

На правах рукописи

УДК 004.021/658/928/942; 004.652.5;
519.213.1; 519.222/254-256;
550.343.42; 550.34.06.013.3.

МИХЕЕВА
Анна Владленовна

ПРОГРАММНО-АЛГОРИТМИЧЕСКИЙ
ИНСТРУМЕНТАРИЙ ПОДГОТОВКИ И АНАЛИЗА
СЕЙСМОЛОГИЧЕСКИХ ДАННЫХ
В ИНФОРМАЦИОННО-ВЫЧИСЛИТЕЛЬНОМ
КОМПЛЕКСЕ EEDB

05.13.11 – математическое и программное обеспечение
вычислительных машин, комплексов и компьютерных сетей

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата физико – математических наук

Новосибирск 2011

Работа выполнена в Институте нефтегазовой геологии и геофизики
им. А.А. Трофимука СО РАН

Научный руководитель: Дядьков Петр Георгиевич,
кандидат геолого – минералогических наук,
доцент

Официальные оппоненты: Фатьянов Алексей Геннадьевич,
доктор физико – математических наук

Блинов Игорь Николаевич,
кандидат физико – математических наук

Ведущая организация: Учреждение Российской академии наук
Институт физики Земли им. О.Ю. Шмидта

Защита состоится 28 декабря 2011 г. в 15 ч. 00 мин. на
заседании диссертационного совета ДМ 003.032.01 в Институте
систем информатики им. А.П. Ершова Сибирского отделения РАН по
адресу:
630090, г. Новосибирск, пр. ак. Лаврентьева, 6.

С диссертацией можно ознакомиться в читальном зале ИСИ СО РАН
(г. Новосибирск, пр. ак. Лаврентьева, 6).

Автореферат разослан _____ ноября 2011 г.

Ученый секретарь
Диссертационного совета,

к.ф.-м.н.



Мурзин Ф.А.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность проблемы.

Предметная область представляемой информационно-программной разработки довольно широка и охватывает не только прикладные, но и фундаментальные задачи геодинамики - науки о природе глубинных сил и процессов, возникающих в результате планетарной эволюции Земли. Геодинамика лежит в основе тектоники литосферных плит, исследование которой, является одной из фундаментальных задач физики Земли. Нелинейная геодинамика изучает процессы, связанные с нерегулярными, хаотическими импульсами в земных глубинах (например, сейсмического происхождения), а также с воздействиями внеземных факторов (движением комет, падением метеоритов и т.д.). Сейсмогеодинамический режим, исследуемый по данным сейсмичности, является индикатором состояния земной коры, его изучение необходимо и для решения прикладных задач геодинамики, таких как сейсмический прогноз и сейсмозонирование территории. Поиск новых методов прогноза сильных сейсмических активизаций представляет собой весьма актуальную часть данного исследования, т.к. в ряду прочих природных катастроф землетрясения занимают особое место, как представляющие наибольшую угрозу для строительной-хозяйственной деятельности человека.

Учитывая большой объем необходимой информации, значительное число характеристик сейсмического процесса, необходимость опробования на конкретном материале новых теоретических моделей сложных геодинамических систем, важной задачей сейсмогеодинамического исследования стало привлечение новых компьютерных технологий. Несмотря на то что сегодня известно множество программ комплексного исследования сейсмогеодинамического процесса, информационное обеспечение этих программ неполно. Отсутствуют представительные каталоги землетрясений и достаточно детальные географические данные, которые при исследованиях в планетарном масштабе составляют наиболее весомую часть информационного обеспечения и требуют решения проблемы их хранения, представления и анализа. Кроме узкой территориальности и других недостатков информационного и программного обеспечения, можно отметить недостаточное удобство многомодульных интерфейсов. Все это делает необходимым развитие информационных технологий: использование новых ГИС-технологий, новых средств для разработки интерфейсов.

Кроме того, известные вычислительные системы исследования сейсмичности являются узко-ориентированными на специфические задачи (например, прогноз времени и места крупных событий или выявление тектонических структур), т.е. недостаточно универсальны для сейсмогеодинамических исследований.

Таким образом, существующие программные разработки теоретически и практически не устраивает исследователей сейсмогеодинамического процесса по следующим причинам:

- 1) недостаточной эффективности алгоритмов обработки данных,
- 2) отсутствия единой среды для исследования,
- 3) специфической направленности решаемых задач,
- 4) узкой территориальности,
- 5) недостаточной гибкости и наглядности инструментов анализа.

Цель настоящей работы – на основе развития программных и информационных технологий повысить качество подготовки, анализа и интерпретации сейсмологических данных, т.е.:

- повысить скорость обработки данных;
- унифицировать форматы в рамках единого цикла обработки и анализа сейсмологических данных.
- исключить человеческий фактор при выполнении рутинных трудоемких операций в работе с данными;
- повысить качество предоставляемой информации о сейсмическом процессе;
- создать эффективный многооконный человеко-машинный интерфейс для сейсмо-геодинамических исследований.

Научная задача – разработка интегрированного компьютерного информационно-аналитического комплекса изучения сейсмичности в трех частях: информационного и программно-алгоритмического обеспечения сейсмологической базы данных, развития информационного обеспечения географической оболочки, математического и программного обеспечения подсистемы пространственно-временного анализа сейсмологических данных.

Теоретической и методологической основой диссертационного исследования явились основные положения теории систем, теории вероятности и математической статистики, достижения отечественных и зарубежных авторов в области разработки методов исследования сейсмичности, методов картографии, развитию геоинформационных систем и баз данных, геоинформационному моделированию.

Методы исследования, используемые в диссертации: прикладные методы теории систем – планирование, проектирование и

конструирование системы «человек-машина»; системно-структурный и статистический анализ; графический и картографический методы визуализации результатов анализа; различные методы графического и пространственно-временного анализа характеристик сейсмического процесса: угла наклона графика повторяемости ($b(t)$, $b(s)$), изменения суммарной сейсмической энергии ($E(t)$, $lg E_s/E_n(s)$), сейсмической активности ($A10$, $A15$), критерия плотности сейсмогенных разрывов ($K_{cp}(t)$, $K_{cp}(s)$), стабильности выделения сейсмической энергии ($1/\sigma_-(t)$, $\sigma(s)$), методы выявления группирования землетрясений; методы теории вероятности; стандартные компьютерные методы: формирования реляционных баз данных в среде стандартных СУБД, создания программных приложений и баз данных в объектно-ориентированной среде на базе библиотеки MFC; общенаучные методы: тестирования, верификации, сравнительный анализ.

В работе использовались **фактические материалы** по сейсмичности отдельных регионов – оперативные и детальные каталоги землетрясений, предоставленные Алтае-Саянским и Байкальским филиалами Геофизической службы СО РАН, мировые и региональные каталоги ведущих геофизических служб (USGS, Harvard Seismology Group, Мирового Центра Данных в Москве и т.д.), а также цифровые модели рельефа (ЦМР) различной детальности (USGS).

Научная новизна.

Впервые создан интегрированный информационно-аналитический комплекс EEDB (Expert Earthquakes Data Base - «Экспертная база данных землетрясений»), собравший в единой среде алгоритмы предобработки сейсмологических данных (выборки событий, выбора области произвольной формы, проверки качества каталогов, очистки от афтершоков и роев) с функциями анализа данных и их визуализации. В частности:

- 1) для сейсмологической базы данных созданы: новые структуры хранения и комплекс программ обработки и конвертирования каталогов землетрясений различных исходных форматов; структура размещения полей каталога механизмов; наиболее полный сводный каталог землетрясений сейсмоактивных районов юга Сибири;
- 2) для системы управления сейсмологической базой данных предложены новые алгоритмы предобработки данных, в частности, идентификации афтершоков в последовательности событий на основе алгоритма А.Г. Прозорова, модифицированного с целью увеличения быстродействия и эффективности;

- 3) для географической системы разработаны: новые варианты визуализации очагов на географической карте и вертикальном разрезе; алгоритмы расчета и построения зональных карт; программы регионализации сейсмичности для проведения одновременного анализа нескольких тектонических областей;
- 4) внедрены географические данные с высокой разрешающей способностью для оперативного проведения исследований на разных масштабных уровнях;
- 5) реализованы варианты расчета параметра плотности сейсмогенных разрывов K_{cp} для теоретических случаев стабильного развития сейсмического процесса;
- 6) программно реализован уникальный комплекс исследования суммарной нормированной сейсмической энергии $lgE_s/E_n(t,s)$ (и её градиента) с оценкой площади выделяемых областей сейсмических затиший и возможностью выбора конкретной аномалии для временного анализа её значений;
- 7) разработан наиболее актуальный в своей предметной области программный блок гибкого анализа сейсмологических данных, включающий в себя необходимый набор характеристик (перечисленных в *методах анализа характеристик сейсмического процесса*) для поэтапного их рассмотрения.

Практическая ценность.

Создана единая программная среда, позволившая организовать проведение коллективного многопланового исследования сейсмогеодинамических процессов в литосфере: пространственно-временных аномалий сейсмического режима; процессов подготовки крупных сейсмических событий; межрегиональной корреляции сейсмичности; взаимного влияния удаленных землетрясений. В статьях ведущих журналов подтверждено успешное использование программного комплекса EEDB для выдвижения и обоснования геомеханических моделей литосферы и моделей подготовки сильных землетрясений, а также для выявления среднесрочных сейсмопрогностических критериев. В частности: создана геодинамическая модель подготовки Чуйского землетрясения (27.9.2003, $M_s=7.5$); проанализирован сейсмический режим различных регионов юга Сибири; исследованы процессы подготовки Суматра-Адаманского мега-землетрясения (26.12.2004, $M_w=9.1$) и их влияние на землетрясения в Тибете (14.11.2001, $M=8.1$), Токачи-оки (25.9.2003, $M_s=8.3$) и Чуйское землетрясение; выявлены факты «переклички» весьма удалённых друг от друга землетрясений БРЗ и их корреляция с событиями Дальневосточной зоны субдукции.

Особую значимость имеют технические, математические и информационные характеристики данной программной разработки. Сейсмологическая база данных и ГИС-система EEDB содержат информационное и программное обеспечение соответствующей детальности и содержания для различного масштабного уровня и большого разнообразия требований научных исследований. Эффективно организованное взаимодействие обеспечивающих и функциональных подсистем позволяет проводить полный цикл исследований, включающий все этапы предварительной обработки, анализа и визуализации данных сейсмологических каталогов в одной среде. Данный опыт можно использовать для проектирования различного рода новых геоинформационно-аналитических систем регионального и субрегионального масштаба.

Апробация работы и публикации.

Работа выполнена в Лаборатории естественных геофизических полей (ИНГГ СО РАН). Программный комплекс EEDB является основой системы GIS-EEDB, успешно внедренной и применяемой для исследовательской работы на нескольких компьютерах лаборатории.

Первые разработки по базам данных землетрясений и информационно-экспертным программам, имеющие частичное отношение к данной работе, представлялись на Совещании по природным и антропогенным катастрофам (Томск, октябрь 1991) – с демонстрацией первой версии программы по проблеме землетрясений и цунами, а также на Международных ассамблеях (Австрия, 1991; Япония, 1993; США, 1995, Австралия, 1997). Результаты настоящей работы хорошо известны научной общественности, они получили одобрение специалистов: на региональных (Армения, 2004; Тайланд, 2006), всероссийских (Иркутск, 2005; Новосибирск, ИНГГ, 2009; Москва, ИПУ, 2010; Петрозаводск, 2011) и международных (Переславль – Залесский, ИПС, 2009; Новосибирск, ИСИ, 2009 – с демонстрацией EEDB; Париж, 2011) конференциях.

По теме диссертации соискатель имеет 17 публикаций и 2 регистрационных свидетельства (базы данных EEDB и программно-алгоритмического комплекса EEDB), участвовала от ИНГГ СО РАН в 5 проектах РФФИ (с 2000 по 2012 гг.), 5 интеграционных проектах ИП и МИП СО РАН, программе Президиума РАН и программе ОНЗ РАН.

Структура и объем работы

Диссертационная работа состоит из введения, четырех глав, заключения, списка литературы и трех приложений. Объем

диссертации – 176 страниц. Список литературы содержит 217 наименований. Работа включает 66 рисунков, 7 таблиц и 2 приложения.

СОДЕРЖАНИЕ РАБОТЫ

Во **введении** обосновывается актуальность темы исследований и приводится краткая характеристика работы.

В первой главе приведён исторический обзор использования специализированных программ и автоматизированных комплексов при решении геолого-геофизических задач в области сейсмогеодинамики и особенности вновь созданного программного комплекса. Общая оценка возможностей современных информационных комплексов применительно к решению актуальных задач хранения, обработки и интерпретации результатов сейсмологических исследований подтверждает общепризнанный факт отсутствия на данный момент достаточно универсального программного обеспечения для изучения процессов подготовки землетрясений, а также невозможности использования имеющегося программного обеспечения для решения специализированных задач по изучению сейсмогеодинамического режима южных районов Сибири.

Большинство систем изучения сейсмического режима из-за своей узкой территориальности (MERP, PGA, RDBMS NGDC,...) не позволяют учитывать взаимного влияния удаленных, но геодинамически связанных регионов. Кроме того, известные вычислительные системы исследования сейсмичности являются узко-ориентированными на специфические задачи:

- «детерминистский» прогноз (prediction) сильных событий с жестко определенными способами интегрирования параметров сейсмического процесса, фиксированными требованиями к их средним значениям и порогам (известные алгоритмы КН, М8, Мsc диагностики ППВ, ПВОПВ, расчет интегрального критерия RTL), обладающие излишней автоматизацией на этапе принятия решений,

- прогноз в универсальном смысле (проверка статистических гипотез), общие подходы для задач прогнозирования (ГЕО В.Г.Гитиса),

- «вероятностный» прогноз (forecasting) поведения сейсмичности по анализу комплекса предвестников имеет недостатки, связанные с используемыми в алгоритмах математическими моделями (КОЗ А.Д.Завьялова - Г.А.Соболева; ETAS Д.Джексона -Я. Кагана)

Интегрированный анализ, предлагаемый в этих и других программах как наиболее эффективный, в ряде случаев затрудняет исследование. Например, для универсальных задач исследования сейсмогеодинамического режима на меняющемся масштабном уровне

от узкой области подготовки землетрясения до уровня межрегиональных взаимодействий неприменимы узкоориентированные и четко формализованные подходы «детерминистского» прогноза. Более гибкими являются вероятностные подходы анализа комплекса предвестников, но эти алгоритмы имеют, по словам П.Н. Шебалина, недостатки, связанные с используемыми в них математическими моделями: Байесовский подход к распознаванию (в алгоритме КОЗ) не эффективен для анализа зависимых признаков, а известная модель Огаты «эпидемического самовоспроизводства землетрясений» (в ETAS) не учитывает наличия их общих причин.

Кроме недостатков информационного (Imp.Itrис и др.) и программного обеспечения (ISC, SeismoGIS, OpenSHA и др., не содержащие собственной ГИС-системы), в существующих разработках можно отметить недостаточное удобство многомодульных интерфейсов (например, «EQ» и «Prediction» В.В.Ружича).

Всё это побуждает начать работу по созданию собственного, более актуального программного комплекса для исследования сейсмогеодинамического процесса, позволяющего использовать развитые информационные технологии и необходимый математический аппарат для выполнения следующих требований: наукоёмкости; возможности экспериментирования с использованием новых методик; отказа от излишней автоматизации (т.к. нужна система экспертных оценок, информационной помощи исследователю); возможности оперативного проведения исследований на разных масштабных уровнях; адаптированности к конкретным сейсмоактивным районам юга Сибири.

Предложена функциональная архитектура и иерархия классов (~240 классов) нового комплекса, на уровне логического обобщения представляющие собой совокупность трёх взаимодействующих между собой программных блоков: сейсмологической базы данных, географической подсистемы и подсистемы анализа данных (рис. 1). Подсистема анализа является наиболее наукоемким функциональным элементом EEDB. Её первый слой составляют процедуры проверки полноты и качества каталогов землетрясений. Следующий слой связан с визуальным анализом комплекса характеристик и состоит из двух подуровней: графического и картографического.

Далее описываются информационные характеристики программной системы GIS-EEDB, интегрированной автором на базе комплекса программ EEDB и ГИС-технологий. Кроме элементов ГИС-системы (представляемой также в других, более ранних разработках, например, WinHTDB, WinITDB и т.д. В.К.Гусякова - Ан.Г.Марчука),

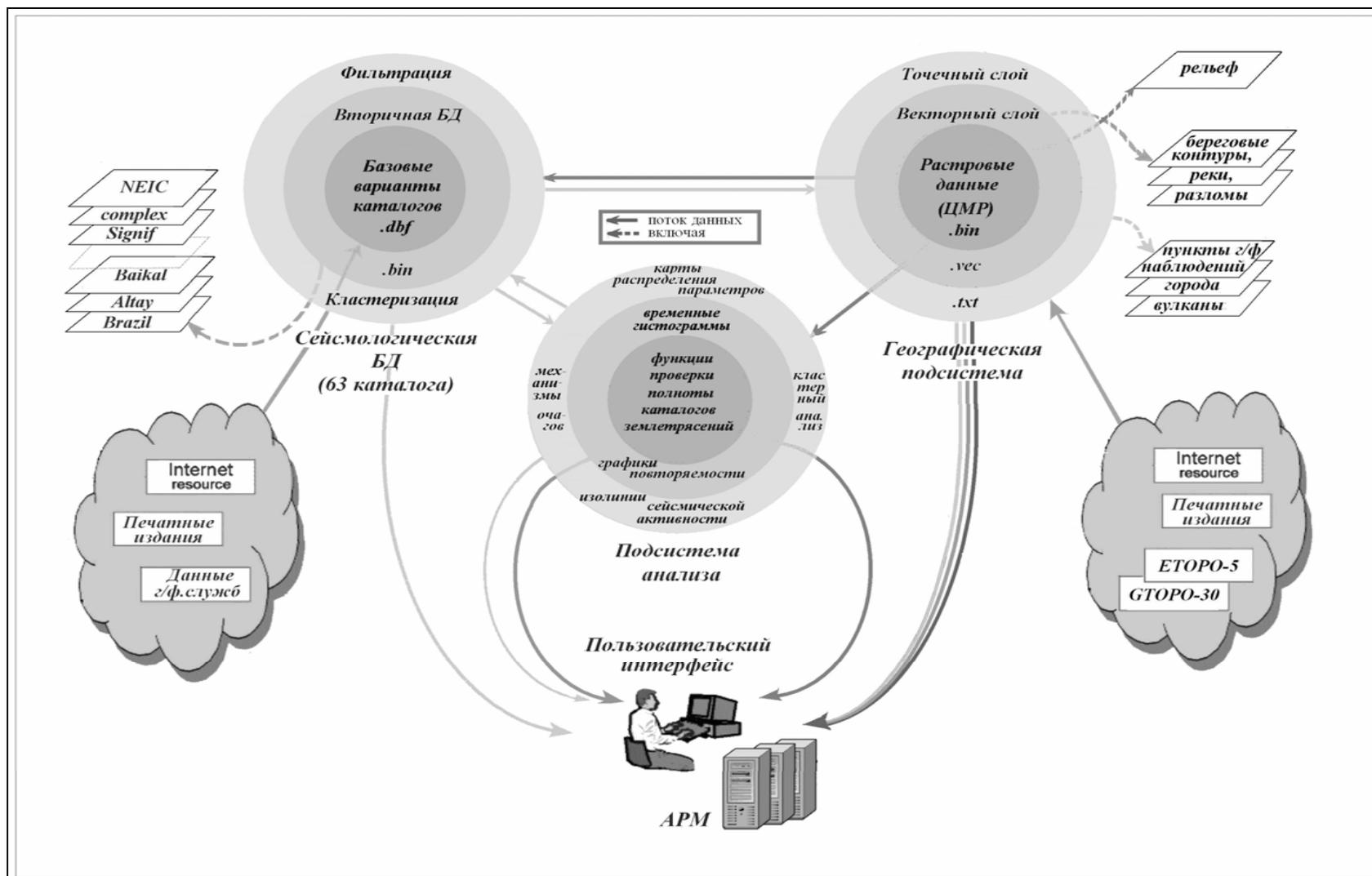


Рис. 1. Архитектура систем, входящих в комплекс EEDB.

система GIS-EEDB содержит также элементы экспертных (модельных) систем и специализированных систем применительно к конкретному кругу задач. Вследствие этого GIS-EEDB можно отнести скорее к классу систем обеспечения принятия решений (СОПР) 1-го типа, т.е. автоматизированной информационно-аналитической системе (ИАС), включающей в себя совокупность математических алгоритмов обработки и анализа массовых геоданных на основе ГИС-технологий.

В отличие от информационной модели, являющейся неформальной частью и средством представления знания об изучаемом объекте, физико-математическая модель составляет формальную обобщенно-теоретическую часть описания изучаемого объекта.

Математическую модель исследуемой проблемы (сейсмического процесса) можно сформулировать следующим образом:

1. эмпирические энергетические и геометрические закономерности сейсмической регионализации: закон повторяемости землетрясений [В.Gutenberg - Ch.Richter], закон затухания интенсивности афтершокового процесса [закон Омори] (степенные законы), фрактальная размерность множества сосредоточения эпицентров землетрясений [М.А.Садовский; В.Г.Кособоков] (самоподобие – в глобальном масштабе);

2. «канонические» физические законы распределения временных рядов данных: Гауссовского (нормального), Пуассоновского (экспоненциального) ввиду недостаточной статистической обеспеченности эмпирических функций; а также пространственного распределения афтершоков: двумерного нормального распределения.

Таким образом, реализованные в программно-алгоритмическом инструментарии EEDB математические модели имеют логико-вероятностный характер. Эмпирические уравнения являются качественными, выявляющими поведение рассматриваемой динамической системы в предельных случаях. Задача эксперта, принимающего решение на основе данных EEDB, состоит в квалифицированной интерпретации предоставляемой программным комплексом информации. Обобщенная пользовательская схема анализа сейсмологических каталогов в среде EEDB приведена на рис. 2.

Физическая структура системы, созданной на основе программного комплекса EEDB, представляет собой совокупность трех базовых файлов: исполняемого модуля, файла путей размещения наборов данных, файла с описанием меню комплекса; и пяти папок, содержащих файлы географической и сейсмологической баз данных. В целом прикладной пакет GIS-EEDB содержит более 150 файлов. Объем сгенерированного и авторского программного кода ~23 Мбт.

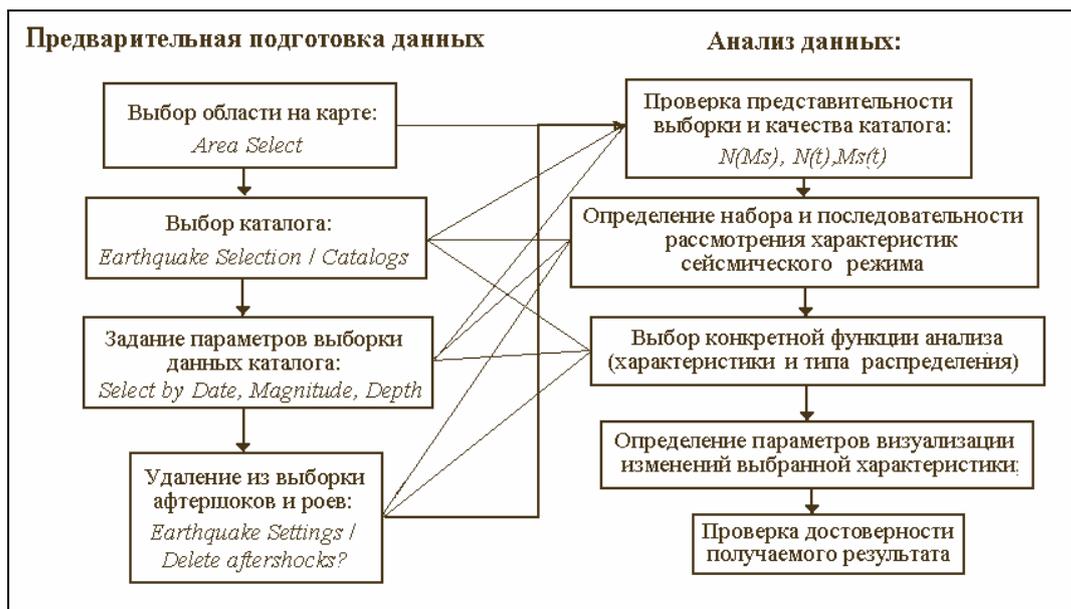


Рис. 2. Обобщенная схема анализа данных в среде EEDB.

Вторая глава посвящена описанию системы управления базами сейсмологических данных. Создана реляционная структура хранения и комплекс программ обработки и конвертирования каталогов землетрясений различных исходных форматов, программы автоматизированного пополнения каталогов и их последующей проверки с целью исключения дублирующей и неиспользуемой информации, а также программы классификации данных для их рациональной обработки и представления. Логическая структура сейсмологической базы данных (СБД) может быть представлена в виде двукорневой схемы со стандартной СУБД (использующей дисковую память) на *базовом* уровне [Гусяков, **Осипова (Михеева)**, 1992] и структурой размещения данных (табл. 1) в программной среде EEDB (использующей оперативную память) на уровне *прикладном* (для загрузки данных из исходного бинарного файла в ГИС-среду). Это разделение сделано и из соображений эффективности: постоянное обращение к базе данных на диске замедлило бы работу программы.

В рамках СБД созданы *комплексные сводные каталоги* (в частности, наиболее полный каталог землетрясений сейсмоактивных районов юга Сибири), разработана соответствующая программная поддержка их создания. В результате составлен расширенный список каталогов в среде EEDB (63 каталога), включающий в себя как опубликованные в литературе каталоги, преобразованные в единый формат (каталоги службы USGS: NEIC – с 1973 г. по сей день, $\sim 655 \cdot 10^3$ записи, SIGN – с -2000 по 1993, $\sim 6.7 \cdot 10^3$ зап.; каталоги ГС СО РАН: Байкальский $\sim 150 \cdot 10^3$ зап., Алтайский $\sim 32 \cdot 10^3$ зап.), так и

комплексные каталоги, собранные из множества публикаций (ВАИК ~167*10³ зап. и COMPLEX >1*10⁶ зап.). Проведено исследование полученных каталогов на полноту и представительность.

В программный блок системы управления СБД включена *первичная подготовка исходных каталогов*, заключающаяся в выборе текущего каталога и последующей выборке подмножества землетрясений по параметрам запроса: диапазону времени, магнитуд, пространственному диапазону и т.д.

Дальнейшая обработка данных к исследованию предполагает очищение выбранной части каталога от афтершоков и роев, чему служит отдельная функция, реализующая три метода выделения афтершоков по выбору пользователя (рис. 3).

Первый метод, условно названный *эмпирическим*, основан на эмпирически установленной зависимости пространственно-временных параметров поискового окна (*dT* и *dS*) от величины главного толчка.

Второй (*эллиптический*) метод опирается на основные принципы алгоритма А.Г.Прозорова, модифицированного следующим образом:

Таблица 1.

Структуры размещения сейсмологических данных в ПС EEDB.

Структурное описание гипоцентра толчка	Атрибут	Структурное описание механизма	Атрибут
<pre>typedef struct _EVENTSTRUCT { SHORT nYear; SHORT nMonth; SHORT nDay; SHORT nHour; SHORT nMinute; SHORT nSecond; CDate nTime; int nLat; int nLon; int nDepth; int ndDepth; int nBook; LONG nMs; LONG nKl; LONG nL; SHORT radius; CPen *pen; CBrush *brush; COLORREF color; }EVENTSTRUCT;</pre>	<p>Год Месяц День Час Минута Секунда Дата Широта Долгота Глубина Погрешн.глуб. Библ.ссылка Магнитуда Класс Длина очага Радиус Перо Заливка Цвет</p>	<pre>typedef struct _MEXSTRUCT { INT nYear; SHORT nMonth; SHORT nDay; SHORT nHr; SHORT nMn; SHORT nSc; CString nT; CString nN; CString nP; CString nNP1; CString nNP2; SHORT nRange; }MEXSTRUCT;</pre>	<p>Год Мес. День Час Мин. Сек. Ось растяж. Промеж.сеть Ось сжатия Нодал.пл.1 Нодал.пл.2 Ранг</p>

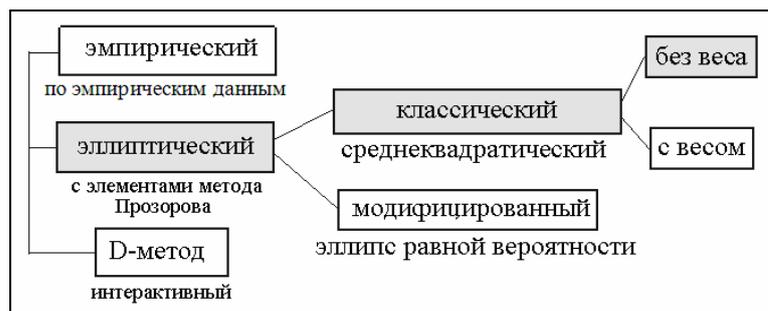


Рис. 3. Схема алгоритмов выделения афтершоков. Фоном выделены методы, работающие по умолчанию.

все выделяемые афтершоковые последовательности ведутся параллельно за один проход каталога (что значительно улучшает временные характеристики алгоритма), минимальный размер прямоугольной метрики задаётся интерактивно (повышает гибкость алгоритма), кроме того, в отличие от *классического* варианта расчета эллиптических метрик, предложено построение эллипса равной вероятности (рис.4). Третий метод назван *интерактивным* - значения dT и dS задаются пользователем.

Эллиптический метод предполагает задание следующих параметров: порог сигнал/шум $R_{s/n}$, минимальная магнитуда главного толчка, минимальная магнитуда афтершоков, минимальный размер прямоугольной метрики и т.д.

На примерах показана эффективность *модифицированного* метода, фиксирующего пространственную структуру распределения афтершоков с помощью эллипса равной вероятности:

$$\varphi(x,y) = \frac{1}{1-\rho_{12}^2} \left(\frac{x^2}{\sigma_1^2} - 2\rho_{12} \frac{xy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2} \right) = \lambda^2, \quad \text{где } \lambda^2 \approx 2 \cdot \left(1 - \frac{1}{9} + 3.29 \cdot \frac{1}{3} \right)^3 -$$

аппроксимация квантилей для распределения с двумя степенями свободы при $P = 0.9995$; $\sigma_1^2 = DX$, $\sigma_2^2 = DY$ – дисперсии x и y , а ρ_{12} – коэффициент корреляции между x и y . Результат применения данной оценки параметров эллипса при выявлении афтершоков землетрясения 16.9.2003 в северной части Байкальской рифтовой зоны (БРЗ) превзошел *классический* вариант как по числу выделенных событий (263 и 246 соответственно), так и по длительности афтершоковой последовательности (3,9 и 1,6 года, соответственно).

Преимуществом *модифицированного* метода является также его независимость от порогового соотношения сигнал/шум $R_{s/n}$.

Сравнение эффективности описываемых алгоритмов проведено с помощью оценки статистических свойств получаемых множеств.

Результаты сравнения работы *классического* и *модифицированного* алгоритмов (рис. 5) показывают, что наблюдаемые значения числа событий превышают теоретическую кривую

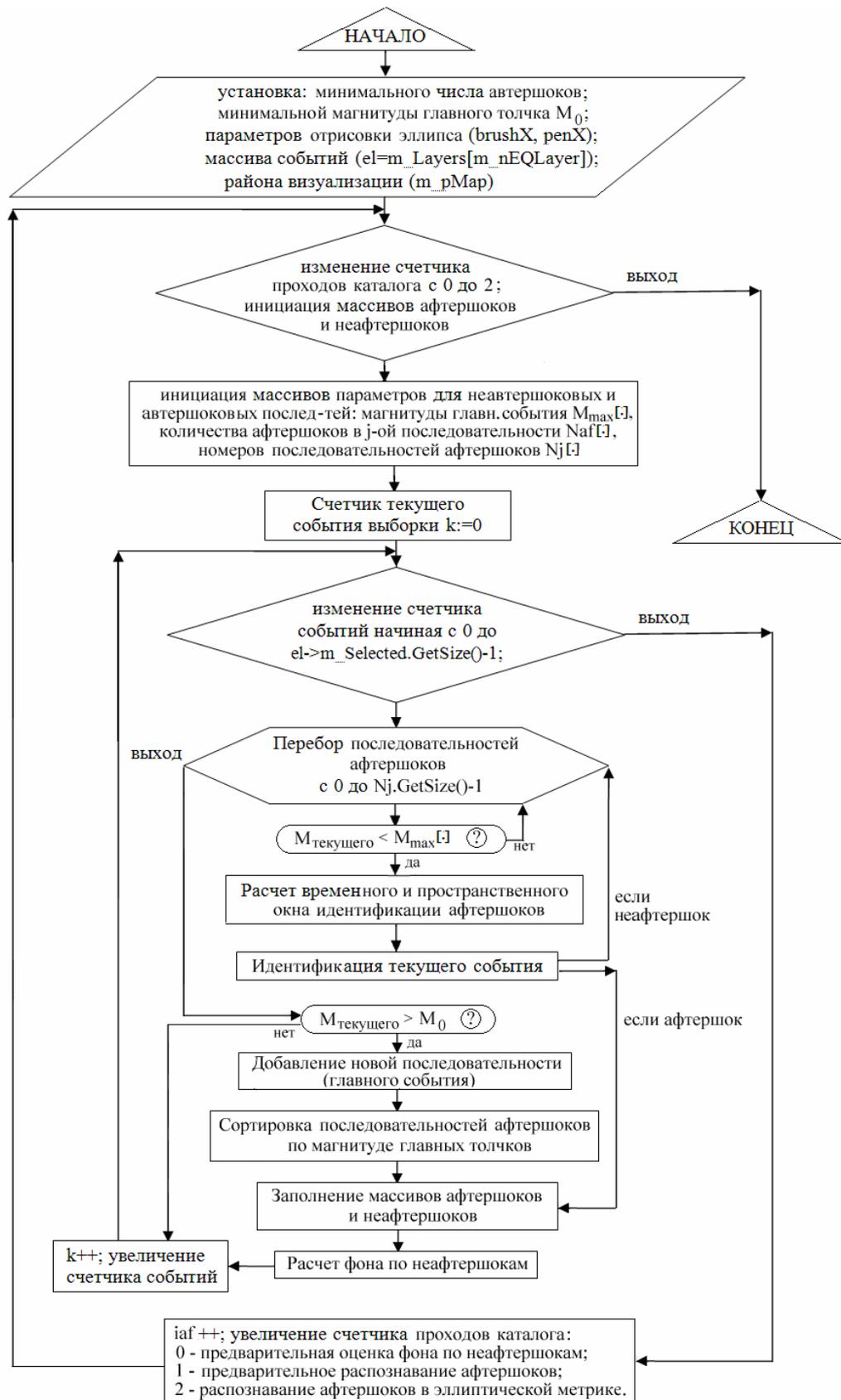


Рис. 4. Блок-схема модифицированного алгоритма выделения автершоков.

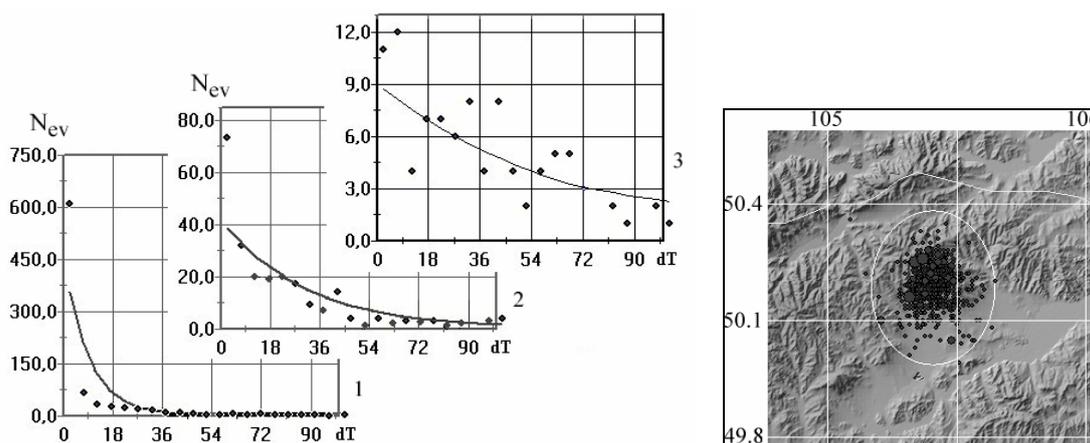


Рис. 5. Отклонение наблюдаемого числа землетрясений каталога ВАИК ($M \geq 1.5$ с 1987г. в области $49,7-51,1^\circ\text{с.ш.}; 104-106,9^\circ\text{в.д.}$) от теоретической кривой распределения Пуассона: 1 - до удаления афтершоков Северо-Монгольского землетрясения 13.05.1989, $M=5.8$ (852 событий); 2 - после удаления афтершоков классическим алгоритмом (257 событий, $R_{s/n}=20$); 3 - после удаления афтершоков модифицированным алгоритмом (120 событий). Приведена карта 733 афтершоков, выявленных модифицированным методом.

экспоненциального распределения для случайных величин в первый 5-дневный интервал как до удаления афтершоков (рис. 5.1), так и после их удаления классическим методом (рис. 5.2). Это говорит о недостаточной эффективности классического метода. Экспоненциальное распределение потока событий (рис. 5) описывается формулой Пуассона: $f_T(\tau) = \lambda \cdot \exp(-\lambda\tau)$, где $\tau = dT$, λ – скорость событий: $\lambda = N/\tau_{\text{ср}}$, где $\tau_{\text{ср}}$ – среднее время между событиями изучаемого процесса.

Третья глава описывает географическую систему, создающую картографический планшет для визуализации сейсмологических данных и использующую различные форматы картографических данных: растровые, векторные и точечные. Для визуализации растровой информации используется метод цифровой картографии, заключающийся в тонировании (закрашивании оттенком оговоренного цвета) точек поверхности в зависимости от их яркости при заданном векторе бокового освещения (разновидность Гуро-тонирования). Для обеспечения возможности построения карт различных масштабов – от обзорной карты всего мира до карт отдельных регионов автоматически выбирается цифровой массив рельефа с оптимальным разрешением.

В качестве ЦМР в графический пакет GIS-EEDB внедрены глобальные банки данных рельефа земной поверхности GTOPO-30 и SRTM-90 с, соответственно, 30-секундной и 3-секундной (для юга Сибири) сетками данных. Затем на растровое изображение

накладываются векторные (разрешающая способность векторных данных речной сети юга Сибири достигает 500 м) и точечные слои.

В процедурах географической системы реализованы различные варианты отображения очагов на географической карте и вертикальном разрезе, в том числе с использованием анимации. Анимация дает исследователю дополнительное знание о характере процесса, используя два вида компьютерной информации: видео и звук.

Организована процедура регионализации сейсмичности с помощью произвольным образом ориентированных рамок на карте для проведения сравнительного анализа нескольких тектонических областей, реализована графическая визуализация последовательности зональных карт по равномерным интервалам времени (рис. 7).

В алгоритме тонирования зональных карт применена двумерная формула Бесселя линейной интерполяции по задаваемой сетке для z_{ij} :

$$Z(x, y) = \frac{1}{4} (z_{00} + z_{10} + z_{01} + z_{11}) + \frac{1}{2} (u - \frac{1}{2}) (z_{10} - z_{00} + z_{11} - z_{01}) + \\ + \frac{1}{2} (v - \frac{1}{2}) (z_{01} - z_{00} + z_{11} - z_{10}) + (u - \frac{1}{2}) (v - \frac{1}{2}) (z_{11} - z_{10} - z_{01} + z_{00}) + \dots,$$

где $u = (x - x_0) / \Delta x$; $v = (y - y_0) / \Delta y$; $\Delta x = \Delta y = h$ – фиксированное приращение.

В четвертой главе описана система анализа, условно названная экспертной системой, первый слой которой составляют процедуры анализа данных на полноту и качество $N(t)$ и $M_s(t)$. Приводятся примеры их использования для сравнительной оценки полноты и представительности каталогов СБД. По результатам исследования этого уровня выявляются временные области, различающиеся по качеству регистрации и диапазонам регистрируемых магнитуд. Делается вывод о том, что анализ графиков $N(t)$ и $M_s(t)$ не даёт глубокого знания о полноте каталогов и может отражать лишь качество данных отдельных участков охватываемой каталогами территории. Только в совокупности с картографическим анализом пространственного распределения параметров N и M_s , анализ графиков создаёт полное представление о качестве регистрации землетрясений.

Методы графического и картографического анализа данных, составляющие следующие два слоя системы позволяют исследовать ряд характеристик сейсмического процесса:

- наклон графика повторяемости $b(t)$, $b(t,s)$ на основе метода наименьших квадратов МНК (рис. 6) с учетом допустимого среднеквадратичного отклонения и минимального объема выборки (показано его преимущество по сравнению с ММП);
- параметр плотности сейсмогенных разрывов $K_{cp}(t,s)$, $K_{cp}(t)$ с широким набором вариантов: по экспериментальным данным, для теоретического случая равномерного приращения длин трещин;

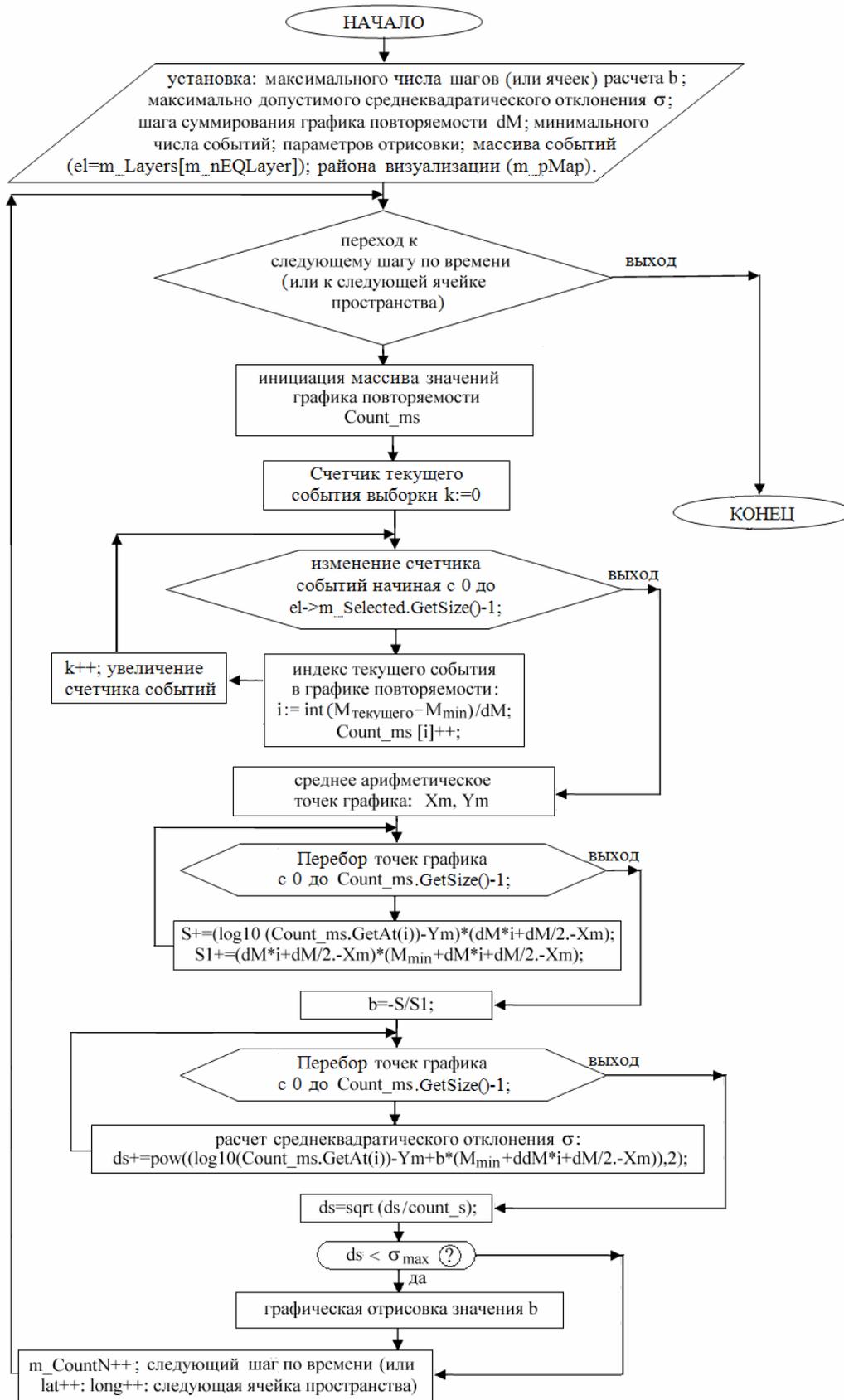


Рис. 6. Блок-схема алгоритма расчета угла наклона графика повторяемости $b(t)$ (или $b(s)$) на основе МНК с учетом дисперсии σ .

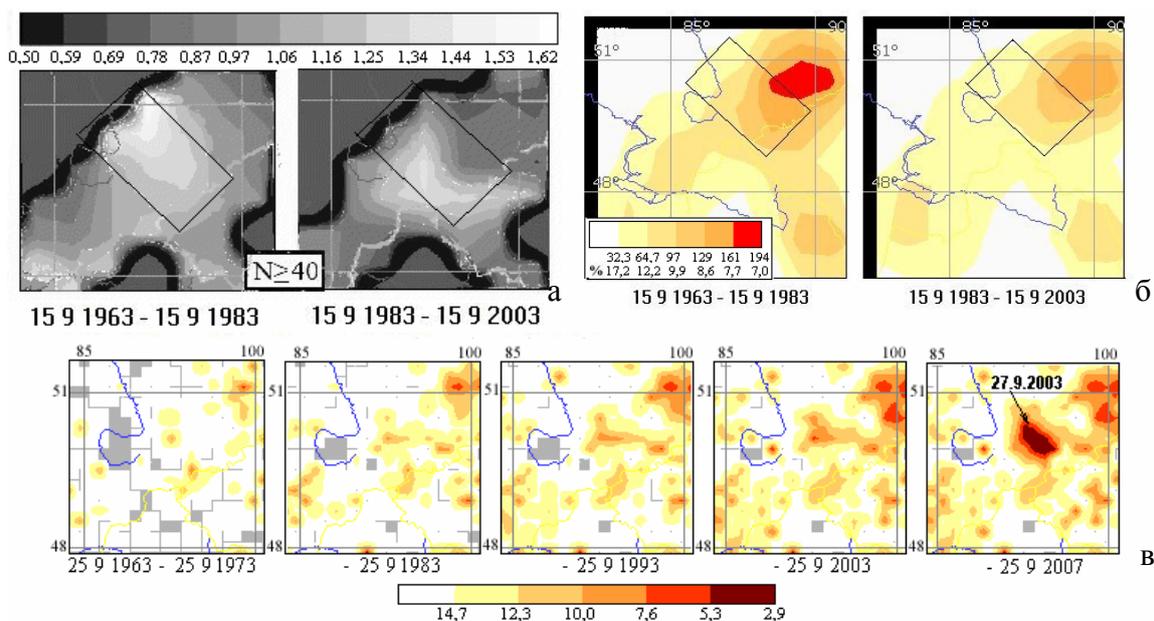


Рис. 7. Пространственное распределение характеристик сейсмического режима в области Чуйского землетрясения 27.09.2003 г. за 40-летний период перед событием: а – угла наклона графика повторяемости b по выборке с $M=2.6-4.6$ с двойным перекрытием по пространству; б – соответствующая карта количества землетрясений (ϵ указывается в %); в – кумулятивной плотности сейсмогенных разрывов K_{cp} землетрясений с $M_s=2.0-4.5$ через каждые 10 лет, размер сетки $0.2 \times 0.3^\circ$. БЗП отмечена рамкой.

- суммарная нормированная сейсмическая энергия $lgE_s/E_n(t,s)$ (и её градиент) с оценкой площади выделяемых областей сейсмических затиший и возможностью выбора конкретной аномалии для временного анализа её значений;
- три варианта расчета сейсмической активности: как плотности распределения землетрясений конкретной энергетической величины, распределения числа всех событий и традиционным методом A_{10} (A_{15}).

Для основных характеристик организовано окно одновременного вывода ряда гистограмм: по последовательным интервалам времени, по тектоническим областям, по выбранным статистическим функциям.

Приводятся результаты использования отдельных методов в исследовании сейсмического режима ближней зоны подготовки (БЗП) Чуйского землетрясения (27.09.2003, $M=7.5$), выявлены аномалии распределения концентрационного критерия K_{cp} и угла наклона графика повторяемости землетрясений b на зональных картах (рис. 7).

В частности, на графиках и картах b -value выявлено аномальное понижение значений параметра на протяжении 20-лет перед Чуйским событием, что свидетельствует о самоорганизации среды в жёсткие структуры, т.е. является предвестником сильного землетрясения.

Для статистически точного определения необходимого объёма выборки n и учёта её влияния на погрешность ε получаемых значений характеристик использовался вывод из теоремы Муавра-Лапласа о доверительном интервале для выборочной оценки: $\varepsilon = |u| / (2\sqrt{n})$, где $|u|_{0.95} = u_{0.975} \cong 1.96$. Кроме требования к объёму выборки программой осуществляется контроль качества результата по линейности графика повторяемости во всех ячейках путем задания максимально допустимого среднеквадратичного отклонения σ . На основании этих характеристик предполагается подбор пользователем параметров расчета b -value (размера ячеек, интервала времен и т.д.) с целью установления оптимального распределения сейсмичности в ячейках и выявления устойчивых к этим параметрам аномалий. Оптимальность распределения землетрясений определяется не только количественной достаточностью событий, попадающих в скользящее окно, но и статистической однородностью выборки.

Для дополнительного контроля достоверности получаемых аномалий в комплексе EEDB создана процедура формирования *синтетического* каталога по пространственным точкам реальных данных с заменой их dT на случайные значения (рандомизацией реальных данных) в заданных интервалах времени.

Относительно метода расчета b делается вывод, что алгоритмы на основе метода максимального правдоподобия ММП, традиционно используемые при расчете b , являются менее надежными, т.к. не позволяют осуществлять контроль качества исходных данных по дисперсии точек графика повторяемости и, кроме того, являются слишком чувствительными к выбору начального класса K_{min} .

При изучении приуроченности сейсмического процесса к основным тектоническим структурам и при учете миграции событий во времени требуются *кластерные методы* выявления *цепочек* (рис. 8) связанных событий, отвечающих критериям П.Н. Шебалина. В комплексе EEDB реализовано несколько кластерных методов. Для анализа полученных кластеров созданы алгоритмы расчета функций распределения числа связанных пар по времени и азимуту (рис. 9).

Оценку “скупенности” эпицентров можно проводить с помощью индекса Моришита. Для проверки их “связности” по времени (dT) в EEDB предусмотрена возможность сравнительного анализа распределений реальных данных с классическим случайным распределением (например, рис. 5-1).

Миграцию сейсмичности до и после заданного момента времени можно также проследить по цепочке географических центров распределения сейсмичности на заданных временных интервалах.

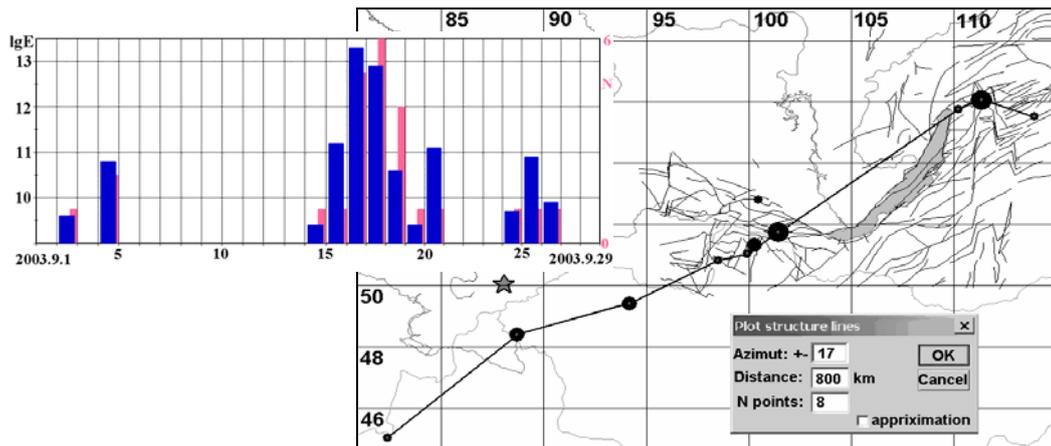


Рис. 8. Цепочка событий 14-20 сентября 2003 г. с $M \geq 3$ перед Чуйским толчком (обозначен звездочкой), выявленная алгоритмом поиска линейных структур. Нанесена система разломов Байкальского региона (по Шерману).

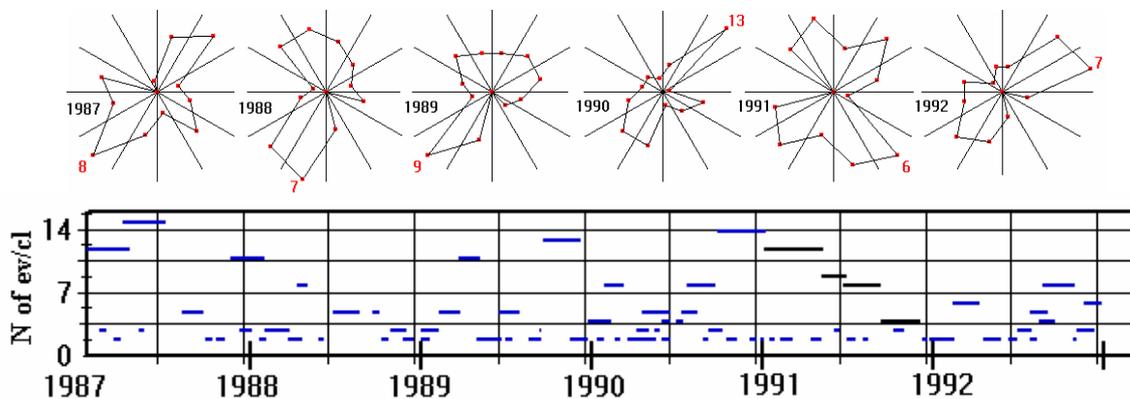


Рис. 9. Анализ выборочных данных каталога ВАИК (1987-1992г, $2 \leq M \leq 4$, без афтершоков): вверху – розы распределения азимутов кластеризованных пар, отчётливо выявляющие ЮВ–СЗ направленность в 1991 г.; внизу – протяженность и мощность (число событий) выявленных кластеров: самые протяженные в 1991 г. кластеры образуют последовательности из 12, 9, 8 и 4 событий.

Создан блок статистического анализа параметров *механизмов* землетрясений в виде временного распределения: значений угла между простиранием плоскости разлома и вектором подвижки, относительного количества механизмов разного типа и т.д.

В заключении перечислены основные результаты работы.

О ЛИЧНОМ ВКЛАДЕ АВТОРА.

Создание системы GIS-EEDB явилось результатом многолетней работы большой группы исследователей и программистов. Общие подходы формирования баз данных были заложены в 90-х годах во время работы автора в лаборатории ММВЦ

ИВМиМГ СО РАН [Гусьяков, **Осипова (Михеева)**, 1992]. Первые информационно-экспертные программные комплексы обработки и визуализации данных разрабатывались автором диссертации в рамках автоматизированных ГИС по моделированию волн цунами [Gusyaikov, Marchuk, **Osipova (Mikheeva)**, 1994; Gusiakov, Ivaikin, Marchuk, 1999].

Выбор, создание и исследование методов анализа сейсмологических данных в процессе разработки Экспертной базы данных по землетрясениям (EEDB) происходило совместно с научным руководителем [Dyadkov, **Mikheeva**, 2009, 2010] и другими сотрудниками лаборатории [Методы и программные средства..., 2009]. Вклад автора диссертации в создание системы «GIS-EEDB» состоит в разработке сейсмологической базы данных, адаптации прежних программных разработок (ГИС, СУБД) к новой предметной области, в создании, реализации и интегрировании комплекса математических и программных алгоритмов EEDB, перечисленных в пунктах «*Научная новизна*» и «*Основные результаты*», разработке наглядного, соответствующего предметным требованиям, человеко-машинного интерфейса.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ

В рамках создания информационно-вычислительного комплекса «Экспертная база данных землетрясений» (EEDB):

1. Разработаны классы и структуры, адекватно отражающие формат, состав и содержание сейсмологических каталогов, необходимых для расчета основных характеристик сейсмогеодинамического процесса;
2. Оптимизированы известные математические алгоритмы и реализованы новые программные алгоритмы предварительной обработки сейсмологических данных, в частности, в процедуре удаления афтершоков предложено фиксировать пространственную структуру их распределения с помощью эллипса равной вероятности;
3. Предложен и реализован математический алгоритм расчета и построения последовательности зональных карт с использованием линейной интерполяции;
4. Разработаны новые алгоритмы и программы анализа сейсмологических данных на основе методов статистического анализа, учитывающие допустимые погрешности оценок и минимальный объем выборки;
5. Организовано функциональное взаимодействие систем комплекса EEDB (баз данных, программ их управления и изучения), путем их интеграции с географической программной оболочкой, реализующей псевдо - трехмерную графику;

6. Предложена новая система управления исследованием сейсмологических данных на основе произвольного комплексирования характеристик сейсмического процесса, и одновременного рассмотрения их поведения на различных участках пространства и времени;
7. Создан эффективный, многооконный и оптимально информативный человеко-машинный интерфейс для разномасштабного сейсмогеодинамического исследования.

СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Возможное влияние землетрясений в Северном Тибете и близ о. Хоккайдо на процесс подготовки Алтайского землетрясения 2003 года / П.Г. Дядьков, Л.А. Назаров, Л.А. Назарова, **А.В. Михеева** [и др.] // Физическая мезомеханика. – 2006. – Т.9, - № 1. – С. 67-72.
2. Стадии подготовки Алтайского землетрясения 27.09.2003г., $M_w=7.3$, и связанные с ними изменения состояния сейсмогенной среды / П.Г. Дядьков, О.А. Кучай, **А.В. Михеева** [и др.] // Физическая мезомеханика. – 2010. – Т.13, - № 6. – С. 78-82.
3. **Osipova (Mikheeva) A.V.** Expert tsunami database for the Kuril-Kamchatka region / V.K. Gusiakov, An.G. Marchuk, A.V. Osipova // Computing Center Bulletin, Series: Mathematical Problem in Geophysics Novosibirsk. NCC Publisher, 1994. - Вып. 1. - N.1. – С. 65-77.
4. **Mikheeva A.V.** The EEDB - Expert earthquake database for seismic-geodynamic research / P.G. Djadkov, A.V. Mikheeva // Bulletins of the Novosibirsk ICMG. Math. model. in geoph. – 2010. – N.13. – P. 15-30.
5. **Mikheeva A.V.** Research methods of the global and regional earthquakes catalogues / P.G. Djadkov, A.V. Mikheeva // Materials of the international conference "Electronic Geophysical Year: State of the Art and Results" (г. Переславль-Залесский, 3 – 6 июня 2009), edited by V. Nechitailenko. - GC RAS, Moscow, 2009. – P. 117.
6. **Михеева А.В.** Применение экспертной системы EEDB при изучении процессов подготовки сильных землетрясений / П.Г. Дядьков, А.В. Михеева // 7-я международная конференция памяти А.П.Ершова. Рабочий семинар «Научно-программное обеспечение» (Новосибирск, Академгородок, 15-19 июня 2009). – Новосибирск, ООО «Сиб. науч. изд-во», 2009. – С. 120-125.
7. Методы и программные средства изучения сейсмического процесса в геоинформационной системе EEDB / П.Г. Дядьков, **А.В. Михеева** [и др.] // Научно-практическая конференция ЕАГО «Сейсмические исследования земной коры. Пузыревские чтения – 2009»

- (Новосибирск, 22-25 ноября 2009г.) : Сб. докл. - Новосибирск: Изд-во ИНГГ СО РАН, 2009. – С. 291-294.
8. **Михеева А.В.** Геоинформационно-экспертные методы анализа сейсмологических данных [Электронный ресурс] / П.Г. Дядьков, А.В. Михеева // Сб. тр. Второй российской конференции с международным участием «Технические и программные средства систем управления, контроля и измерения (теория, методы, алгоритмы, исследования и разработки)» – М., 2010. – С. 430-437 - 1 электр. опт. диск (CD-ROM).
 9. **Михеева А.В.** Методы выявления пространственного группирования землетрясений при сейсмо-геодинамических исследованиях в информационной системе GIS-EEDB / **А.В. Михеева**, П.Г. Дядьков, А.Г. Марчук // XI всероссийская конференция с участием иностранных ученых «Проблемы мониторинга окружающей среды» (Кемерово, 24-28 октября 2011 г.) : сб. докл. - Кемерово: КемГУ, 2011. - С. 166-171.
 10. **Mikheeva A.V.** Methods of complex analysis used for learning the seismicity regularities and anomalies [Электронный ресурс] / A.V. Mikheeva, P.G. Djadkov // 25th International cartographic conference of the ICA (Paris, 3-8 July 2011) - Paris, 2011. – 1 электронный носитель (запоминающее устройство для USB). – Digital conference proceedings.
 11. Дядьков П.Г. Методы выявления пространственного группирования землетрясений в сейсмогеодинамическом исследовании районов Центральной Азии / П.Г. Дядьков, **А.В. Михеева** // Математические методы распознавания образов: 15-я всероссийская конференция (Петрозаводск, 11-17 сентября 2011 г.) : сб. докл. - М: МАКС Пресс, 2011. – Р. 560-563.
 12. **Mikheeva A.V.** Expert Earthquake Data Base - geographic shell for seismological researches / A.V. Mikheeva // 24-th international cartographic conference (Santiago de Chile, 15-21st of November, 2009), Santiago de Chile, 2009. – P. 123.
 13. Дядьков П.Г. Аномалии в режиме слабой сейсмичности перед сильными землетрясениями в южных районах Сибири / П.Г. Дядьков, Ю.М. Крупицкая, **А.В. Михеева** // Казахстанско-Российская межд. конференция "Геодинамические, сейсмологические и геофизические основы прогноза землетрясений и оценки сейсмического риска" (Алматы, 22 -24 июня 2004 г.) : тез. докл. - Алматы, 2004. – С. 53-54.
 14. Kuznetsova J.M. Seismic quiescence before strong earthquakes in regions with different geodynamic regime / J.M. Kuznetsova, P.G. Djadkov, **A.V. Mikheeva** // Asian Seismological Commission, VI General Assembly (Bangkok, Thailand, 7-10 November, 2006) : тез. докл. - Bangkok, 2006. – P. 178.

15. Дядьков П.Г. Аномалии слабой сейсмичности / П.Г. Дядьков, Ю.М. Кузнецова, **А.В. Михеева** // Современная геодинамика и опасные природные процессы в Центральной Азии. - Вып. 3 : Материалы всероссийского совещания «Современная геодинамика и сейсмичность Центральной Азии: фундаментальный и прикладной аспекты». – Иркутск: ИЗК СО РАН, 2005. – С. 185-186.
16. Djadkov P.G. Anomalous behavior of seismicity before the Sumatra-Andaman earthquake of 26 December 2004, M=9.0 / P.G. Djadkov, J.M. Kuznetsova, **A.V. Mikheeva** // Asian Seismological Commission, VI General Assembly (Bangkok, Thailand, 7-10 November, 2006) : тез. докл. - Bangkok, 2006. – P. 174-175.
17. **Осипова (Михеева) А.В.** База данных по землетрясениям и цунами Курило-Камчатского региона / В.К. Гусяков, А.В. Осипова // Препринт 976. – Новосибирск, ВЦ СОРАН, 1992. – 25 с.

Михеева А.В.

**ПРОГРАММНО-АЛГОРИТМИЧЕСКИЙ ИНСТРУМЕНТАРИЙ
ПОДГОТОВКИ И АНАЛИЗА СЕЙСМОЛОГИЧЕСКИХ ДАННЫХ
В ИНФОРМАЦИОННО-ВЫЧИСЛИТЕЛЬНОМ
КОМПЛЕКСЕ ЕЕДВ**

Автореферат

Подписано в печать

Объем 1,1 уч.-изд. л.

Формат бумаги 60 × 90 1/16

Тираж 100 экз.

Отпечатано в ЗАО РИЦ «Прайс-курьер»

630090, г. Новосибирск, пр. ак. Лаврентьева, 6, тел. 34-22-02

Заказ №135