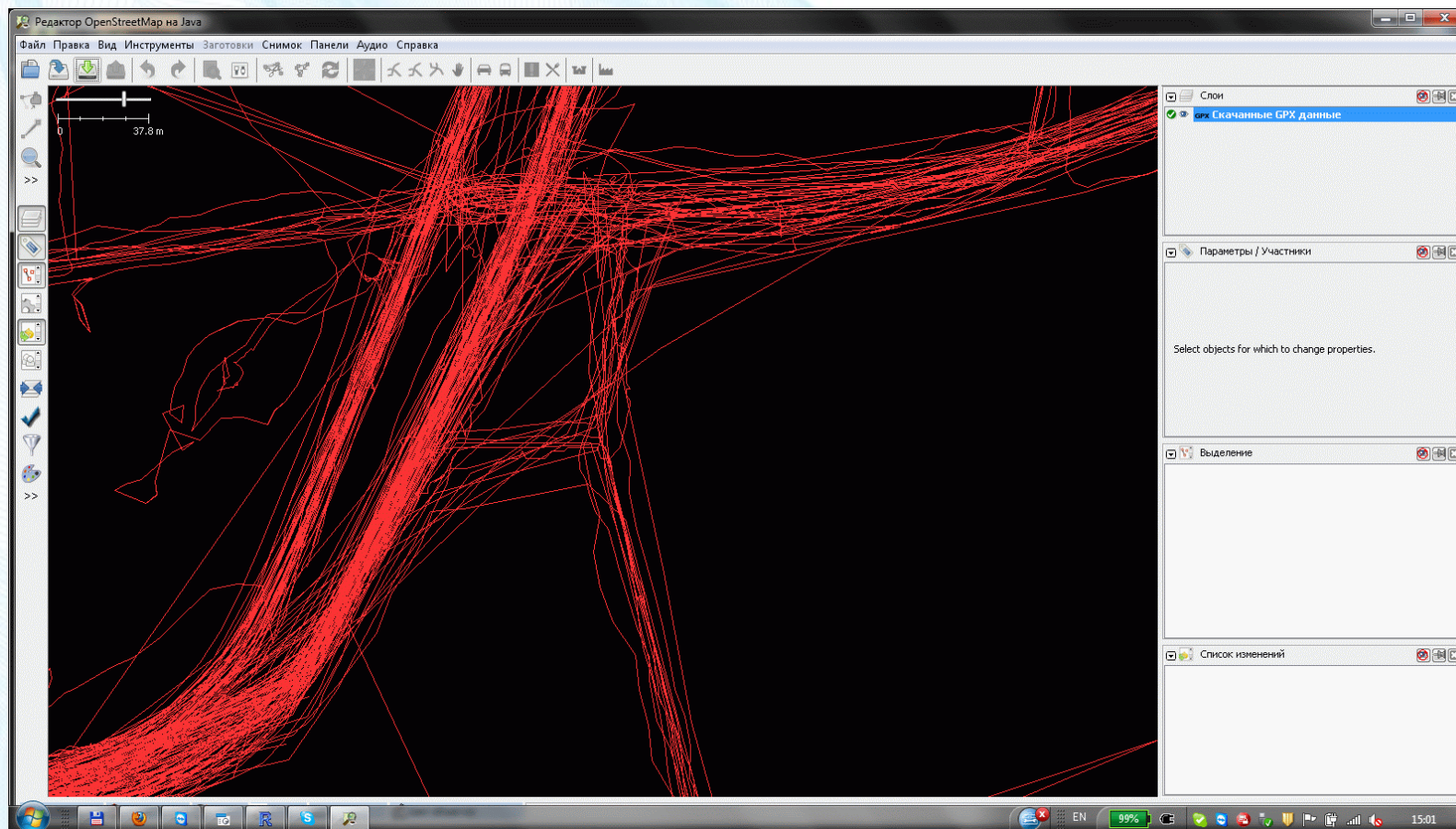


Основные наборы данных

- Наземная съемка: GPS-треки

Пользовательские треки из OpenStreetMap

- Число загруженных точек GPS: 2775 миллионов



Основные наборы данных

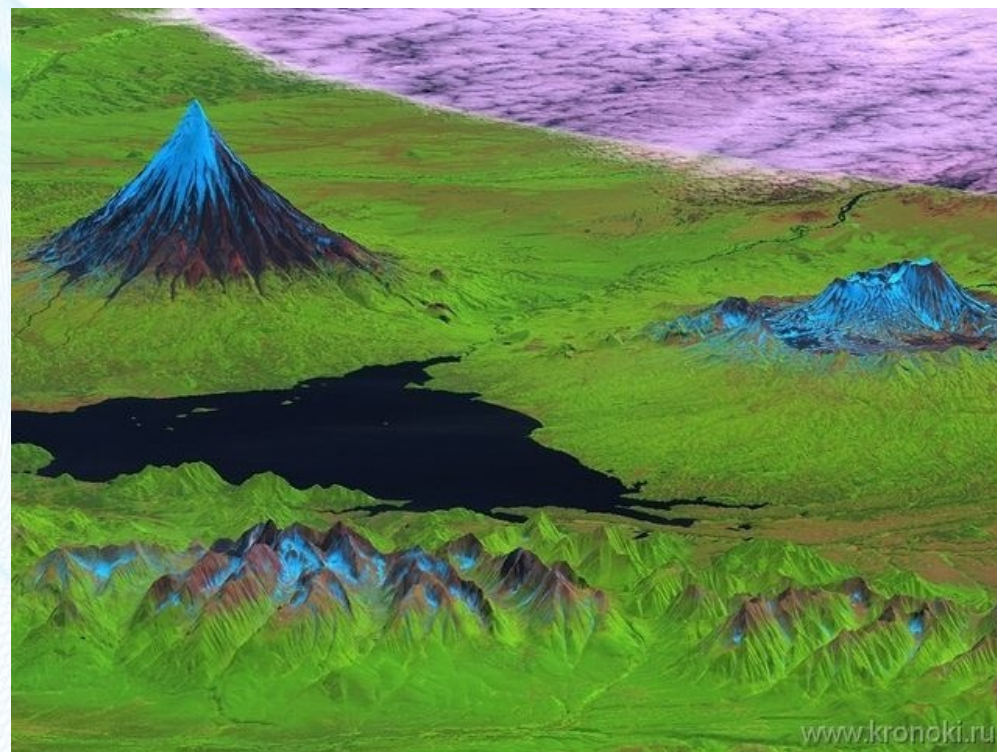
- Рельеф: SRTM, ASTER GDEM

Shuttle Radar Topographic Mission (SRTM), 2000

- разрешение 30/90 метров
- радарная интерферометрия
- охват: до 60 гр. с.ш. и ю.ш.

ASTER GDEM

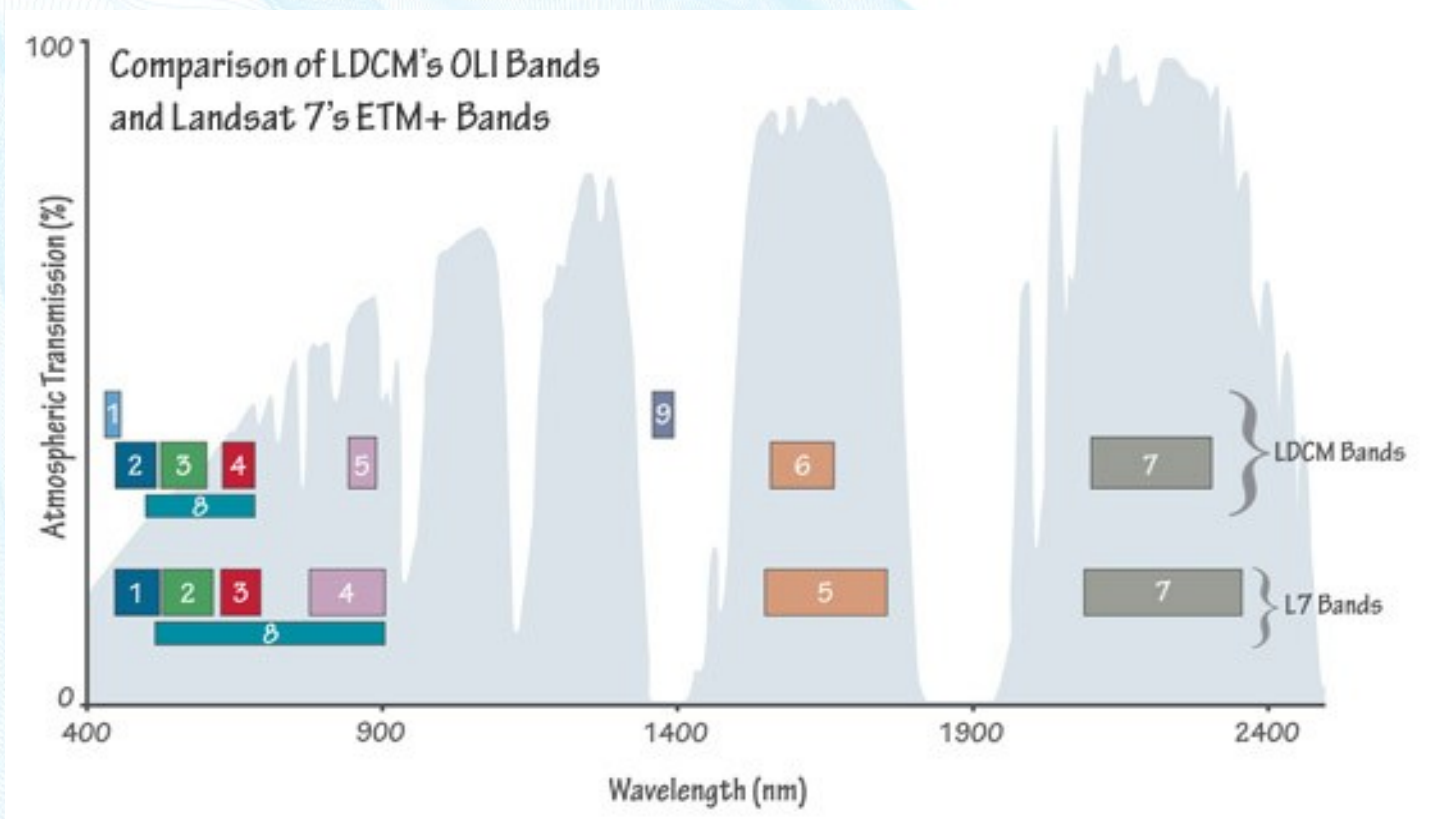
- разрешение 30 метров
- стереоскопическая съемка
- охват: без ограничений



Основные наборы данных

Дистанционное зондирование:

- Источники: наземная, авиационная (БПЛА), космическая,
- Диапазоны спектра: видимые, инфракрасные, РЛС, лидары
- Их количество: монохромная, мульти и гиперспектральная



Основные наборы данных

Дистанционное зондирование

OrbView-3

- 1 м/пиксел
- панхроматическая съемка
- в основном населенные пункты

TM/ETM+ (Landsat 4-7)

- 15,30,60 м/пиксел
- панхром + 6 спектральных каналов + тепловой
- съемка раз в 16 дней

MODIS (TERRA,AQUA)

- 250, 500, 1000 м/пиксел
- 32 спектральных канала
- 4 раза в сутки

Подводные камни

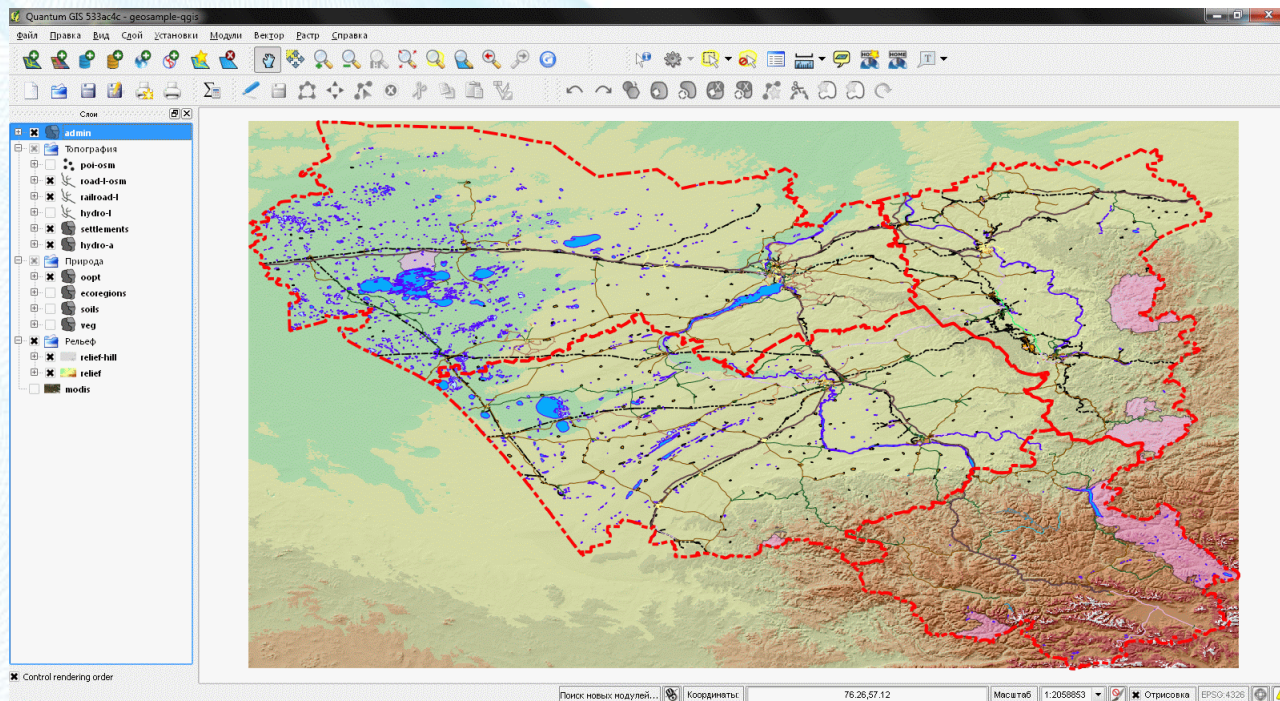
- Нет единых стандартов на получение и создание
- Неравномерность, неполнота, неактуальность
- Многие источники не валидированы
- Большая часть данных недоступна или плохо доступна: лицензионные, коммерческие, технические проблемы

Программное обеспечение

Какое базовое ПО нужно для работы с геоданными?

Программное обеспечение

- ПО ГИС общего назначения: Quantum GIS/GRASS
 - Визуализация данных
 - Обработка данных
 - Формирование геоданных
 - Работа с атрибутивной информацией



Программное обеспечение

- Библиотеки абстракции форматов: GDAL/OGR
- 130 форматов растровых данных (GeoTIFF, IMG, NetCDF, GRASS raster)
- 70 форматов векторных данных (DXF, Shape, GeoJSON, GML, GPX)
- Базовые алгоритмы для вектора и растра:
 - Intersects, Equals, Disjoint, Touches, Crosses, Within, Contains, Overlaps
 - Отмывка рельефа, генерация контуров по рельефу, трансформирование по контрольным точкам, полиномиальное преобразование, RPC ортотрансформирование, мозаицирование, объединение каналов, перепроецирование

Программное обеспечение

- Статистика: R
 - Язык и среда
 - Самый широкий набор пакетов для анализа пространственных данных
 - Не подходит для анализа больших наборов данных
 - Не подходит для сложных визуализаций

Программное обеспечение

- Пространственные базы данных: PostgreSQL/PostGIS, SQLITE/Spatialite
 - Вектор и растр
 - Топологические отношения
 - OGC-совместимость

Тенденции

- Что происходит в мире:
 - Новые средства получения данных: БПЛА, лидары, краудсорсинг
 - Новые средства доставки пространственных данных: sensor-web, geospatial semantic web
 - Открытый исходный код
 - Облачные вычисления и сервисы

Что делать?

- Перспективные направления:
 - Параллельная обработка данных (например ортокоррекция OrbView-3)
 - Открытые алгоритмы (например ландшафтные индексы, классификаторы и т.д.)
 - Открытое полное 3D
 - Пространственно-временной анализ геоданных
 - Обработка изображений с нестабильных платформ

Как?

- Совершенствование существующих алгоритмов
- Наводить мосты между дисциплинами:
 - Компьютерное зрение (OpenCV)
 - Медицинская визуализация (ITK)