

Сетевые аналитические ГИС в фундаментальных исследованиях

Н.А.Кузнецов, В.Г.Гитис

Институт проблем передачи информации, Российская академия наук, Москва, Россия
Поступила в редколлегию 25.08.2004

Аннотация—Даны основы подхода к сетевому анализу пространственно-временной географической информации. Рассматриваются примеры применения сетевых ГИС GeoПроцессор и КОМПАС в фундаментальных исследованиях, связанных с анализом пространственно-временных процессов и пространственным прогнозом по мультидисциплинарным данным.

1. ВВЕДЕНИЕ

Значительные успехи в телекоммуникационных технологиях и создание единого информационного пространства сети Интернет открывают новые возможности для развития фундаментальной науки (Кузнецов, 2001). Около 80% информационных потоков составляет географическая информация (ГИ) и ее объем экспоненциально растет. Для удаленного доступа к распределенным географическим данным, их комплексной обработки и анализа создаются сетевые аналитические географические информационные системы (ГИС). Они представляют собой интегрированную среду для визуального исследования, моделирования, анализа и прогноза пространственно-временных процессов, происходящих в природе и обществе.

В работе рассмотрены основы сетевого геоинформационного анализа, показаны примеры применения сетевых ГИС в научных исследованиях и намечены пути развития.

2. ГЕОГРАФИЧЕСКАЯ ИНФОРМАЦИЯ

Географический мир можно представить с помощью трех базовых понятий: *сущностей*, их *свойств* и *связей* (Рис. 1). Для анализа многих пространственно-временных процессов и явлений достаточно выделять два типа сущностей: *дискретные сущности*, представляющие географические объекты, и *гладкие непрерывные многообразия*, которые представляют свойства природной среды или пространственно-временные свойства множеств географических объектов. *Свойства* представляют собой понятия, которые характеризуют сущности. *Связи* представляют собой понятия, которые описывают отношения *между географическими объектами, между свойствами и между объектами и свойствами*.

В ГИС сущности отображаются в виде цифровых моделей. При этом для географических объектов используются так называемые векторные форматы данных, а для непрерывных многообразий – сеточные или растровые форматы. Для выполнения групповых операций данные структурируются в семантически однородные слои. Обычно исследования выполняются с набором данных, относящихся к одной пространственной области. Этот набор данных принято называть проектом.

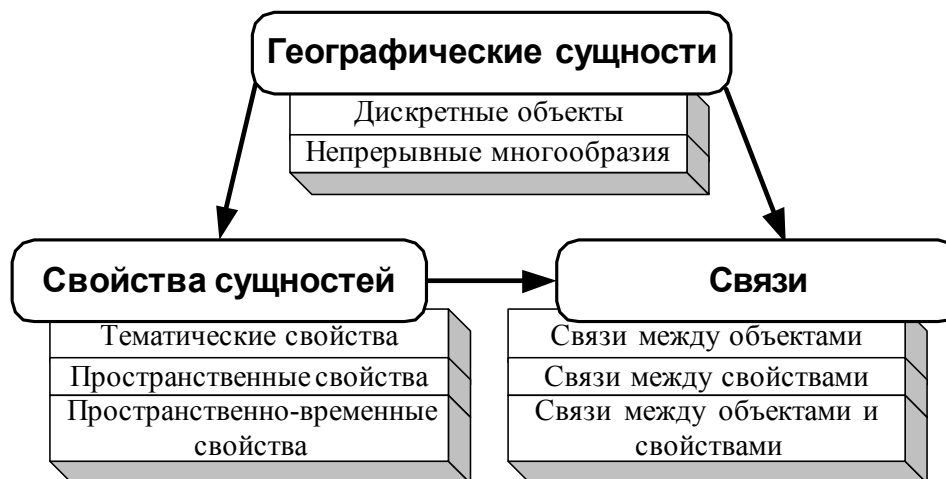


Рис. 1. Геоинформационное представление географического мира.

3. ТИПЫ ЗАДАЧ

Сложность решения задач геоинформационного анализа существенно зависит от полноты имеющейся информации.

Задачи с *полной информацией* обычно встречаются при ознакомлении с географическими понятиями по комплексу всех имеющихся данных. Анализ в таких задачах, как правило, состоит в выявлении качественных пространственных или пространственно-временных особенностей ГИ, в вычислении новых свойств ГИ с помощью заранее известных преобразований и в оценивании стандартных статистических характеристик векторных и сеточных данных.

Задачи с *неполной информацией* возникают при исследовании географических понятий и, в частности, при прогнозе объектов, явлений и пространственно-временных свойств геологической среды. Такие задачи являются очень сложными как с точки зрения формализации и структуризации исходной информации, так и ее дальнейшей обработки и анализа. В задачах с неполной информацией знания о прогнозируемом явлении часто носят описательный характер, инструментальные измерения параметров среды, прямо связанных с математическими моделями прогнозируемых явлений, отсутствуют, а имеющиеся инструментальные косвенные измерения сильно зашумлены. Сложность задач дополнительно возрастает в связи с тем, что инструментальные измерения некоторых параметров среды либо невозможны, либо требуют дополнительной интерпретации, не всегда однозначной, и поэтому часто заменяются экспертными оценками или дополняются экспертными решениями.

Можно сформулировать четыре основные проблемы, для которых разрабатываются методы аналитических геоинформационных технологий.

1. Оценивание и интерпретация связей между свойствами ГИ (например, картографическое исследование взаимного расположения элементов нескольких информационных слоев, оценивание статистических параметров свойств, в частности, оценивание прямых экономических потерь от землетрясений, оценивание запасов месторождений полезных ископаемых).
2. Оценивание и интерпретация отношений между географическими объектами (например, оценивание меры связи между природными процессами или географическими объектами, в частности, оценивание косвенных потерь от землетрясения).
3. Прогнозирование, обнаружение и распознавание целевых заранее неизвестных свойств ГИ и заранее неизвестных географических объектов (например, прогнозирование максимальных магнитуд ожидаемых землетрясений по комплексу геолого-геофизических данных, прогнозирование месторождений, обнаружение предвестников землетрясений).
4. Прогнозирование развития пространственно-временных процессов (например, моделирование сценариев развития взаимодействующих катастрофических процессов, вызванных землетрясением).

4. ПОТРЕБИТЕЛИ СЕТЕВЫХ АНАЛИТИЧЕСКИХ ГИС

Потребителей сетевых аналитических ГИС можно разделить на следующие группы:

- Непрофессиональные пользователи сети Интернет (ГИС-зрители).
- Администраторы всех уровней.
- Учащиеся школ, колледжей и университетов.
- Эксперты и исследователи.

Наиболее многочисленной является группа непрофессиональных потребителей ГИС. Запросы этой группы ГИС-зрителей относятся в основном к получению справочной, мало изменяющейся во времени информации из достаточно больших баз данных, либо к просмотру постоянно обновляемой динамической информации в реальном времени. Этой группе в основном необходимо обеспечить дружелюбный интерфейс к средствам поиска релевантной ГИС и наглядное представление ГИС. Поддержка запросов этой массовой группы пользователей очень важна для построения информационного общества.

Запросы остальных групп потребителей ГИС связаны с обучением, поддержкой принятия решений, производством и исследовательской работой. Обеспечение этих запросов требует существенно более детального анализа ГИС. Поэтому эти группы должны иметь с одной стороны возможность настройки аналитических ресурсов системы под свои задачи вплоть до подключения своих программных средств, а с другой - возможность подключения своих личных данных для их совместной обработки с данными сервера (в том числе и с данными, предназначенными только для абонированных пользователей).

Рассмотрим более подробно запросы основных групп пользователей и системные средства для их поддержки (*Гитис, Вайншток, 2001*).

Задачи группы *непрофессиональных пользователей сети Интернет* состоят в просмотре справочной ГИС, в получении ГИС в реальном времени и ее экспресс-анализе (Рис. 2). Для поддержки этих запросов достаточно обеспечить доступ к средствам картографии, когнитивной графики и мультимедиа.

Задачи группы *администраторов* добавляют к этому списку запросы по детальному анализу ГИС, по подготовке аналитических отчетов и по поддержке принятия решений. Для этого требуется подключение к ГИС средств визуального исследования ГИС с помощью картографии и когнитивной графики, а также средств вычисления новых свойств ГИС с помощью аналитических преобразований.

Задачи группы *учащихся* университетов, колледжей и школ могут относиться к изучению пространственно-временных процессов и явлений, происходящих в природе и обществе. Вместе с тем, задачи этой группы включают в себя и рассмотренные выше запросы. Необходимые дополнительные средства должны поддерживать комплексный анализ свойств ГИС и классификацию географических объектов, а также возможность подключения данных пользователя.

Задачи группы экспертов в различных предметных областях являются наиболее сложными. В дополнение к остальным группам пользователей задачи экспертов могут включать исследование закономерностей в географических данных, подготовку новых данных, а также разработку предметно-ориентированных технологий анализа ГИС. Эти запросы поддерживаются средствами оценивания связей между географическими понятиями, средствами изменения структуры ГИС, моделирования пространственно-временных процессов и средствами подключения программных модулей пользователя.

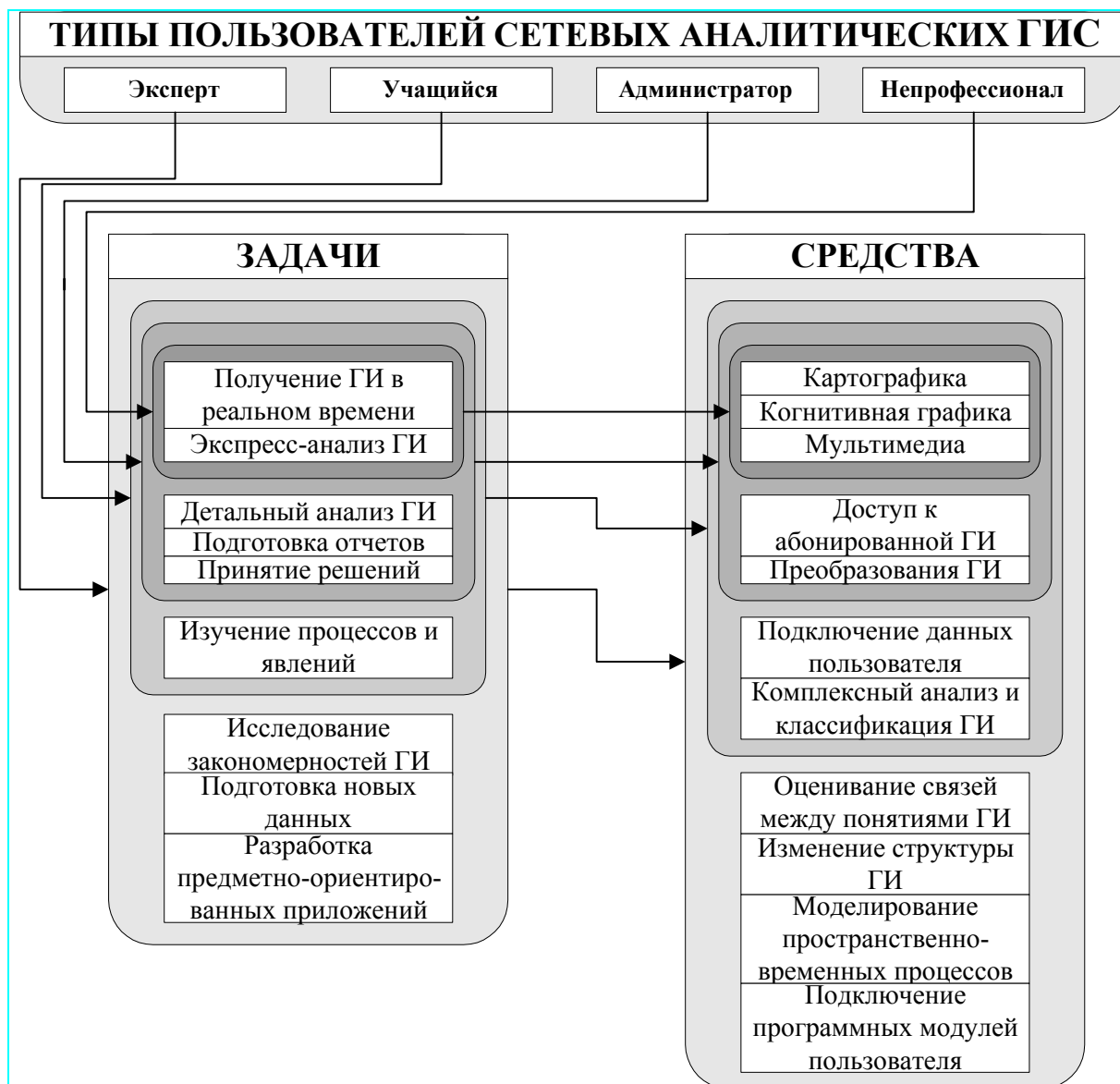


Рис. 2. Запросы и средства поддержки основных групп пользователей сетевых аналитических ГИС

5. СЕТЕВЫЕ АНАЛИТИЧЕСКИЕ ГИС

Рассмотрим две широко распространенных архитектуры сетевых ГИС: серверные системы и системы клиент-сервер.

Структура серверной схемы приведена на рис. 3. По запросу пользователя ГИ выбирается из удаленной БД и обрабатывается на ГИС сервере. Затем картографический результат преобразуется в растровое изображение и пересылается пользователю. Просмотр данных осуществляется с помощью стандартных Web-браузеров.

Эта технология, по-видимому, обеспечивает наиболее простой подход к публикации ГИ в сети Интернет: на сервере можно использовать настольные ГИС, поскольку отсутствует прямой доступ к базе данных, то данные достаточно просто защитить от несанкционированного доступа. Вместе с тем имеется ряд принципиальных ограничений: значительные времена ожидания на каждой транзакции при передаче большого растрового изображения от сервера к клиенту, низкая интерактивность при обслуживании динамических запросов и при выполнении операций, ограниченный интерфейс пользователя, ограниченное качество графики, обусловленное передачей данных в наиболее типовых форматах GIF и JPG.

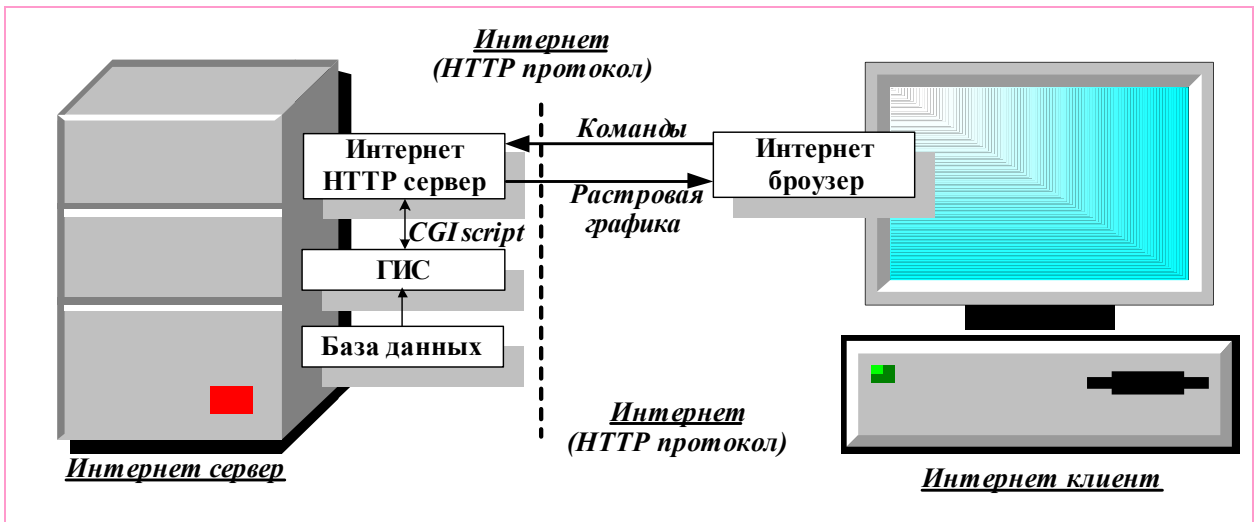


Рис. 3. Взаимодействие клиента с геосервером

Схема клиент-сервер (Рис. 4) позволяет обеспечить неограниченную интерактивность картографического исследования, пространственно-временной обработки и анализа ГИ. Часто клиентская часть сетевой ГИС реализуется в виде программы на языке Java и включается в Web-страницу. При загрузке такой страницы система автоматически загружается на компьютер пользователя и начинает работу. При этом ГИС может использовать данные, расположенные на различных серверах. Такие программы называются апплетами. Апплеты могут работать на всех платформах, на которых установлены стандартные средства сети Интернет. Быстродействие выполнения всех ГИС функций определяется только параметрами компьютера пользователя.

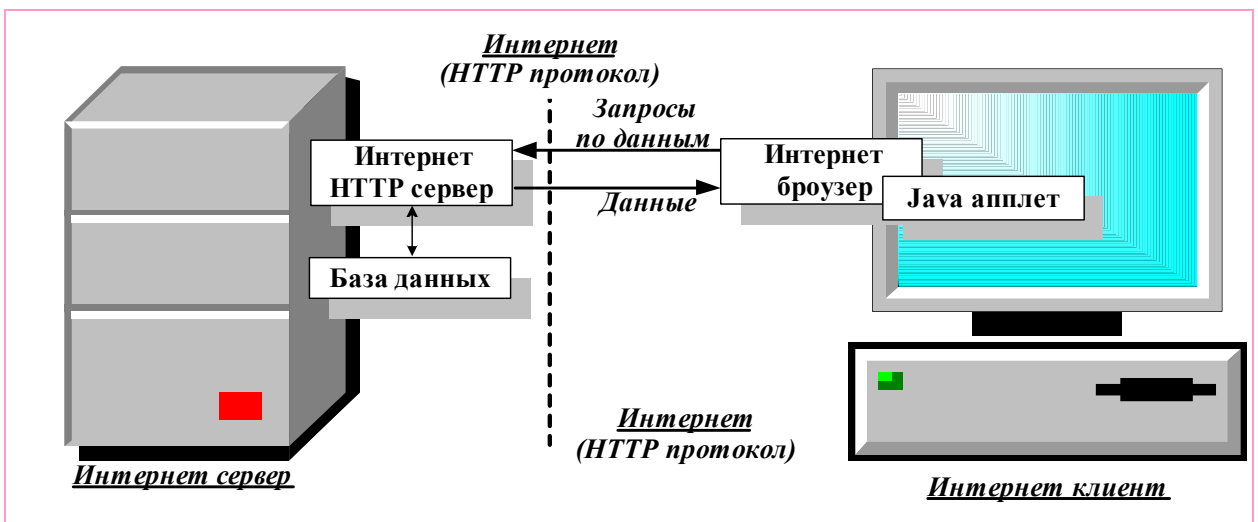


Рис. 4. Взаимодействие по схеме клиент-сервер.

Данная архитектура наиболее приемлема для создания аналитических сетевых ГИС. Ее существенными достоинствами являются высокая интерактивность анализа данных, динамическое управление средствами картографической и графической визуализации, высококачественная графика. Недостатки по сравнению с серверной схемой состоят в необходимости однократной затраты времени на загрузку Java апплета. Кроме того, поскольку данные передаются на компьютер клиента, то нельзя гарантировать их защиту от несанкционированного доступа.

Следует отметить также, что Java апплеты могут быть установлены на машине пользователя и работать в локальном режиме, обеспечивая доступ, как к удаленным информационным ре-

сурсам, так и к данным пользователя. Локальный режим работы позволяет разрабатывать новые ГИС приложения.

Примерами аналитических ГИС в такой архитектуре являются CommonGIS (Descartes) (*Andrienko and Andrienko, 1999, Andrienko et al, 1999, Andrienko and Andrienko 2002*), Lava Magma (*Berg et al, 1999*), ГеоПроцессор и КОМПАС (*Gitis et al, 1998, Gitis et al, 2000*).

6. СРЕДСТВА СЕТЕВЫХ АНАЛИТИЧЕСКИХ ГИС

В геоинформационных технологиях широко используются три метода анализа данных:

1. *Визуализация*.
2. *Аналитические преобразования*: применение заранее заданного оператора к имеющимся географическим данным для вычисления нового представления и новых свойств географической информации.
3. *Правдоподобный вывод*: нахождение заранее неизвестного оператора аналитического преобразования.

Основные отличия сетевых аналитических ГИС от традиционных настольных ГИС обусловлены, главным образом, *изменением типа пользователя от эксперта в области ГИС технологий к эксперту в предметной области*. Это обстоятельство накладывает дополнительные требования к основным операциям сетевых аналитических ГИС. С одной стороны, система должна включать достаточный набор функций, чтобы обеспечить комплексный анализ пространственно-временной информации, относящейся к данной предметной области. С другой стороны, эти функции должны реализовываться с помощью удобных и достаточно понятных для эксперта операций:

1. *Визуализация* должна обеспечивать неограниченную интерактивность при картографическом исследовании данных.
2. *Аналитические преобразования* должны включать в себя расширенный класс заранее подготовленных операторов, ориентированных на предметную область.
3. *Правдоподобный вывод* должен быть основан на интуитивно понятных методах интеллектуальной обработки пространственно-временных данных.

Визуализальное исследование данных поддерживается набором интерактивных картографических и когнитивно-графических операций, которые направлены на выявление образа пространственного распределения свойств исследуемой области, выявление взаимного пространственного распределения нескольких географических слоев и выполнение картографических измерений. Интерактивность позволяет динамически управлять изменением закрашки карт, размерами пиктограмм и трехмерным представлением ГИ, динамически изменять отображение данных на диаграммах при пересчете значений показателей, динамически выделять на карте группы объектов в скользящем временном интервале (например, эпицентры землетрясений), одновременно подсвечивать группы анализируемых объектов в картографических, графических и табличных окнах.

Аналитические преобразования позволяют по имеющимся данным вычислить новые тематические и пространственные свойства географической информации с помощью заранее заданных операторов. В сетевых аналитических ГИС ГеоПроцессор, КОМПАС (ИППИ РАН, <http://www.iitp.ru/projects/geo>, *Gitis, V. u др., 1999, Gitis V. u др., 2000*) и CommonGIS (Fraunhofer AiS при участии ИППИ РАН, <http://www.ais.fhg.de/and>, *Andrienko, et al, 2003*) реализованы четыре типа аналитических преобразований:

1. Преобразование **Сеточные слои** \Rightarrow **Сеточный слой**: например, растровая фильтрация, вычисление произвольных функций от одного или нескольких растров.
2. Преобразование **Сеточные слои и Векторный слой** \Rightarrow **Атрибуты векторного слоя**: например, вычисление статистик сеточных слоев для буферных зон точечных, линейных и зональных географических объектов.
3. Преобразование **Векторный слой** \Rightarrow **Сеточный слой**: например, вычисление растра (взвешенной) плотности точек или линий, вычисление растра близости к географическим объектам.

4. Преобразование **Векторные слои** \Rightarrow **Атрибуты векторного слоя**: например, вычисление метрических свойств географических объектов, вычисление произвольных функций от имеющихся атрибутов слоя.

Методы правдоподобного вывода предназначены для нахождения операторов аналитических преобразований на основании имеющихся данных и знаний. В сетевых ГИС GeoProcessor, КОМПАС и CommonGIS для вывода используются методы построения карт прогноза по комплексу исходных признаков. Эти методы основаны на использовании решающих правил, найденных с помощью правдоподобного вывода по сходству, по функции принадлежности к классам объектов, по эмпирической функции распределения и по функции непараметрической регрессии.

7. ПРИМЕРЫ АНАЛИЗА ПРОСТРАНСТВЕННО-ВРЕМЕННЫХ ДАННЫХ

7.1. Оценка возможного ущерба от землетрясений (Западная Турция)

Метод применения сетевых аналитических ГИС для поддержки принятия решений по оценке прямого ущерба от возможных сильных землетрясений основан только на применении различного рода аналитических преобразований. Иллюстративный пример выполнен с помощью системы CommonGIS для региона Западной Турции (*Andrienko, Andrienko, and Gitis, 2003*). В примере использованы исходные данные, взятые в сети Интернет¹. Этим данным явно недостаточно для постановки задачи по оценке прямого и косвенного сейсмического риска. Вместе с тем, пример показывает на возможность использования системы CommonGIS для поддержки принятия решений для планирования мероприятий по снижению сейсмического риска. Пример выполнен только для одного типа застройки (сооружения типа А) и для одного типа разрушений: сильные повреждения, не подлежащие восстановлению. Обобщение предложенного метода на все типы сооружений и разрушений является очевидным.

Решение строится на применении аналитических преобразований над сеточной цифровой моделью пиковых ускорений a на поверхности Земли, полученной в рамках проекта GSHAPE (*Giardini et al, 1999*), и векторными данными по инфраструктуре: координаты главных городов и административных районов Западной Турции.

При решении по сеточной модели пиковых ускорений a была вычислена сеточная модель доли D строений типа A , которые должны получить сильные повреждения, не подлежащие восстановлению. Для этой цели применялись два преобразования типа **Сеточные слои** \Rightarrow **Сеточный слой**. Сначала выполнялась операция вычисления максимальной интенсивности землетрясений I из уравнения $\log(a)=0.3I+0.014$ (*Trifunac and Brady, 1975*). Затем выполнялась операция линейной интерполяции зависимости $D(I)$, которая опубликована в (*Coburn, 1995*). Зависимость $D(I)$ интерполировалась по следующим точкам: $D(I<5)=0$, $D(I=6)=0.05$, $D(I=7)=0.5$, $D(I=8)=0.75$, $D(I>9)=0.9$.

Средняя доля сильных разрушений строительных сооружений типа A для крупных городов и административных районов оценивалась в предположении о равномерном распределении застройки в радиусе 10 км вокруг центра города и в пределах административных районов. Для оценивания использовались преобразования **Сеточные слои** и **Векторный слой** \Rightarrow **Атрибуты векторного слоя** и интерактивные средства визуального анализа данных. Полученные результаты картографически представлены на рис. 5. На рисунке можно видеть многослойную карту исследуемого региона и ее легенду. Карта представляет цифровую модель рельефа, на которую нанесен полупрозрачный слой доли тяжелых разрушений застройки типа A . Жирными линиями показаны границы административных районов. Высоты столбиков соответствуют доле разрушений в городах, а размеры кружков – доле разрушений в административных районах. Наибольшая доля разрушений среди городов равна 33% и относится к г. Маниса (доля разрушений для этого административного района равна 26%). Наибольшая доля разрушений среди административных районов равна 37% и относится к району Кутайа.

¹ Используются также табличные данные и зависимости, подготовленные в Центральном Геофизическом Центре, г. Потсдам, в рамках проекта IST Program (EU IST - 10536): "Spatial Mining for Data of Public Interest (SPIN!)"

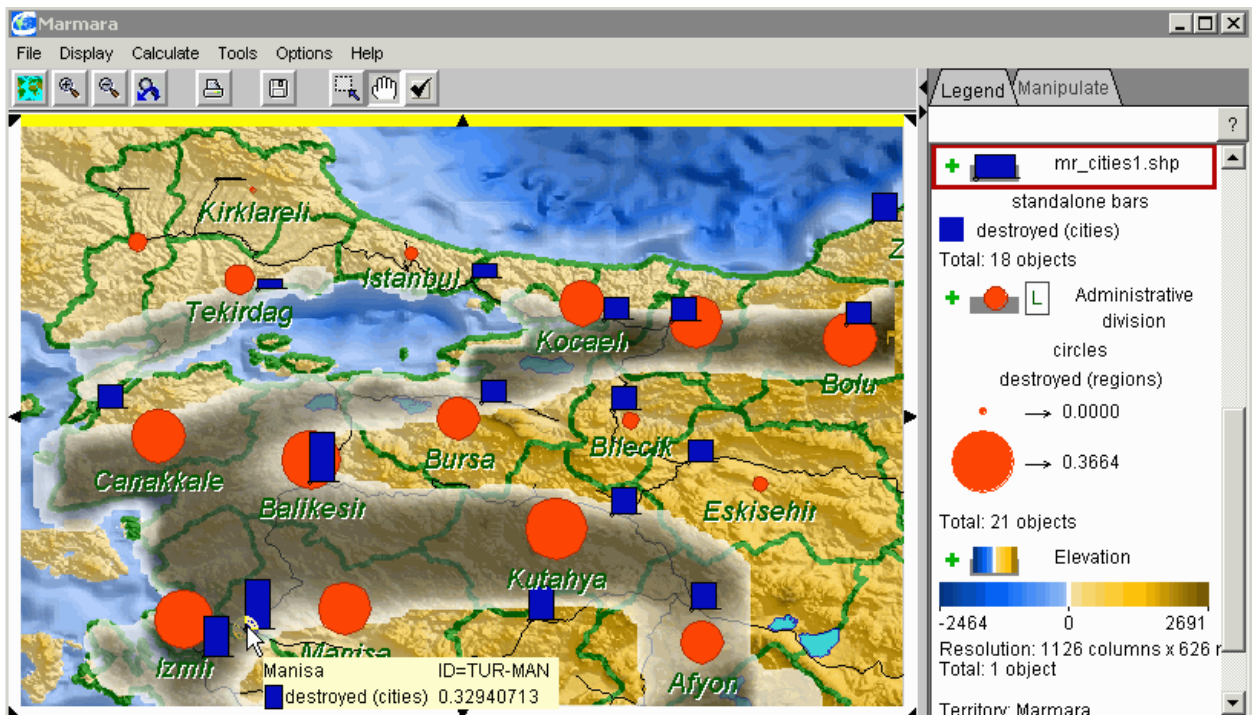


Рис 5. Возможные сильные разрушения застройки типа А в Западной Турции: высоты столбиков соответствуют доле разрушений в городах, а размеры кружков – доле разрушений в административных районах.

7.2. Анализ чувствительности сейсмической сети (Греция)²

В данном примере анализ каталога землетрясений выполнен с помощью настольной ГИС ГеоТайм (*Gitis, et al, 1995*), а растровые операции – с помощью сетевой ГИС ГеоПроцессор. На рис 6. показано распределение количества ежегодных землетрясений Греции с магнитудой $m_l < 4.0$. Можно выделить 3 временных интервала 1964-1980, 1981-1994 и 1995-2003, которые показывают заметное увеличение числа зарегистрированных землетрясений.

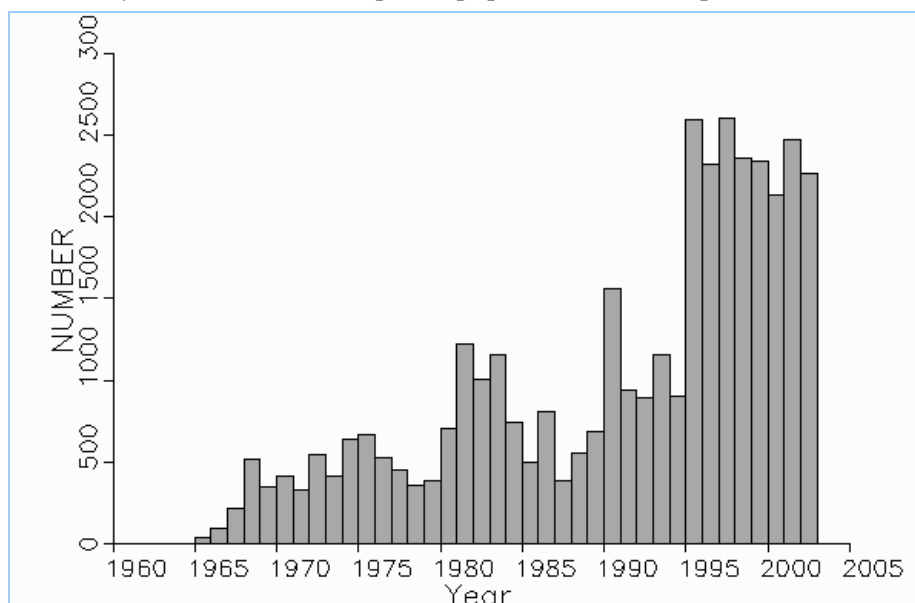


Рис. 6. Гистограмма числа ежегодных землетрясений с $m_l < 4.0$ с 1965 по 2003 гг., 38242 событий

² Использованы данные Национальной Обсерватории г. Афин, подготовленные в рамках двустороннего сотрудничества.

На рис. 7, 8, 9 показаны карты минимальных представительных магнитуд землетрясений для интервалов 1964-1980 (8494 событий), 1981-1994 (14395 событий) и 1995-2002 (20534 событий).

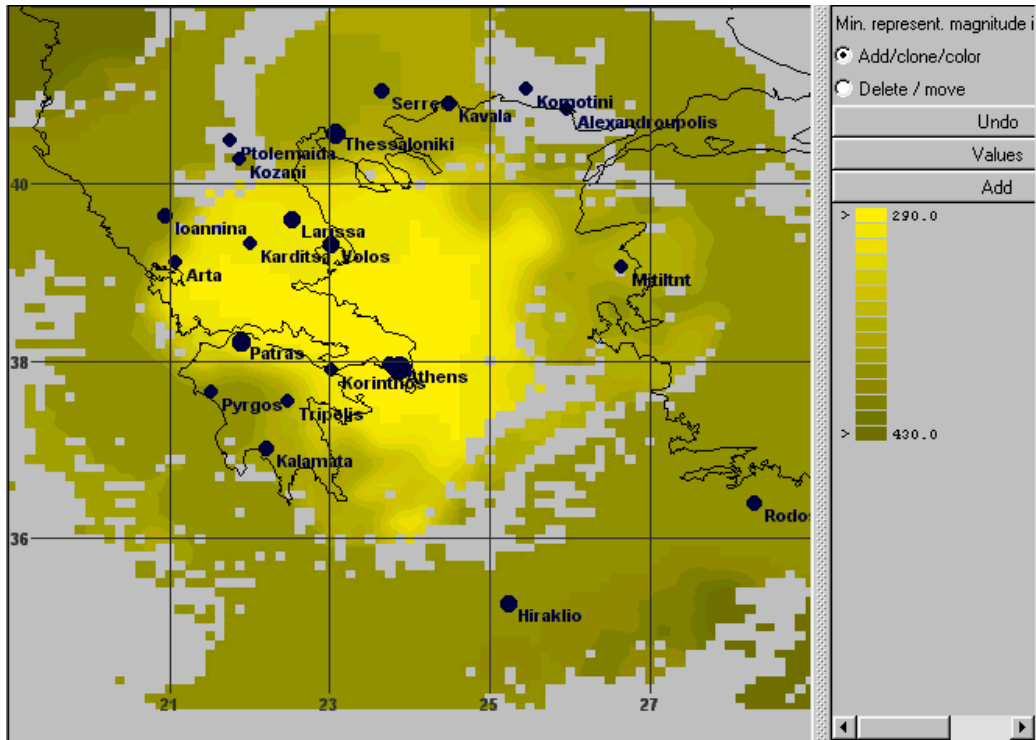


Рис. 7. Пространственное распределение минимальной представительной магнитуды землетрясений (магнитуда умножена на 100) за 1964-1980, 8494 события.

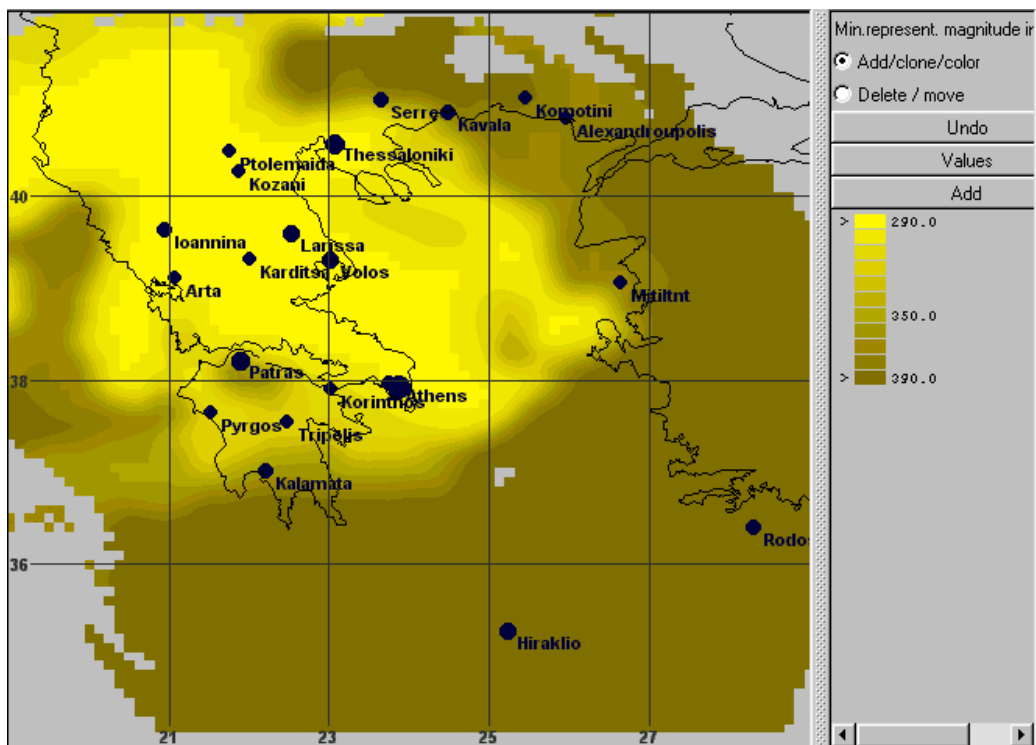


Рис. 8. Пространственное распределение минимальной представительной магнитуды землетрясений (магнитуда умножена на 100) за 1981-1994, 14395 событий.

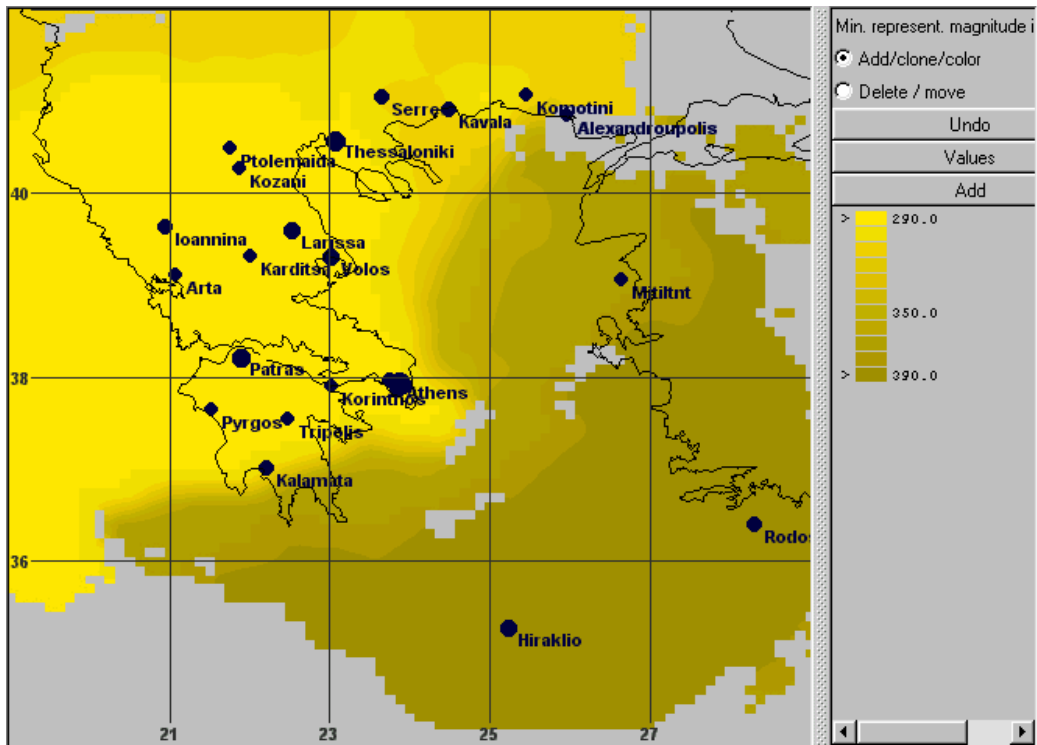


Рис. 9. Пространственное распределение минимальной представительной магнитуды землетрясений (магнитуда умножена на 100) за 1995-2002, 20543 событий.

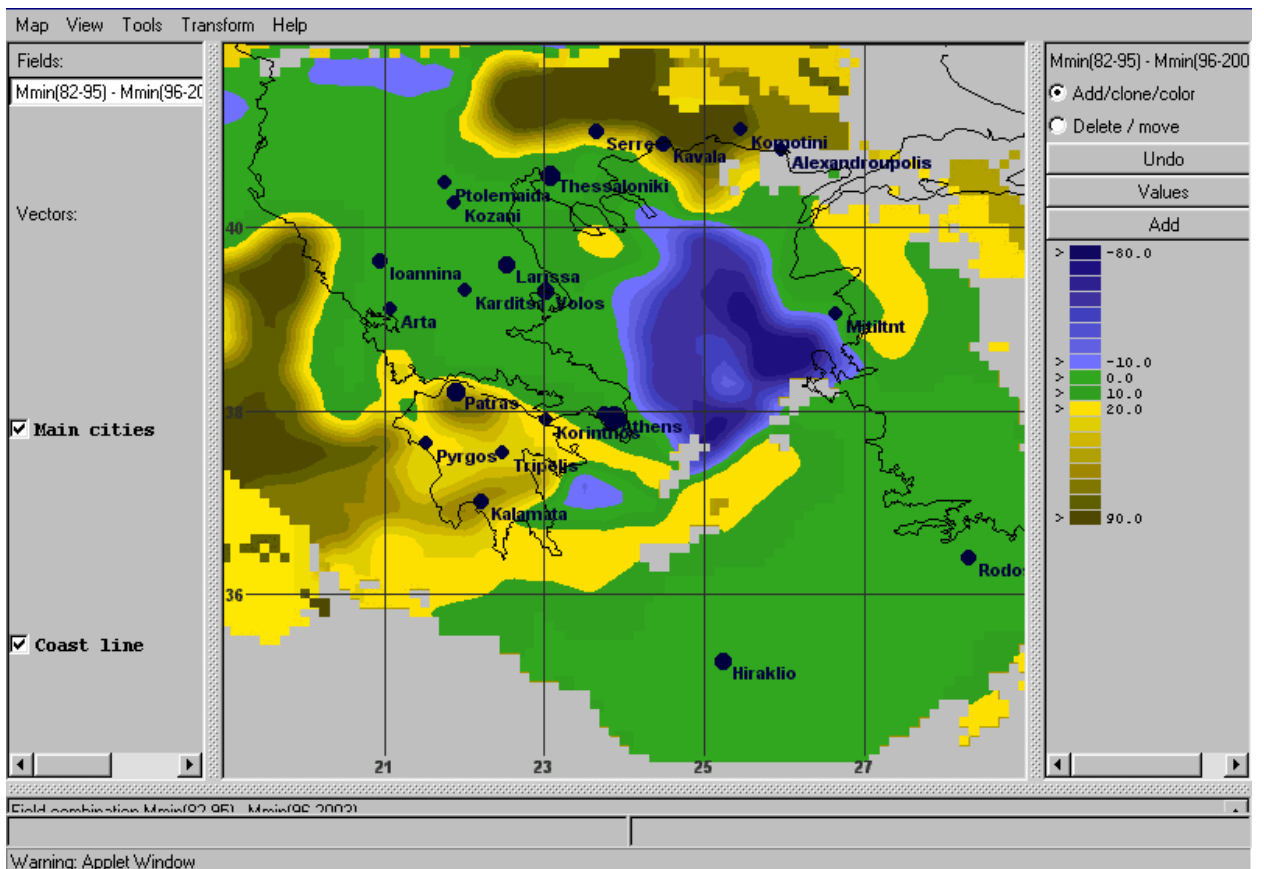


Рис. 10. Пространственное распределение разности между минимальными представительными магнитудами за периоды 1981-1994 и 1995-2002 (магнитуды умножены на 100): синие тона - чувствительность уменьшается, зеленые тона - чувствительность почти не меняется, желто-коричневые тона - чувствительность возрастает.

На рис.10 показана найденная с помощью преобразования **Сеточные слои** \Rightarrow **Сеточный слой** разность между минимальными представительными магнитудами за периоды 1981-1994 и 1995-2002 (магнитуды умножены на 100). Можно видеть, что пространственная чувствительность сейсмической сети изменялась неодинаково: на Юго-Западе и Северо-Востоке Греции минимальная представительная магнитуда уменьшилась на 0.2 - 0.9. В то же время в Эгейском море минимальная представительная магнитуда увеличилась на 0.1 - 0.8.

7.3. Пространственно-временной анализ шумов в каталоге землетрясений (Турция)³

В данном примере анализ каталога землетрясений выполнен с помощью настольной ГИС ГеоТайм (*Gitis et al*, 1995), а растровые операции – с помощью сетевой ГИС CommonGIS. На рис 11. показана гистограмма количества ежегодных землетрясений Западной Турции с магнитудой $m < 3.1$. Можно видеть, что, начиная с 1990 по 1999 годы, число регистрируемых событий существенно увеличилось. Можно ожидать, что увеличение чувствительности сейсмической сети повлечет за собой уменьшение минимальной представительной магнитуды. Однако гистограмма числа событий с 1990 по 1999 годы по времени суток (рис. 12) показывает неоднородность распределения числа событий по времени суток, выраженную в существенном увеличении числа событий в дневные часы.

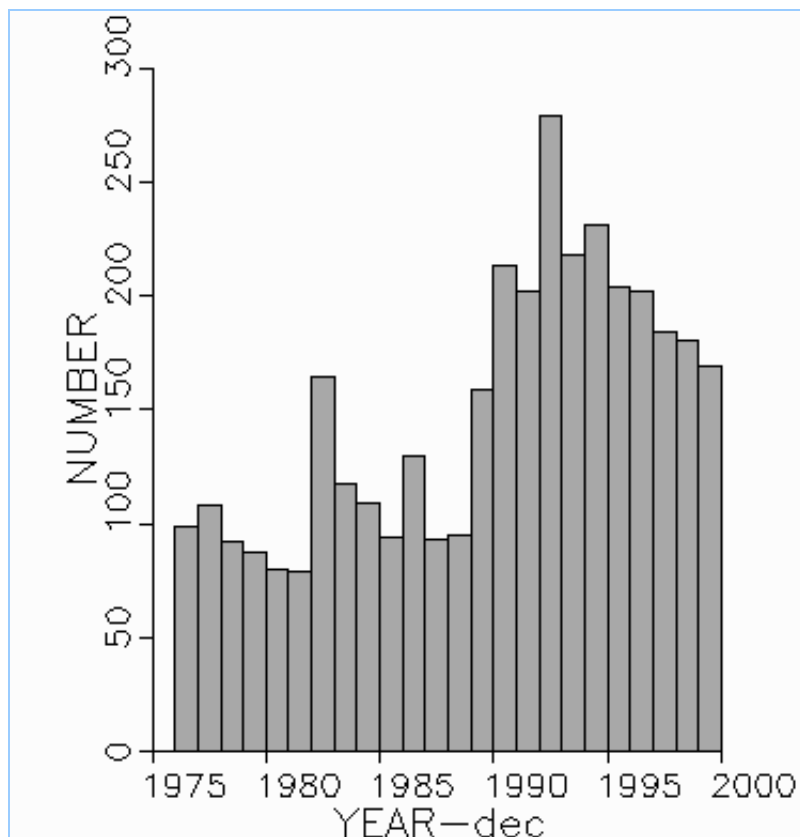


Рис. 11. Гистограмма числа ежегодных землетрясений с $m < 3.1$, 3588 событий с 1976 по 1999 гг.

³ Использованы данные, подготовленные Центральным Геофизическим Центром, г. Потсдам, в рамках проекта IST Program (EU IST - 10536): "Spatial Mining for Data of Public Interest (SPIN!)"

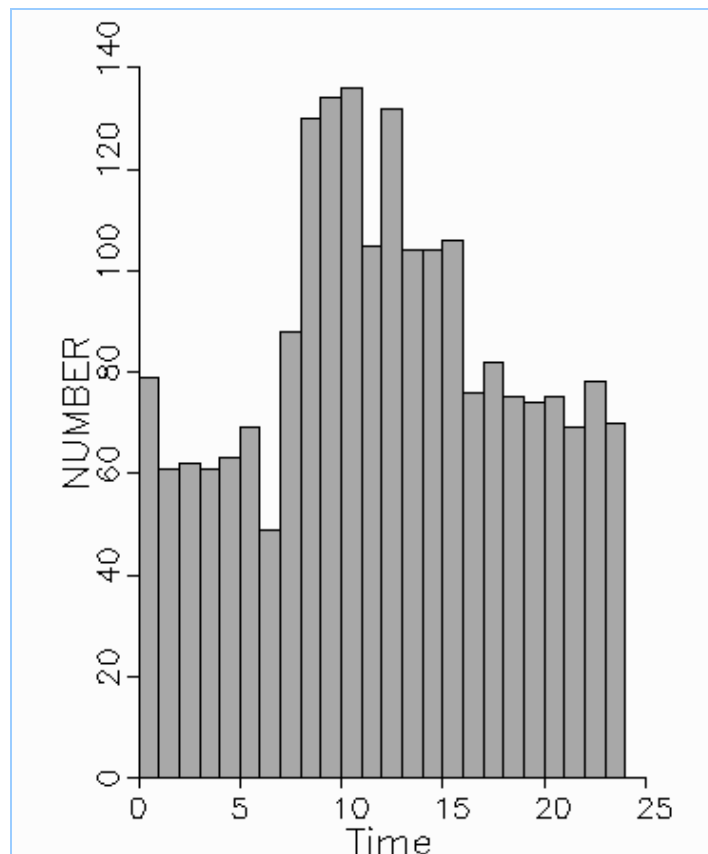


Рис. 12. Гистограмма числа землетрясений с $m < 3.1$ по времени суток, 2082 события с 1990 по 1999 гг.

Для пространственного анализа по каталогу землетрясений за период с 01.01.1990 по 31.12.1999 построены две сеточные модели, в которых для каждого узла регулярной координатной сетки подсчитано число событий, произошедших в круге радиуса $R=30$ км. В первой модели учитывались только события, которые произошли на временном интервале t_d от 8 до 15 часов, а во второй – те события, которые произошли на интервале $t_n = T - t_d$, $T=24$ часа, покрывающем остальное время суток. Пусть некоторому узлу сетки на интервале t_d соответствует N_d событий, а на интервале t_n – N_n событий. Рассмотрим гипотезы H_0 и H_1 . При гипотезе H_0 будем предполагать, что число событий, отнесенных к узлам сетки, имеет биномиальное распределение с вероятностями $p = t_n/T$, $q = 1 - p$. При альтернативной гипотезе H_1 будем предполагать

$q > 1 - p$. Для проверки гипотез используем статистику $v = \frac{N_d - Nq}{\sqrt{Npq}}$, которая при выполнении

гипотезы H_0 в асимптотике имеет стандартное нормальное распределение. На рис. 13 показана цифровая модель рельефа Западной Турции, поверх которой желтыми и коричневыми тонами показаны четыре зоны, для которых гипотеза H_0 отвергается не менее, чем с 1% уровнем значимости. Таким образом, с высокой степенью уверенности можно утверждать, что аномалия в увеличении числа событий в дневное время суток вызвана событиями, приуроченными к выделенным зонам. Скорее всего, эти события не связаны с сейсмичностью, а имеют техногенное происхождение.

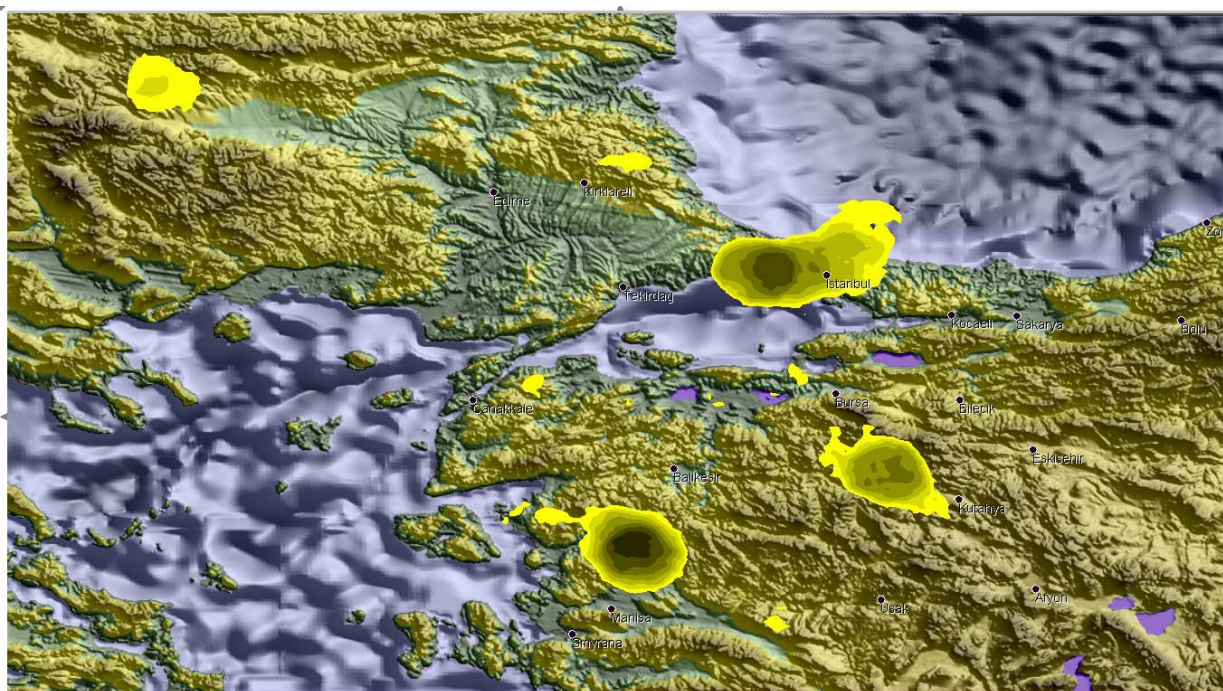


Рис 13. Рельеф Западной Турции, желтыми и коричневыми тонами показаны зоны, для которых гипотеза H_0 отвергается не менее, чем с 1% уровнем значимости.

7.4. Выделение возможных зон возникновения землетрясений (Болгария)⁴

Пример выделения зон возможного возникновения землетрясений по комплексу признаков выполнен с помощью ГИС ГеоПроцессор для области Болгарии, в которой 4.4.1904 произошло сильное землетрясение $M=7.8$ с эпицентром вблизи г. Кресна. Для этого района мы располагали следующими сеточными и векторными данными:

1. Рельеф поверхности Земли с детальностью 30"x30" (рис 14а).
2. Рельеф поверхности Мохоровичича с детальностью 4'x4' (рис 14b).
3. Гравитационные аномалии в редукции Буге с детальностью 4'x4' (рис 14с).
4. Разломы, ранжированные по степени активности в четвертичное время (рис 14d).
5. Каталог землетрясений с 1900 по 1997 годы, представительный с магнитуды $M=1.8$, начиная с 1985 года.

При решении были приняты следующие три предположения:

1. Повторяемость землетрясений: землетрясения могут повторяться в тех местах, где они уже произошли. Потому зарегистрированные землетрясения задают нижнюю границу максимальных магнитуд землетрясений. Однако выделение областей только по максимальным зарегистрированным землетрясениям дает заниженную оценку из-за короткого интервала наблюдений. В каталоге зарегистрированы сильные землетрясения с магнитудами M до 7.8. Следовательно, вблизи очаговых зон зарегистрированных землетрясений возможны землетрясения с такими же магнитудами в будущем.

⁴ Используются данные, подготовленные в рамках проекта INCO-COPERNICUS (IC 15 CT97 0200): "Assessment of Seismic Potential in European Large Earthquake Areas (ASPE-LEA)".

2. Квазистационарность сейсмического процесса: максимальная магнитуда землетрясений зависит от тектонических свойств геологической среды, которые медленно изменяются во времени.
3. Геологическая обусловленность: очаги землетрясений с $M=7-8$ приурочены к глубинным неоднородностям земной коры, а с $M=5-6$ приурочены к менее крупным неоднородностям, которые могут проявиться в рельефе поверхности Земли. Пересечение зон неоднородности земной коры с зонами влияния активных в четвертичное время разломов выделяет области с наибольшими тектоническими напряжениями.

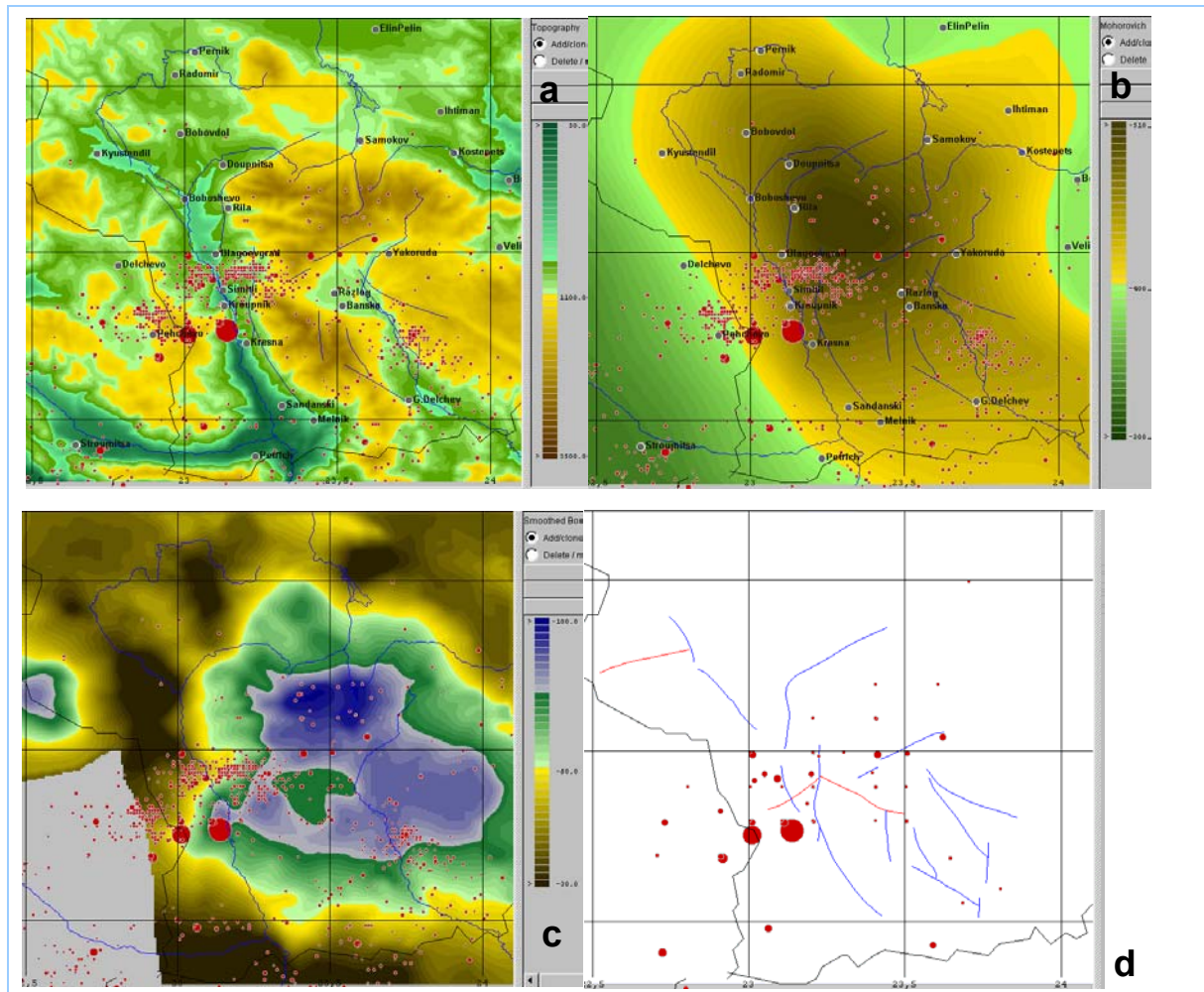


Рис. 14. Карты исходных данных: а) Рельеф поверхности Земли и эпицентры землетрясений $M > 1.8$, 1900-1997 гг.; б) Рельеф поверхности Мохоровичича; в) Гравитационные аномалии в редукции Буге; д) Разломы, активные и частично активные в четвертичное время и эпицентры землетрясений с магнитудой $M \geq 4.5$ с 1900 г по 1997 гг., афтершоки удалены

Признаки прогноза искали в соответствии с предположением 3. Зоны глубинных неоднородностей коры искали по сеточным моделям рельефа поверхности Земли, поверхности Мохоровичича и аномалиям силы тяжести с помощью преобразований типа **Сеточные слои** \Rightarrow **Сеточный слой**, таких как вычисление производных по нескольким направлениям, вычисление модулей градиентов, вычисление максимального разброса значений и среднеквадратичного отклонения (с.к.о) в скользящих окнах с радиусами $R=2.5, 5, 7.5$ и 10 км. Зоны влияния активных разломов находились с помощью преобразований **Векторный слой** \Rightarrow **Сеточный слой**, таких как вычисление плотности линий, близости к линиям и суммарной длины линий в скользящих окнах с радиусами $R=5, 10, 15, 20$ и 25 км. В результате картографического анализа были отобраны следующие 3 признака:

1. Сеточная модель средне-квадратичного отклонения (с.к.о.) значений поверхности Мохоровичича в скользящем окне радиуса $R=7.5$ км (x_1). Из рис. 15а видно, что признак x_1 выделяет несколько областей, в одной из которых происходили землетрясения с $M=7-7.8$.
2. Сеточная модель с.к.о. поверхности рельефа Земли в окне радиуса $R=7.5$ км (x_2). Из рис. 15б видно, что признак x_2 является менее избирательным, но он достаточно хорошо выделяет область, где происходили землетрясения с $M=5-5.5$.
3. Сеточная модель близости к разломам, частично активным в четвертичное время (рис. 15с), вычисленная с окном $R=20$ км (x_3). Близость узла сетки к линии определяется как $S = 1 - \rho(u, l) / R$, при $\rho(u, l) < R$, $S = 0$ при $\rho(u, l) \geq R$, где $\rho(u, l)$ - евклидово расстояние от узла сетки u до множества линий.

Решающее правило искалось по двум выборкам, составленным по эпицентрам землетрясений с $M \geq 4.5$ (рис. 15d). Можно предположить, что увеличение неоднородности геологической среды по с.к.о. поверхности Мохоровичича или поверхности рельефа Земли или большая близость к середине зоны активного разлома при прочих равных условиях увеличивают уверенность в возможности накопления и разрядки высокого сейсмического потенциала. Таким образом, предполагается, что отобранные признаки обладают свойством положительной монотонности, состоящей в том, что увеличение значения любого из признаков при прочих равных условиях увеличивает или не изменяет (но не уменьшает) степень уверенности в возможности того, что данный набор признаков относится к очаговой зоне сильных землетрясений. Можно видеть, что при условии положительной монотонности в качестве функции уверенности можно использовать эмпирическую функцию распределения. Пусть $\{\mathbf{x}^{(n)}\} \subset R^l$, $n = 1, \dots, N$, обозначает множество прецедентов. Эмпирическая функция распределения имеет вид $F(\mathbf{x}) = v_{\mathbf{x}} / N$, где $v_{\mathbf{x}}$ обозначает число прецедентов, для которых $x_i^{(n)} < x_i$ по всем $i=1, \dots, l$.

На рис. 16а, 16б показаны Западная область, выделенная по функции уверенности $F_1(\mathbf{x})$ на выборке 1 и на признаках x_1, x_3 , и Восточная область, выделенная по функции уверенности $F_2(\mathbf{x})$ на выборке 2 и на признаках x_2, x_3 .

При принятии решения на относительно небольших обучающих выборках наиболее важен детерминистский аспект функции уверенности: граница области признаков $A = \{\mathbf{x}: F(\mathbf{x}) > 0\}$, на которой функция принимает положительные значения. Для объяснения решения может быть выведена сейсмо-тектоническая интерпретация этой границы в виде не очень сложных логических высказываний (Гитис и др., 2004).

Решению по Западной области со степенью уверенности 0.75 (рис 16с) соответствует следующее правило:

если (с.к.о. поверхности Мохоровичича ($R=7.5$ км) больше 16.52 км)
и (близость к разломам ($R=20$ км) больше 0.39)
или (с.к.о. поверхности Мохоровичича ($R=7.5$ км) больше 16.94 км)
и (близость к разломам при $R=20$ км больше 0.33),
то возможны землетрясения с магнитудами $M=7-8$.

Восточной области со степенью уверенности 0.75 (рис 16d) соответствует следующее правило:

если (с.к.о. поверхности рельефа Земли ($R=7.5$ км) больше 330 м)
и (близость к разломам при $R=20$ км больше 0.25),
то возможны землетрясения с магнитудами $M=5-6$.

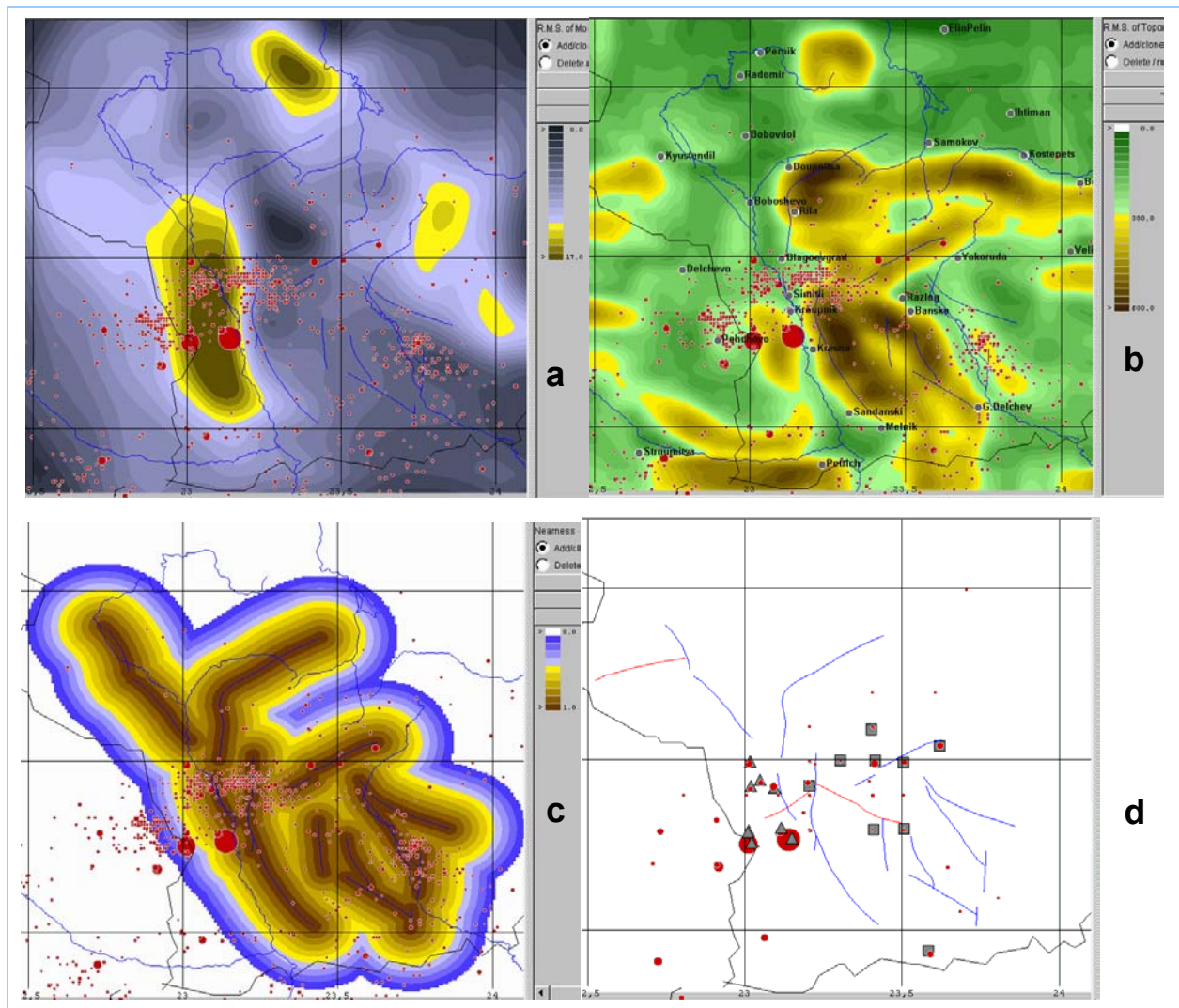


Рис 15. Геолого-геофизические признаки и выборки прецедентов для областей 1 и 2: а) Карта средне-квадратичных отклонений (с.к.о.) значений поверхности Мохоровичича в текущем окне $R=7.5$ км. б) Карта с.к.о значений рельефа поверхности Земли в текущем окне $R=7.5$ км. в) Карта близости к разломам, частично активным в четвертичное время с $R=20$ км. д) Выборки прецедентов 1 (треугольники) и 2 (квадраты).

Интерпретация полученного решения состоит в следующем. Можно предположить, что выделенная Западная область 1 характеризует напряженное состояние верхней мантии, которое выражено высокой дисперсией поверхности Мохоровичича. Эта область могла генерировать активные разломы и являться причиной наличия очагов с большими энергиями землетрясений. Выделенная Восточная область 2 относится к очаговым зонам землетрясений с меньшими энергиями. Эти очаги могут быть связаны с неоднородностями верхней части Земной коры, которые выражены в рельефе поверхности Земли и зонах активных разломов.

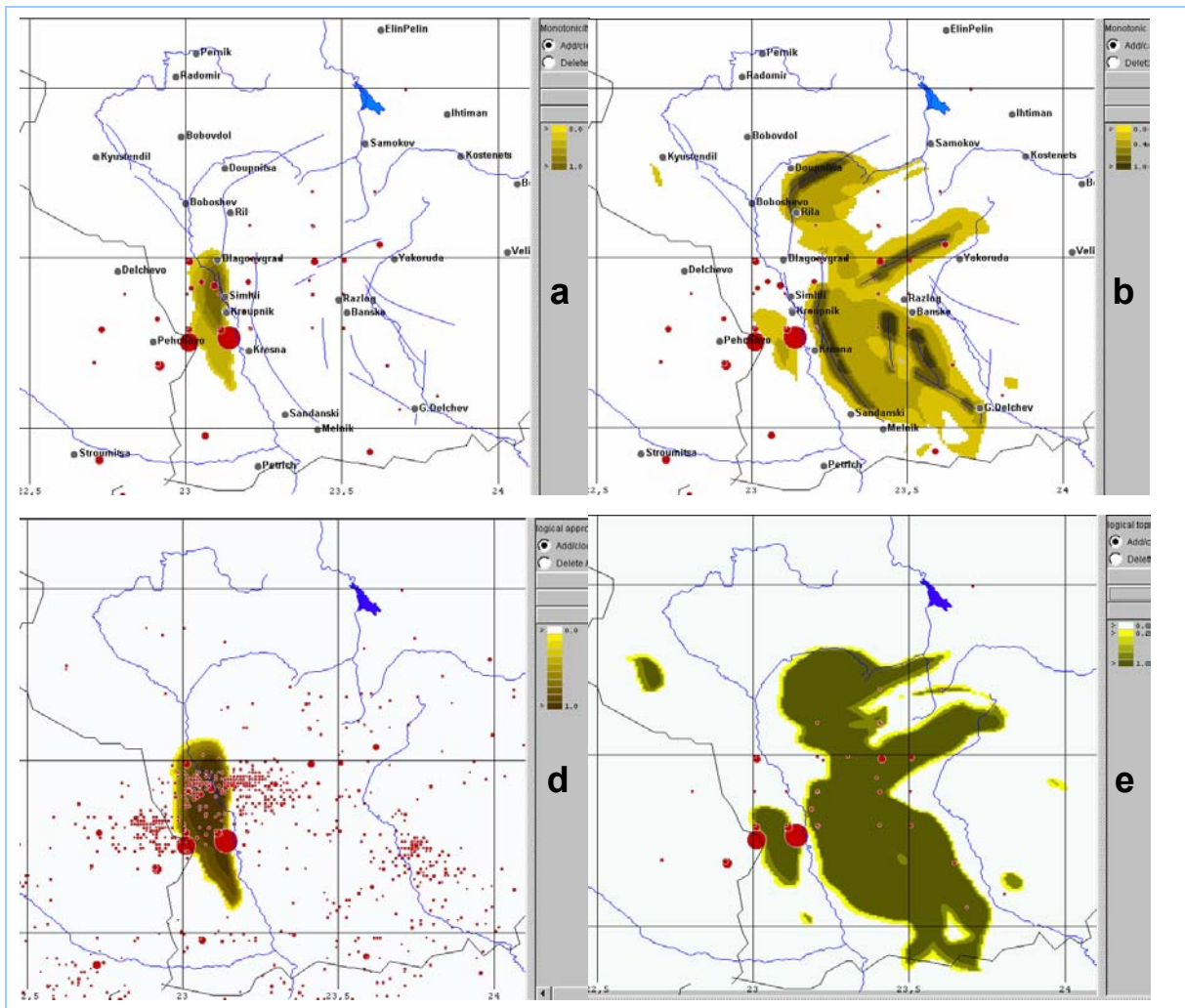


Рис 16. Результаты районирования по прецеденту и функциям нечеткой логики: а) и б) Области, выделенные по функциям уверенности $f_1(x)$ и $f_2(x)$ соответственно. с) и д) Области, выделенные по функциям нечеткой логики для $\Psi_1(x)$ и $\Psi_2(x)$ соответственно.

7.5. Природопользование: районирование России по структуре сельских земель, находящихся в частной собственности.

Рассмотрим один из примеров применения ГИС КОМПАС для анализа структуры использования сельскохозяйственных земель России.

В примере использованы данные по природопользованию для субъектов РФ (Думнов, Рыбальский и др., 2001), в частности данные за 1998 г. по площадям сельских земель, находящихся в частной собственности: площади фермерских (ξ_1), приусадебных (ξ_2) и дачных (ξ_3) земель.

Анализ структуры земель производился в два этапа:

1. Вычисление долей каждого типа земель x_i с помощью преобразования **Векторные слои** \Rightarrow **Атрибуты векторного слоя**: $x_i = \xi_i / (\xi_1 + \xi_2 + \xi_3)$, $i=1, 2, 3$.

2. Выбор эталонных объектов классов структуры земель, оценка функции принадлежности к классам и построение картограммы. Для классификации использован вывод по методу ближайшего соседа. Реализованный в ГИС КОМПАС алгоритм сводится к следующему. Пусть R^l - пространство признаков, нормированных по среднеквадратичным отклонениям, тогда для

точек $\mathbf{x} \in R^l$ среднеквадратичные отклонения $\sigma_i \equiv 1$, пусть $\rho(\mathbf{x}, \mathbf{x}^{(n)})$ - метрика в R^l , C - константа. Классы заданы выборками прецедентов $\{\mathbf{x}^{(q,n)}\}$ и $\{\mathbf{x}^{(\bar{q},m)}\}$, $n = 1, \dots, N$, $m = 1, \dots, M$. Функция принадлежности для двух классов имеет вид:

$$f(\mathbf{x}, \{\mathbf{x}^{(q,n)}\}, \{\mathbf{x}^{(\bar{q},m)}\}) = -\frac{\rho(\mathbf{x}, \mathbf{x}^{(q,u)}) - \rho(\mathbf{x}, \mathbf{x}^{(\bar{q},v)})}{\rho(\mathbf{x}^{(q,u)}, \mathbf{x}^{(\bar{q},v)})} \text{ для } \min_{\langle n \rangle} \rho(\mathbf{x}, \mathbf{x}^{(n)}) \leq C,$$

где $\mathbf{x}^{(q,u)} = \arg \min_{\langle n \rangle} \rho(\mathbf{x}, \mathbf{x}^{(q,n)})$, $\mathbf{x}^{(\bar{q},v)} = \arg \min_{\langle m \rangle} \rho(\mathbf{x}, \mathbf{x}^{(\bar{q},m)})$.

В качестве эталона 1-го класса выбрана Московская область, а в качестве эталонов 2-го класса – Краснодарский край и Оренбургская область. В верхней части рис. 17 показаны диаграммы значений признаков для выбранных эталонов. В нижней части показана картограмма функции принадлежности.

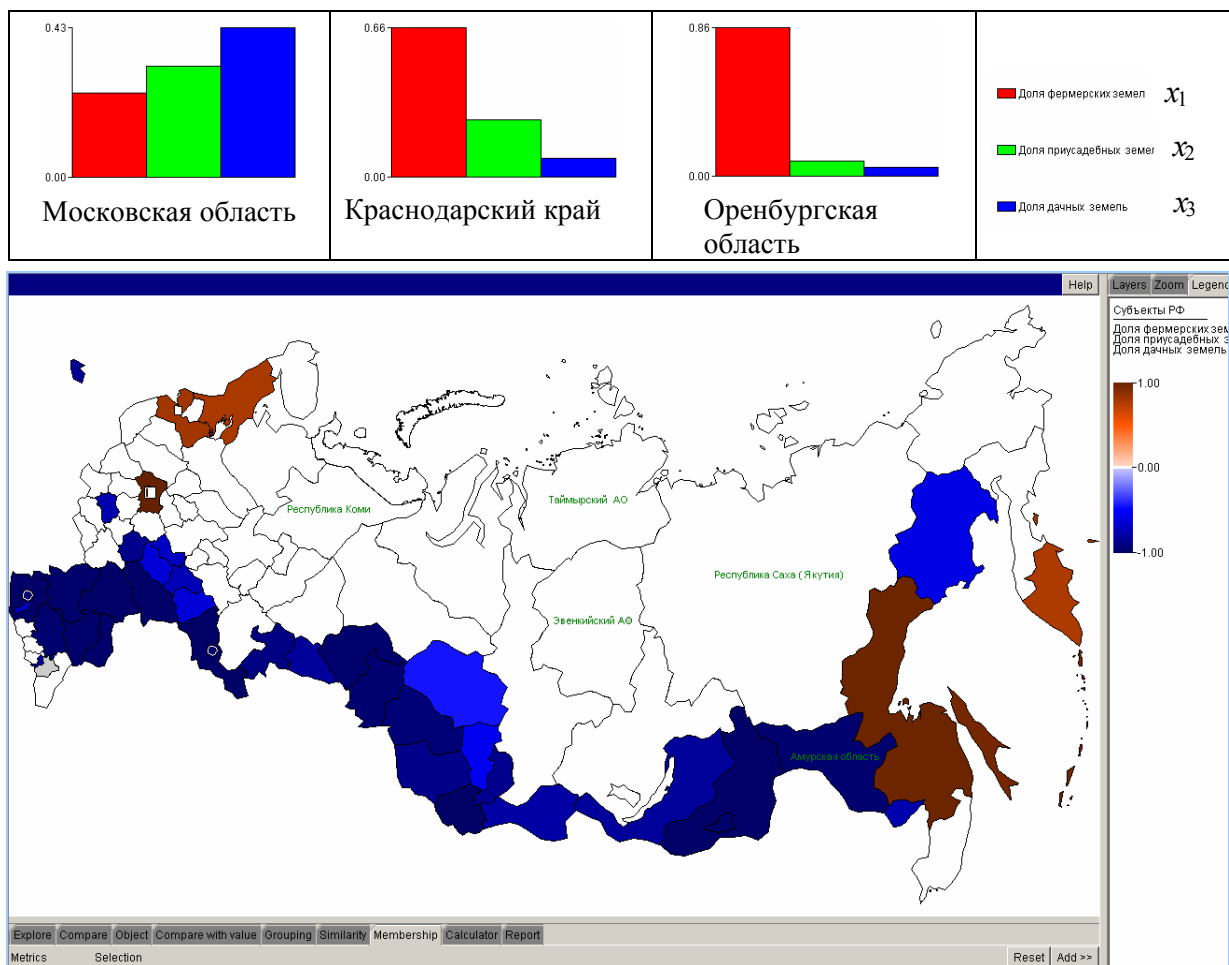


Рис. 17. Диаграммы значений признаков для выбранных эталонов и картограмма функции принадлежности. Объекты, принадлежащие к 1-му и 2-му классам, выделены коричневыми и синими тонами соответственно. Белым цветом показаны несопоставимые объекты, для которых расстояния до эталонов в нормированном пространстве признаков больше 1.

На диаграмме видно, что структуры долей фермерских, приусадебных и дачных земель для эталонов разных классов существенно отличаются. В Московской области преобладают доли дачных и личных земель, а в эталонах 2-го класса доминирует доля фермерских земель. На картограмме видно, что кроме Московской области по структуре использования земель к 1-му

классу относятся Ленинградская область, Республика Карелия, Хабаровский край, Сахалинская и Камчатская области. Структура использования земель 2-го класса характерна в основном для Южных регионов России.

8. ЗАКЛЮЧЕНИЕ

Около 80% информационных потоков составляет географическая информация (ГИ) и ее объем экспоненциально растет. Это требует разработки методов, технологий и систем для автоматического и автоматизированного анализа ГИ с целью извлечения из нее существенной информации и знаний. Сетевые геоинформационные технологии направлены на решение следующих задач комплексного анализа пространственно-временных данных:

1. Оценивание и интерпретация связей между свойствами ГИ.
2. Оценивание и интерпретация отношений между географическими объектами.
3. Прогнозирование, обнаружение и распознавание целевых заранее неизвестных свойств ГИ и заранее неизвестных географических объектов.
4. Прогнозирование развития пространственно-временных процессов.

В ИППИ РАН разработаны сетевые аналитические ГИС ГеоПроцессор и КОМПАС в архитектуре клиент-сервер на языке Java (<http://gis.iitp.ru>). Эти системы могут использоваться для выполнения проектов, в которых требуется поддержка следующих функций:

1. Комплексный анализ пространственно-временных процессов и пространственный прогноз по распределенным данным с удаленным доступом.
2. Обеспечение сетевого пользователя возможностями комплексного анализа пространственно-временной информации, полученной в результате научных исследований (например, клиентская часть ГРИД).
3. Публикация географической информации в электронных библиотеках с возможностью ее интерактивного визуального исследования и анализа.

Пятилетний опыт создания сетевых ГИС и сетевых ГИС-приложений показывает, что системы ГеоПроцессор и КОМПАС являются эффективным инструментом для проведения научных исследований.

ИППИ РАН предлагает Институтам РАН и Университетам научное сотрудничество для создания примеров сетевых ГИС приложений на базе их географических ресурсов. После этого ИППИ РАН согласен безвозмездно передать партнерам ГИС ГеоПроцессор и КОМПАС.

В настоящее время в ИППИ РАН ведется разработка методов и инструментальных средств для сетевого анализа и прогноза пространственно-временных процессов.

СПИСОК ЛИТЕРАТУРЫ

- Гитис В.Г., Вайншток А.П. 2001. Сетевые аналитические ГИС, часть 1// ГИС обозрение № 2, 14 – 16.*
- Гитис В.Г., Вайншток А.П. 2001. Сетевые аналитические ГИС, часть 2. – ГИС обозрение, № 3, 8-12.*
- Гитис В.Г., Андриенко Г.Л., Андриенко Н.В. 2004. Исследование сейсмологической информации в сетевых аналитических ГИС // Физика Земли №3. С. 43-53.*
- Думнов А.Д., Рыбальский Н.Г., Самотесов Е.Д., Максимов Ю.И. и др. 2001. Аналитический доклад "ПРИРОДНЫЕ РЕСУРСЫ И ОКРУЖАЮЩАЯ СРЕДА РОССИИ" // М.: НИИ-Природа, РЭФИА, - 572 с.*
- Кузнецов Н.А., 2001. Информационное взаимодействие в технических и живых системах // Информационные процессы, Том 1, №1. С 1-9.*

- Andrienko, G.; Andrienko, N. 1999a.* Interactive Maps for Visual Data Exploration // International Journal of Geographical Information Science. 13. P. 355-374.
- Andrienko, G., Andrienko, N., Gitis, V. 2003.* Interactive Maps for Visual Exploration of Grid and vector geodata. - ISPRS Journal of Photogrammetry & Remote Sensing 1257 1-10.
- Berg, C., Tuinman, F., Vijbrief, T., Meijer, C., Oisterom, P., Uitermark, H. 1999.* Multiserver Internet GIS: standartization and practical experiments // Goodchild M., Egenhofer M., Feageas R., Kottman C. (eds) Interoperating Geographical Information Systems. Boston: Kluwer Academic Publishers. P. 365-377
- Coburn, A. 1995.* Disaster Prevention and Mitigation in Metropolitan Areas: Reducing Urban Vulnerability in Turkey // Parker, R., Kreimer, A., Munasinghe, M., Editors, Informal Settlements, Environmental Degradation, and Disaster Vulnerability. The Turkey Case Study. World Bank and International Decade for Natural Disaster Reduction, P. 66-94.
- Giardini, D., Grunthal, G., Shedlock, K., and Zhang, P. 1999.* Global Seismic Hazard Map assembled by, GSHAPE.
- Gitis, V., Dovgiallo, A., Osher, B., Gergely, T. 1998.* GeoNet: an information technology for WWW on-line intelligent Geodata analysis // Proceedings of the 4th EC-GIS Workshop. Hungary. Joint Research Centre of European Commission. P. 124-135.
- Gitis, V., Osher, B., Dovgiallo, A., Vainshtok, A. 2000.* COMPASS: Cartography On-line Modeling, Presentation and Analysis System // Proc. of the 5th EC-GIS Workshop. Stresa. Italy, EC JRC. P. 487-497.
- Gitis, V., Jurkov, E., Osher, B., Pirogov, S., Ponomarev, A., Sobolev, G. 1995.* Dynamic fields analysis system. Cahiers du Centre Europeen de Geodynamique et de Seismologie. Vol. 9. P. 129-140.
- Trifunac, M.D. and A.G. Brady. 1975.* On the correlation of seismic intensity with peaks of recorded strong ground motion // Bull. Seismol. Soc. Amer., 65. P. 139-162.