ТЕОРИЯ ОБРАБОТКИ ГЕОФИЗИЧЕСКИХ ДАННЫХ

Учебно-методическое пособие

Министерство науки и высшего образования Российской Федерации
Сибирский федеральный университет

ТЕОРИЯ ОБРАБОТКИ ГЕОФИЗИЧЕСКИХ ДАННЫХ

Учебно-методическое пособие

Красноярск
СФУ
2021

УДК 550.837(07)
ББК 26.2я73
Т338

Т338 **Теория обработки геофизических данных :** учеб.-метод. пособие / сост.: К. В. Симонов, С. А. Перетокин, Т. В. Рублева. – Красноярск : Сиб. федер. ун-т, 2021. – 100 с.

Изложены теоретические основы математических и информационных методов обработки геофизических данных при решении геофизических задач, а также основные принципы и актуальные задачи исследования природной геосреды методами математического и информационного моделирования по данным геомониторинга и натурного эксперимента геофизических полей. Рассмотрены основные сведения о моделях геофизических объектов, алгоритмическом и программном обеспечении. Включены вычислительные методики по обработке и анализу реальных данных геомониторинга очаговых зон сильных землетрясений, а также способы эффективной визуализации данных геомониторинга опасных природных процессов.

Предназначено для магистрантов направления подготовки 03.04.02 «Физика» профиль 03.04.02.04 «Физика Земли и планет», выполняющих практические работы, аспирантов и специалистов, изучающих применение методов математического и информационного моделирования в геофизических исследованиях.

УДК 550.837(07)
ББК 26.2я73

Электронный вариант издания
см.: http://catalog.sfu-kras.ru

© Сибирский федеральный
университет, 2021

СОДЕРЖАНИЕ

Введение 7
1. Технология вычислительного эксперимента 9
2. Комплексный анализ сейсмических сигналов 14
2.1. Метод многопараметрической нелинейной регрессии 15
2.2. Корреляционный анализ возможности восстановления сейсмограмм сильных землетрясений 18
2.3. Предварительный анализ сейсмических данных нейросетями 21
3. Вейвлет-преобразование сейсмического сигнала 29
3.1. Вейвлет-анализ данных для исследования сейсмических колебаний от землетрясений 29
3.2. Быстрое вейвлет-преобразование сейсмических сигналов 36
4. Этапы получения вейвлет-спектра исследуемого процесса 41
4.1. Интерпретация вейвлет-спектра сейсмического события, произошедшего в районе г. Дивногорска 42
5. Алгоритмы выделения и обработки слабых сейсмических сигналов на базе ЛСГ 49
5.1. Применение методов фильтрации для выделения и обработки сейсмического сигнала 50
5.2. Обработка геофизических данных сейсмического события 51
6. Исходные данные сейсмических сигналов ЛСГ, схема их обработки и моделирования 60
6.1. Исходная информация для моделирования сигналов локальной сейсмической группы 60
6.2. Обработка и анализ сейсмических сигналов на основе ЛСГ 63
7. Моделирование данных каталога землетрясений нейросетевыми технологиями 65
7.1. Анализ предвестников землетрясений 65
7.2. Моделирование каталога землетрясений 68
7.3. Общий принцип построения самонастраивающегося датчика случайных величин 68
7.4. Нейросетевой анализ каталога землетрясений 70
7.5. Статистические характеристики модельных распределений сейсмических событий 73

8. Моделирование связей геолого-геофизических факторов
с сейсмичностью 78
8.1. Модель и алгоритм метода упругих сеток 78
8.2. Составление карт сейсморайонирования 82
Заключение 88
Глоссарий 89
Список литературы 95

СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ

h_t – оценка

$\widetilde{\alpha}_t$ – измеренные выходные переменные

$\alpha_t(x, A_t)$ – прогнозируемые выходные переменные

FT – преобразования Фурье

WT – вейвлет-преобразование

$\varphi(t)$ – материнский вейвлет

$f(t)$ – вейвлет-преобразование сигнала

$\psi(t)$ – отцовский вейвлет

$\{\varphi_{jk}(t)\}$ – базис *WT*

DWT – дискретное преобразование *WT*

БПФ – быстрое преобразование Фурье

IDWT – обратное вейвлет-преобразование

CWT – непрерывное вейвлет-преобразование

s – параметр масштабирования

$E_w(a,b)$ – плотность энергии

$E_w(a)$ – спектр энергии

$E_F(\overline{w})$ – подобные максимумы спектра Фурье

$I_W(a,t)$ – мера локальной перемежаемости

$C_W(a,t)$ – мера контрастности анализируемого сигнала

N – длина записи во временных отсчетах

A_n – массив данных

a – масштабный коэффициент

b – параметр сдвига

N – число дискретных отсчетов на промежутке

M – величина выборки масштабов

$L(a_i)$ – окно сходимости базисного вейвлета

δ – порог сходимости

T_i – период

δ_W – приемлемая погрешность

Δx – оптимальный шаг дискретизации

ЛСГ – локальная сейсмическая группа

f – набор частот

m – размер частотной выборки

F_{jl} – взаимный спектр Фурье сигналов в точках j и l

x_i, y_i – координаты i-ой сейсмической станции

τ_i – станционная поправка

c – скорость волны

φ – азимутальный угол

ДПД – детектор прямолинейного движения (*ДПД*-фильтр)

Z – вертикальная составляющая сейсмического сигнала

R – радиальная составляющая сейсмического сигнала

P_P – оператор поляризации

K – нормирующий множитель

$C(T)$ – оператор фильтра для каждого момента времени t

I – интенсивность по шкале MSK-64

E – энергетический класс землетрясения

R – расстояние от рассматриваемой точки до эпицентра землетрясения

λ, μ – коэффициенты упругости, отвечающие за растяжение и изогнутость сетки

D_1 – мера близости расположения узлов сетки к данным

D_2 – мера растянутости сетки

D_3 – мера изогнутости (кривизны) сетки

ВВЕДЕНИЕ

Компьютерные методы обработки геофизических данных – активно развивающееся направление современных вычислительных технологий. Его актуальными задачами являются: разработка методов аппроксимации данных наблюдений на основе нейросетевых технологий, вейвлет-преобразование пространственно-временных рядов, а также способы эффективной визуализации данных геомониторинга опасных природных процессов. Современные информационные и вычислительные технологии, создание специализированных баз данных и картографического материала на основе геомониторинга геофизических процессов, с учетом последствий чрезвычайных ситуаций природно-техногенного характера, позволяют решать актуальные проблемы безопасности в науках о Земле на новом качественном уровне.

В качестве объектов изучения, в рамках учебно-методического пособия, становятся данные, полученные при регистрации геофизических полей различной природы. Предметом изучения являются математические модели и методы математического и информационного моделирования сложных геофизических явлений и процессов с применением современной вычислительной техники.

После построения математических моделей проводится их численное исследование в виде вычислительных экспериментов. Важно отметить, что численное исследование моделей позволяет определять разнообразные характеристики геофизических процессов. В то же время вычислительный эксперимент имеет определенные преимущества перед физическим экспериментом. Он предполагает окончательную проверку и должен рассматриваться как средство сокращения объемов и времени экспериментальной обработки.

Особое внимание уделяется следующим вопросам:

- современное изучение геосреды;
- теоретические основы методов обработки данных геомониторинга;
- модельные и экспериментальные исследования геофизических явлений и процессов;

- вейвлет-преобразования сейсмических сигналов и геоструктур;
- спектрально-временной анализ слабых сигналов;
- построение нейросетевых моделей предвестников;
- нейросетевой анализ каталога землетрясений;
- обработка, восстанавление и интерпретация данных наблюдений с учетом специфики решаемых геофизических задач;
- описание вычислительных методик обработки геоданных в очаговых зонах сильных землетрясений.

Учебно-методическое пособие составлено с учетом современных научных достижений и исследований в этой предметной области, а также в соответствии с требованиями Федерального государственного образовательного стандарта высшего профессионального образования. Освоение учебной дисциплины «Теория обработки геофизических данных» базируется на материале предшествующих естественно-научных, общепрофессиональных и специальных дисциплин подготовки бакалавров.

Настоящее учебно-методическое пособие рекомендуется при изучении курса лекций по дисциплине «Теория обработки геофизических данных», выполнении практических заданий и научно-исследовательской работы магистрантами, а также при подготовке магистерских диссертаций направления 03.04.02 «Физика» профиля 03.04.02.04 «Физика Земли и планет».

1. ТЕХНОЛОГИЯ ВЫЧИСЛИТЕЛЬНОГО ЭКСПЕРИМЕНТА

Изучение геофизических явлений и объектов базируется на модельных и численных исследованиях. В связи с этим методом постижения сложных геофизических процессов на основе математического моделирования является *вычислительный эксперимент*.

Понимая вычислительный эксперимент в узком смысле, как создание и изучение математических моделей исследуемого геофизического объекта с помощью вычислительных средств, можно выделить в качестве основы триаду *модель – алгоритм – программа*. В широком (методологическом) смысле под вычислительным экспериментом понимается новая технология научных исследований. Отметим, что исторически вычислительный эксперимент как новая методология исследования «состоялся» только после того, как удалось на каждом из этапов традиционной цепочки эффективно использовать вычислительную технику.

Основные этапы вычислительного эксперимента показаны на рис. 1.

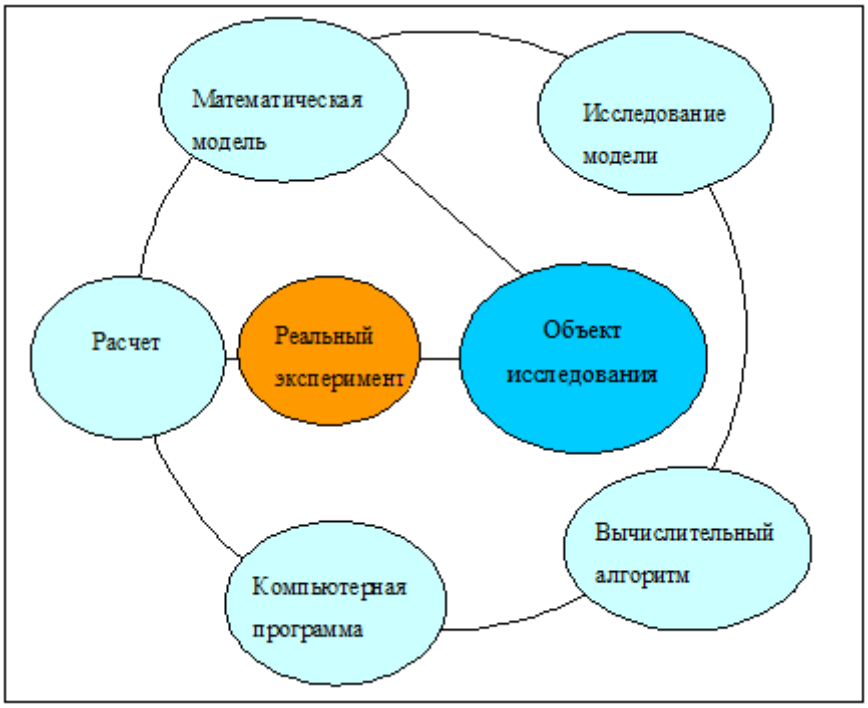

Рис. 1. Технология вычислительного эксперимента

Для исследуемого геофизического объекта (явления) сначала строится математическая модель, которая базируется на известных фундаментальных физических моделях и учитывает основные закономерности и связи, присущие изучаемому геофизическому явлению. В этом случае сам объект (или явление) называется *исходной моделью*.

Вычислительный эксперимент предусматривает исследование группы близких моделей. Вначале задается простая модель, но с определенными допущениями. Во-первых, данная модель может быть достаточно содержательной и полной с точки зрения описания исследуемых процессов. А во-вторых, она, очевидно, должна быть близка к экспериментальным данным.

На последующих этапах, в процессе проведения вычислительного эксперимента, модель уточняется с помощью новых факторов и т. д. Поэтому можем говорить об упорядоченном наборе (об иерархии) математических моделей, каждая из которых с той или иной точностью описывает действительность. После построения математической модели традиционными средствами прикладной математики проводится предварительное исследование математической модели.

В рамках каждой используемой наиболее простой модели необходимо добиваться согласия с экспериментом. Это и является целью вычислительного эксперимента. На этом этапе с доступной полнотой, на принятом в математике уровне строгости решаются вопросы о корректности полной задачи в математическом смысле.

Основное содержание предварительного исследования математической модели состоит в выделении более простых (модельных) задач и их всестороннем исследовании. Модельные математические задачи в цикле вычислительного эксперимента строятся для двух различных целей:

- для качественного исследования полной задачи (следовательно, и исследуемого объекта);
- для проверки и тестирования вычислительных алгоритмов приближенного решения полной задачи.

При качественном исследовании модельных (упрощенных) задач изучаются вопросы множественности решения, его устойчивости и т. д. Большое значение имеют также точные частные решения существенно нелинейных задач, асимптотические решения и т. д. Таким образом, здесь применяется обычный математический арсенал теоретического исследования проблемы. На следующем этапе вычислительного эксперимента строится дискретная задача и численный метод решения этой дискретной задачи. Математическая модель включает в себя, как правило, уравнения с частными производными (ядро математической модели), системы дифференциальных и алгебраических уравнений.

Вычислительный эксперимент характеризуется двумя особенностями, которые необходимо учитывать при создании адекватного программного обеспечения. Это, во-первых, *многовариантность* расчетов в рамках фиксированной математической модели и, во-вторых, *многомодельность*. Здесь нельзя обойтись одной программой на компьютере, нужно иметь возможность легко менять ее для решения близких задач.

Программное обеспечение вычислительного эксперимента базируется на использовании комплексов и пакетов прикладных программ. Комплекс программ предназначен для решения близких по своей математической природе задач из данной предметной области. Комплекс программ включает в себя библиотеку программных модулей (в большей или меньшей степени независимых), из которых комплектуются рабочие программы. В комплексах прикладных программ сборка программ из модулей осуществляется в соответствии с поставленной задачей.

В пакетах прикладных программ для сборки используются системные средства компьютера, что позволяет в значительной степени автоматизировать этот процесс. Пакеты прикладных программ, рассматриваемые как технология решения задач в рамках вычислительного эксперимента, позволяют наиболее эффективно использовать созданный программный продукт. В наибольшей степени основные особенности вычислительного эксперимента учитываются при использовании объектно-ориентированного программирования, а также современных языков программирования.

Затем в цикле вычислительного эксперимента проводится серия расчетов на компьютерах при изменении тех или иных параметров задачи. Полученные данные анализируются и интерпретируются с участием специалистов в данной прикладной области. Обработка результатов проводится с учетом имеющихся теоретических представлений и экспериментальных данных. Она осуществляется во многом в традициях классического натурного эксперимента. Экспериментальные данные представляются в виде таблиц, графиков, карт, видео и т. д.

Необходимо иметь в виду, что объемы обрабатываемой информации, в связи с детализацией полученных результатов, в рамках вычислительного эксперимента весьма значительны. В вычислительном эксперименте проблемы хранения и обработки информации имеют все возрастающее значение. На этапе анализа результатов становится ясным, удачно ли выбрана математическая модель и ее вычислительная реализация. Если есть необходимость, то модели и численные методы уточняются, и весь цикл вычислительного эксперимента повторяется.

Таким образом, основными элементами вычислительного эксперимента являются:

- проведение натурного эксперимента;
- построение математической модели;
- выбор и применение численного метода для нахождения решения;
- построение программной реализации математической модели;
- нахождение численного решения;
- обработка результатов вычислений;
- сравнение с результатами натурного эксперимента;
- принятие решения о продолжении натурных экспериментов;
- продолжение натурного эксперимента для получения данных, необходимых для уточнения модели;
- накопление экспериментальных данных.

Видоизмененная (в зависимости от специфики задачи) цепочка традиционных этапов вычислительного эксперимента, реализованная в виде единого программного комплекса, и составляет «технологию» вычислительного эксперимента.

Контрольные вопросы и задания

1. Как можно обосновать суть вычислительного эксперимента?
2. Какие основные требования должны учитываться при построении математической модели?
3. Перечислите и проанализируйте основные этапы вычислительного эксперимента.
4. Какова роль группы близких моделей в вычислительном эксперименте?
5. Для каких целей строятся модельные математические задачи?
6. Опишите связи вычислительного и натурного экспериментов.
7. Как организованы базы данных в рамках вычислительного эксперимента?
8. Какие методы обработки данных используются в вычислительном эксперименте?

Литература: раздел 1, основная литература п. 1–8; дополнительная литература п. 1, 9, 12.

2. КОМПЛЕКСНЫЙ АНАЛИЗ СЕЙСМИЧЕСКИХ СИГНАЛОВ

Геодинамические задачи весьма сложны и их решение в общем виде с учетом всего многообразия факторов невозможно. На начальном этапе научного исследования строится математическая модель реального геофизического объекта. Для этого необходимо корректно сформулировать задачу и сделать целенаправленный выбор математической модели.

В случае оценки сейсмической опасности малоизученной в сейсмическом отношении области необходимо решить задачу эффективного и экономичного анализа данных геомониторинга очаговых зон. Как правило, при оценке сейсмического режима используются данные о предвестниках землетрясений различной природы, а изучаемый объект и система наблюдений имеют определенную пространственную структуру. Также необходимо отметить определенную «зашумленность» изучаемых сигналов и «стохастичность» исследуемых геофизических процессов.

Для решения данной задачи предлагается следующая процедура анализа (фильтрации) сейсмических сигналов, регистрируемых набором датчиков с использованием современных информационно-вычислительных технологий.

На первом этапе предварительного анализа выполняется многопараметрический нелинейный регрессионный анализ каждого сигнала с использованием нейросетевой вычислительной технологии обработки данных. Он позволяет выделять сигнал на фоне помех во временной области без каких-либо априорных предположений о частотном составе анализируемого сигнала. При этом за счет шумовой составляющей происходит частичная фильтрация данных и строится нейросетевой образ амплитудно-частотной характеристики процесса. Отметим также, что на этом этапе целесообразна процедура компактизации данных без потери в частотном диапазоне для хранения и транспортировки больших массивов данных наблюдений.

На втором этапе для описания иерархичной структуры колебаний во всем частотном диапазоне применяется современный математический аппарат вейвлет-анализа сигналов. Масштабно-временная разверстка,

получающаяся в результате вейвлет-преобразования сигнала, позволяет выявлять не только осцилляции с хорошо фиксированными частотами, но и нестационарные колебания, а также локализованные периодичности и др.

Дисперсия коэффициентов вейвлет-преобразования анализируемых данных дает представление о распределении энергии процесса по масштабам. По значениям локальных экстремумов на вейвлет-диаграммах вычисляется спектр размерности анализируемого множества. Фильтрационные свойства этого преобразования позволяют на этом этапе оперировать информацией без потери ее значимой части.

На третьем этапе для обработки наблюдений предлагается два вычислительных метода, реализованных в программном обеспечении: пространственно-временной и спектрально-временной анализ данных. Первый метод позволяет изучать пространственную структуру поля наблюдаемого геофизического процесса и ее изменения по времени для выделения пространственных неоднородностей. Второй метод дает возможность исследовать дисперсионные характеристики распространения изучаемых колебаний в среде.

На конкретных примерах демонстрируются приемы обработки реальных записей, полученных региональной сетью сейсмических станций от местных землетрясений. Также анализируются данные каталога землетрясений в регионе Центральной Сибири. Результаты этого комплексного анализа представляются обнадеживающими для достижения поставленной цели – изучение сейсмического режима в очаговых зонах и оценка степени сейсмической опасности.

2.1. Метод многопараметрической нелинейной регрессии

Рассмотрим вычислительную технологию построения моделей временного сигнала объекта с помощью нейросетей. Известна универсальность такого подхода к анализу данных различной природы – достаточно совокупности одного нелинейного и многих линейных преобразований для аппроксимации любой однозначной функции с любой заданной точностью. Как уже отмечалось, нейросетевой подход – это отыскание частных производных с помощью метода множителей

Лагранжа – позволяет строить такие аппроксимации достаточно эффективно.

Постановка задачи многопараметрической нелинейной регрессии, используемая в данном нейросетевом алгоритме, предполагает, что ищется зависимость между известными входными переменными и прогнозируемыми выходными переменными. Качество аппроксимации определяется величиной функции оценки

$$H = \sum_t h_t(x, A_t), \tag{1}$$

где h_t является величиной оценки для задачи t, x – представляет собой набор параметров регрессии, A_t – входные переменные для задачи t.

Оценка h_t для задачи t чаще всего определяется квадратом расстояния между измеренными выходными переменными $\tilde{\alpha}_t$ и прогнозируемыми выходными переменными $\alpha_t(x, A_t)$:

$$h_t = (\alpha_t(x, A_t) - \tilde{\alpha}_t)\varepsilon(\alpha_t(x, A_t) - \tilde{\alpha}_t). \tag{2}$$

В формуле 2 матрица ε задает коэффициенты неотрицательно определенной квадратичной формы; если эта матрица диагональная, то эти коэффициенты обратно пропорциональны допустимым среднеквадратичным отклонениям прогнозируемых выходных переменных от измеренных. Коэффициенты ε могут быть названы коэффициентами точности. При неточно известных входах естественное обобщение состоит в переходе к функции оценки:

$$H' = H + \sum_t (A_t - \tilde{A}_t)\mu(A_t - \tilde{A}_t), \tag{3}$$

где введены измеренные входные параметры $\tilde{A}_t$, не совпадающие с наилучшими предполагаемыми параметрами $A_{t,}$, определяемыми в ходе оптимизации функции оценки H' по этим параметрам, а также коэффициенты точности входных переменных μ. Специального

рассмотрения заслуживает предельный случай, когда некоторым или всем входам соответствует нулевая точность.

Для сходимости процедуры обучения, реализующей минимизацию оценки, желательно чтобы множество изменяемых при обучении параметров было заведомо компактным. Для этой задачи можно выбрать варианты требований априорной компактности, допускающие содержательную интерпретацию в терминах гладкости зависимости α_t от A_t. Могут быть использованы два варианта условий:

$$\left\langle (\nabla a_t)^2 \right\rangle \le const; \tag{4}$$

$$\left\langle (\nabla a_t)^2 \right\rangle / \left\langle a_t^2 \right\rangle \le const, \tag{5}$$

накладываемых на средние по пространству входов величины квадратов производных входных переменных по выходным – первый вариант (4), либо на отношение этих средних к среднему квадрату величины входных переменных – второе соотношение (5).

Использовалась нейросетевая аппроксимация зависимости выходов от входов следующего вида:

$$a_t^i = b^i + c^i \sum_q \sin(\varphi^{iq} + \sum_j k^{qj} \times A_t^j). \tag{6}$$

Если выбрана данная аппроксимация, то первый вариант условия компактности соответствует требованию:

$$(\sum_i (c^i)^2) \times \sum_{q,j} (k^{qj})^2 \le const,$$

а второй вариант – требованию:

$$\sum_{q,j} (k^{qj})^2 \le const.$$

При выборе варианта нейросетевой аппроксимации часто исходят из требования асимптотической универсальности нейросети – возможности как угодно точно приближать произвольные достаточно гладкие функции входных переменных при увеличении числа подстроечных параметров сети. Можно показать, что при увеличении числа подстроечных параметров данное представление сети переходит в вариант интегрального представления Фурье, аппроксимационные возможности которого хорошо изучены.

При модифицированном функционале оценки реализовывался вариант на основе метода сопряженных градиентов, когда оптимизация проводится на прямом произведении пространства параметров модели и пространства входных переменных по всем задачам, либо вариант, при котором модификация нейросетевой модели и модификация входных переменных проводятся поочередно.

Основанием к использованию второго варианта служит возможность проведения адаптации по каждой задаче поочередно, что экономит вычислительные ресурсы. Согласно результатам тестовых проверок, в ряде случаев удается удовлетворительно восстановить недостающую информацию о входах. В то же время при низкой точности входных данных алгоритм работает неустойчиво и требуется дополнительная регуляризация на множестве входных переменных.

2.2. Корреляционный анализ возможности восстановления сейсмограмм сильных землетрясений

При сильных сейсмособытиях аппаратура, настроенная на слабые сигналы, не записывает сейсмограмму в области больших амплитуд, выходящих за пределы шкалы. Принципиально эту сложность можно обойти установкой дополнительной регистрирующей системы со слабым усилением. Однако, при этом требуются определенные затраты, а дополнительная аппаратура будет почти все время простаивать, используясь только при сильных сейсмособытиях – раз в несколько лет, так что даже контроль за ее работоспособностью может быть утрачен, не говоря уже о возможности настройки на реальные сигналы.

Поэтому представляет интерес альтернативный подход, когда дополнительная аппаратура не устанавливается, а получаемый обрезанный сигнал восстанавливается математическими методами. Возможность данного подхода апробировалась на первых афтершоках Караганского землетрясения. Эти афтершоки были достаточно слабы и укладывались в диапазон регистрирующей аппаратуры.

Записанный сигнал был искусственно обрезан на уровне порядка 15 % от максимума сигнала, т. е. примерно на уровне верхней границы шумового фона. Затем по обрезанному сигналу восстановлен исходный сигнал, результаты показаны на рис. 2. Как можно видеть, восстановленный сигнал по положению и амплитуде пиков соответствует исходному, хотя некоторые искажения имеются.

Для оценки связей между исходным сигналом и обработанными сигналами рассчитывались коэффициенты корреляции. Количественные характеристики следующие. Коэффициент корреляции между исходным и обрезанным сигналом составляет +0.894. Возможно это связано с тем, что значительная часть сигнала имеет амплитуду ниже уровня обрезания. Тем не менее корреляция исходного сигнала с восстановленным заметно выше и составляет + 0.947.

Рис. 2. Первый и второй афтершоки Караганского землетрясения 27.10.2000 г., использованные для контроля качества методики восстановления (сейсмостанция «Красноярск», канал Z)

Коэффициент корреляции положительной части обрезанного сигнала с восстановленной его положительной частью составляет + 0.579,

коэффициент корреляции отрицательной части обрезанного сигнала с восстановленной его отрицательной частью составляет + 0.651. И коэффициент корреляции не по всему сигналу, а только по всем восстановленным данным, составляет + 0.913.

Значительное повышение коэффициента корреляции в процессе обработки обусловлено тем, что в выборке присутствуют как положительные, так и отрицательные амплитуды. Причем их знак при восстановлении всегда указывается верно, поскольку аппаратура лишь обрезает амплитуду колебаний, но не искажает знак отклонений.

Очевидно, восстановленные сейсмограммы представляют определенный практический интерес. В ходе исследований были опробованы различные варианты алгоритмов восстановления. Относительно лучшие результаты дает четырехточечная интерполяция многочленами Лагранжа, когда две точки берутся до интервала искаженных данных, а две точки – после.

Метод восстановления данных применен для обработки сейсмограммы Караганского землетрясения, когда шкала была многократно перекрыта. Эпицентр Караганского землетрясения, произошедшего 27 октября 2000 г., находился на линии Главного Саянского разлома. В этом случае единственная возможность «увидеть» форму сигнала обеспечивается посредством интерполяции. На рис. 2, *а–в* представлены реальная запись Караганского землетрясения магнитудой M=5,2 и восстановленный сигнал (сейсмостанция «Красноярск») в трех каналах.

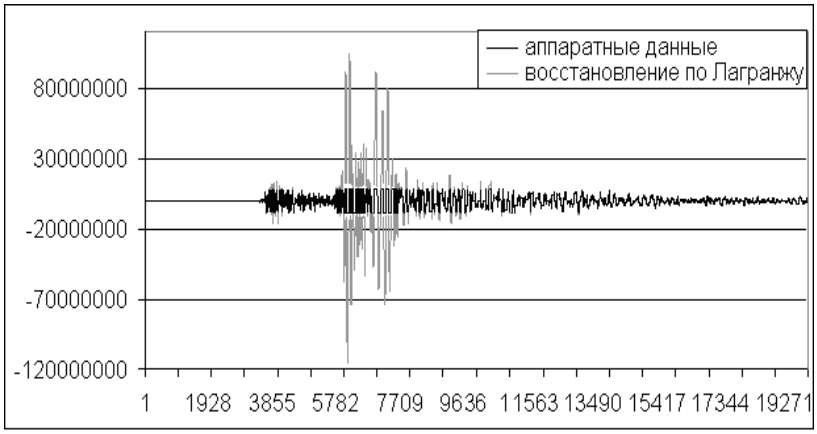

Рис. 2, *а* – Канал 1

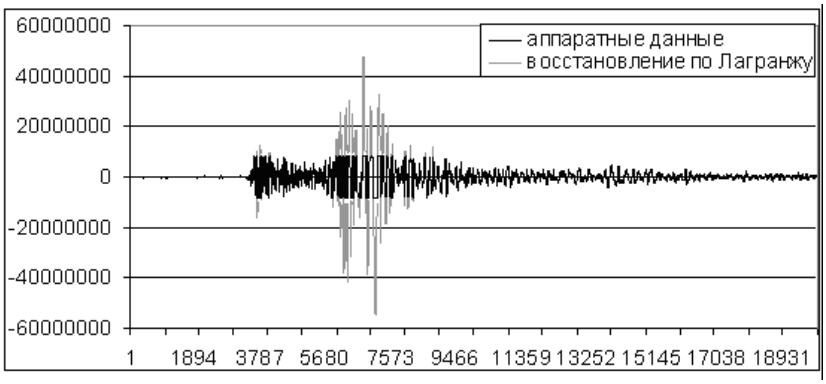

Рис. 2, *б* – Канал 2

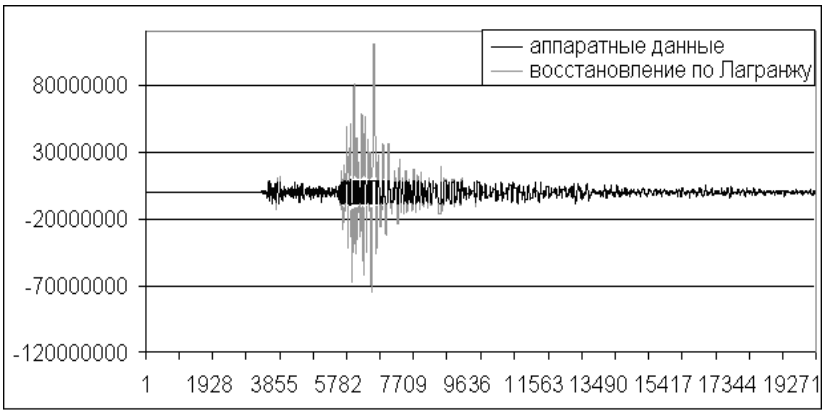

Рис. 2, *в* – Канал 3

2.3. Предварительный анализ сейсмических данных нейросетями

Для эффективной обработки сейсмических данных разработан ряд новых методик анализа записей сейсмических волн во временной области, а именно:

1) процедура нелинейной многопараметрической фильтрации для выделения низкочастотного сигнала на фоне сильных помех;

2) вычислительная схема быстрой компактизации (в 5–10 раз) и декодирования сейсмических данных без потери информации в заданном динамическом диапазоне сейсмического сигнала;

3) методика выделения времени вступления фаз различных сейсмических волн для детального анализа записи сейсмических событий в реальном режиме времени, изучение частотно-временных свойств источника возмущения и трасс распространения.

Результаты реализации процедуры нелинейной многопараметрической фильтрации с целью выделения низкочастотного сигнала на фоне сильных помех детально показаны на рис. 3.

На рис. 3, *а* приведен пример неотфильтрованного сигнала при отсутствии сейсмических событий. Здесь же показан результат нейросетевого сглаживания сигнала.

Графики сглаженных огибающих максимумов и минимумов сигнала в момент отсутствия сейсмических событий показаны на рис. 3, *б*.

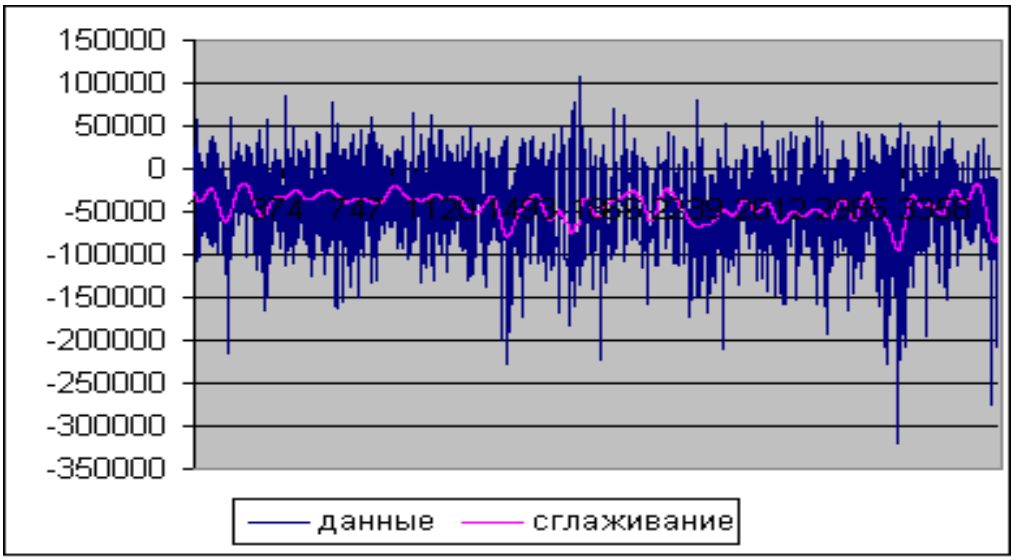

Рис. 3, *а* – пример неотфильтрованного сигнала в отсутствие сейсмических событий и результат нейросетевого сглаживания сигнала

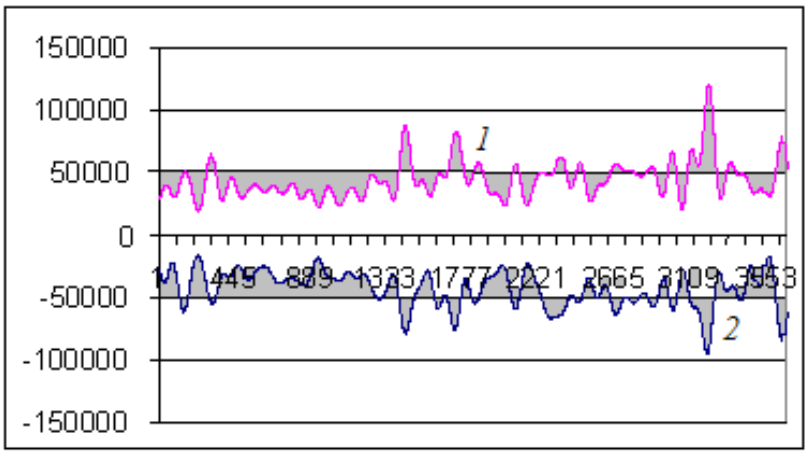

Рис. 3, *б* – сглаженные огибающие максимумов (*1*)
и минимумов (*2*) сигнала в отсутствие сейсмических событий

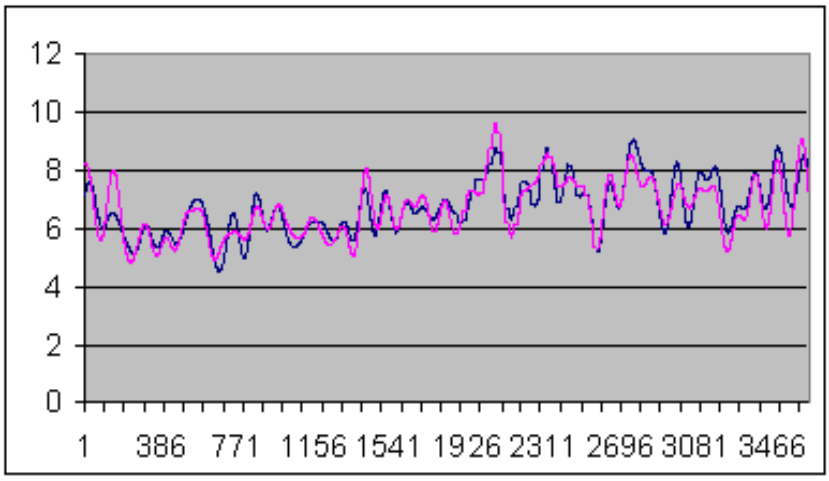

Рис. 3, *в* – сглаженная последовательность интервалов между максимумами и интервалов между минимумами для неотфильтрованного сигнала

На рис. 3, *г* можно видеть зависимость амплитуды сигнала от интервала между локальными экстремумами.

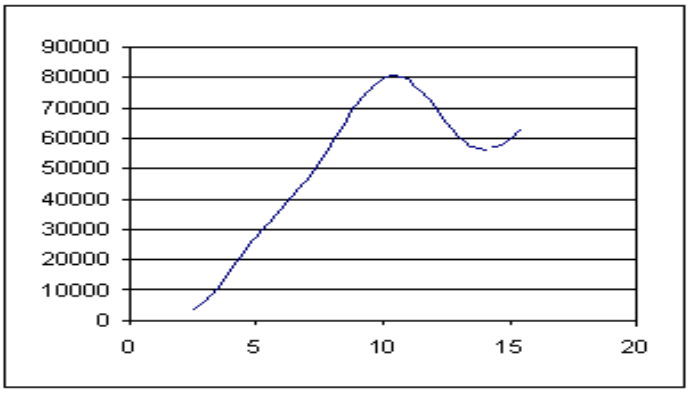

Рис. 3, *г* – зависимость амплитуды сигнала *(вертикальная шкала)* от интервала между локальными экстремумами *(горизонтальная шкала)*

Результаты применения вычислительной схемы быстрой компактизации (в 5–10 раз) и декодирования сейсмических данных без потери информации в заданном динамическом диапазоне сейсмического сигнала продемонстрированы на рис. 4, *а–г*.

На рис. 4, *а* показан отфильтрованный сигнал от зарегистрированного промышленного взрыва на юге Красноярского края.

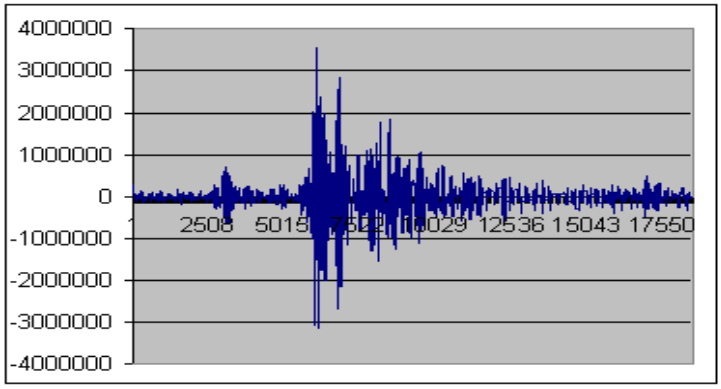

Рис. 4, *а* – отфильтрованный сигнал от зарегистрированного промышленного взрыва на юге Красноярского края

На рис. 4, *б* приведен результат компрессии (*сжатия*) данных и последующего их восстановления. По вертикальной оси отложена нормированная амплитуда сигнала, а по горизонтальной – нормированное время. При компрессии сохранялись только величины локальных экстремумов и временные интервалы между ними. Коэффициент корреляции R между восстановленными и исходными данными составил $R = -0.998$.

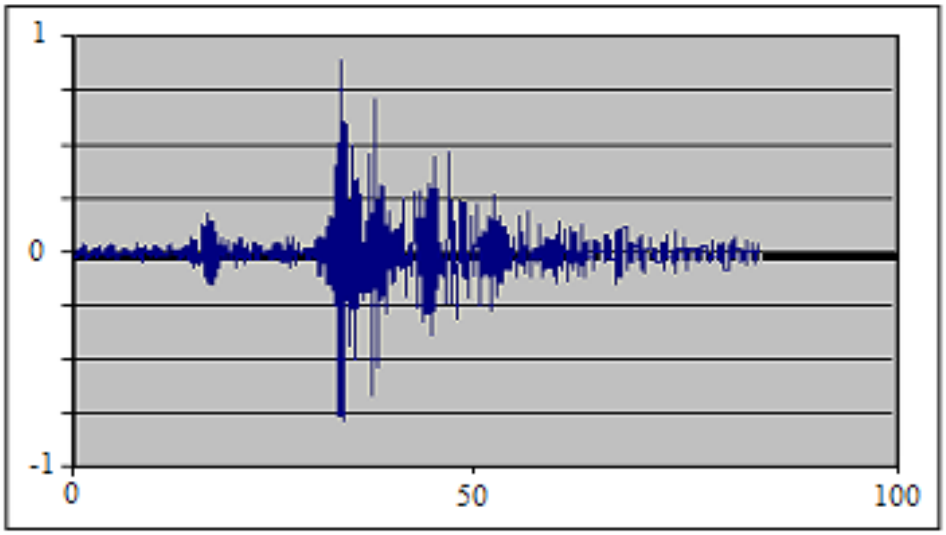

Рис. 4, *б* – результат компрессии данных и последующего восстановления

На рис. 4, *в* показаны графики сглаженной последовательности временных интервалов (периодов) между максимумами (ряд *1*) и минимумами (ряд *2*).

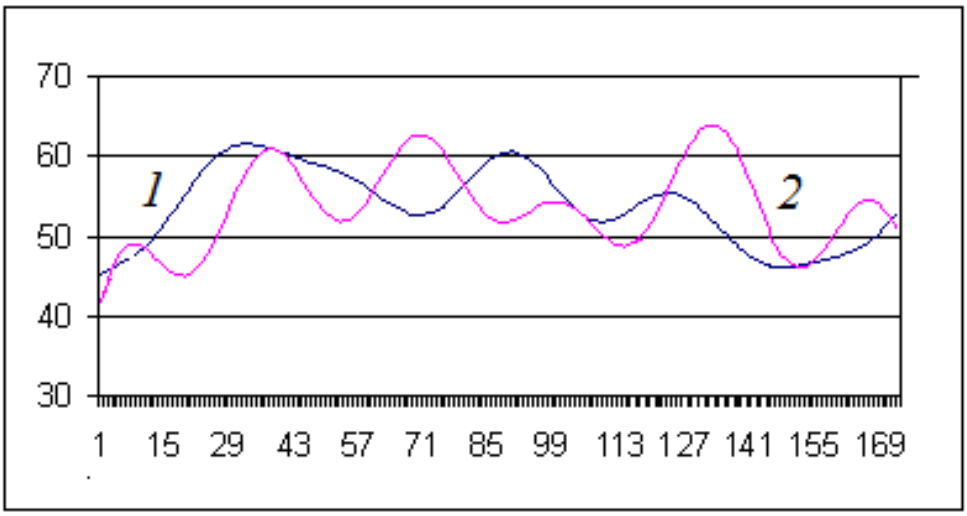

Рис. 4, *в* – сглаженная последовательность временных интервалов между максимумами (ряд *1*) и минимумами (ряд *2*)

На рис. 4, *г* приведены графики сглаженных огибающих максимумов (ряд *1*) и минимумов (ряд *2*) сигнала.

На рис. 4, *д* проиллюстрирован результат восстановления сигнала по сглаженным данным.

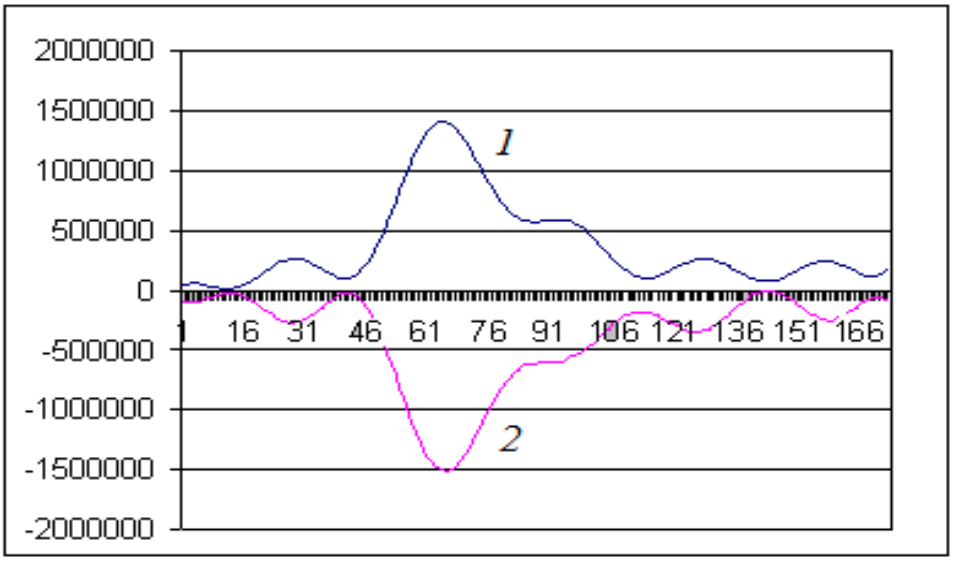

Рис. 4, *г* – сглаженная огибающая максимумов (ряд *1*) и минимумов (ряд *2*) сигнала

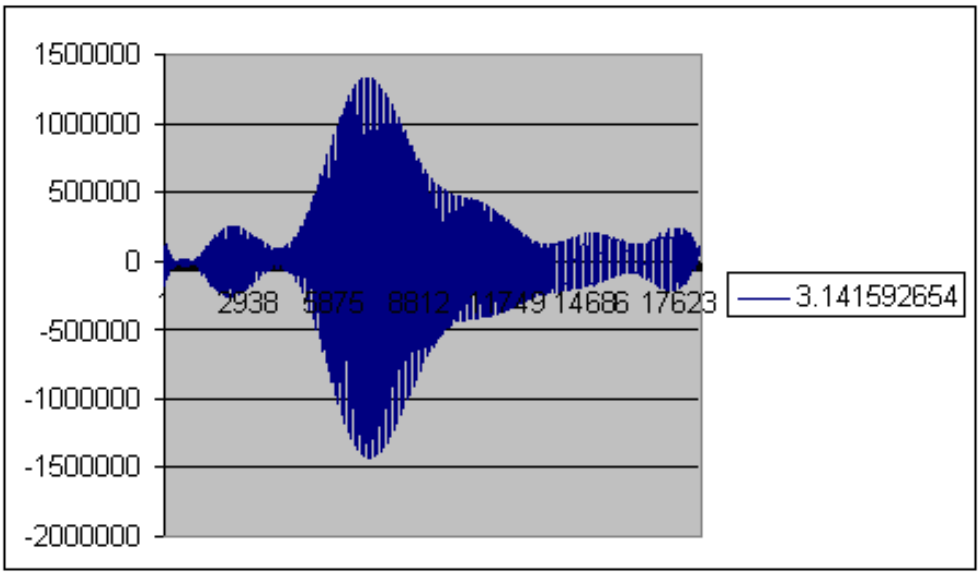

Рис. 4, *д* – результат восстановления сигнала по сглаженным данным

Результаты реализации методики выделения времени вступления фаз различных сейсмических волн с целью детального анализа записи сейсмических событий в реальном режиме времени и изучения частотно-временных свойств источника возмущения и трасс распространения иллюстрируют рис. 5, *а–г*. На рис. 5, *а* показан сейсмический сигнал афтершока Караганского землетрясения (X канал).

На рис. 5, *б* представлены частотно-амплитудная зависимость для сигнала афтершока Караганского землетрясения на основе усредненных исходных данных и график сглаженной зависимости.

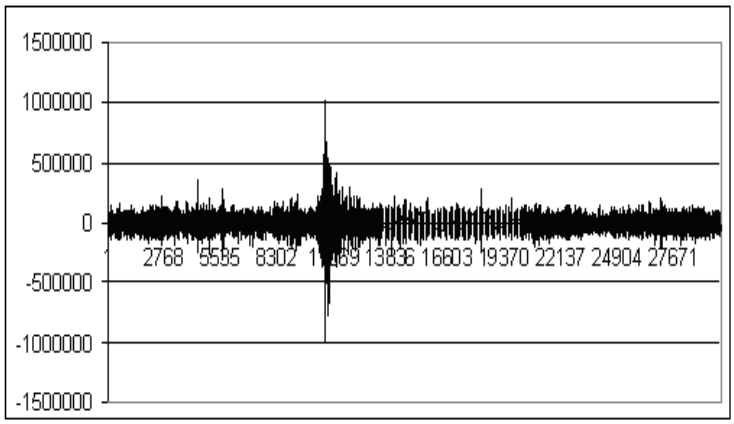

Рис. 5, *а* – сейсмический сигнал афтершока Караганского землетрясения (X канал)

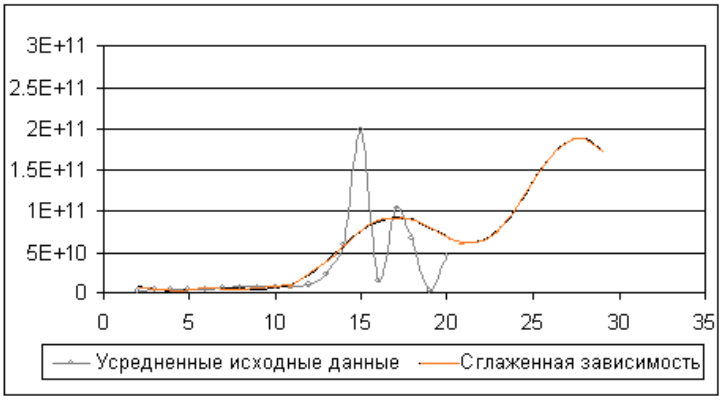

Рис. 5, *б* – частотно-амплитудная зависимость для сигнала афтершока Караганского землетрясения и график сглаженной зависимости

Зависимость квадрата амплитуды от периода колебаний для Караганского землетрясения показана на рис. 5, *в*. Здесь на вертикальной шкале отложены значения квадрата амплитуды, а на горизонтальной – период времени между двумя соседними максимумами.

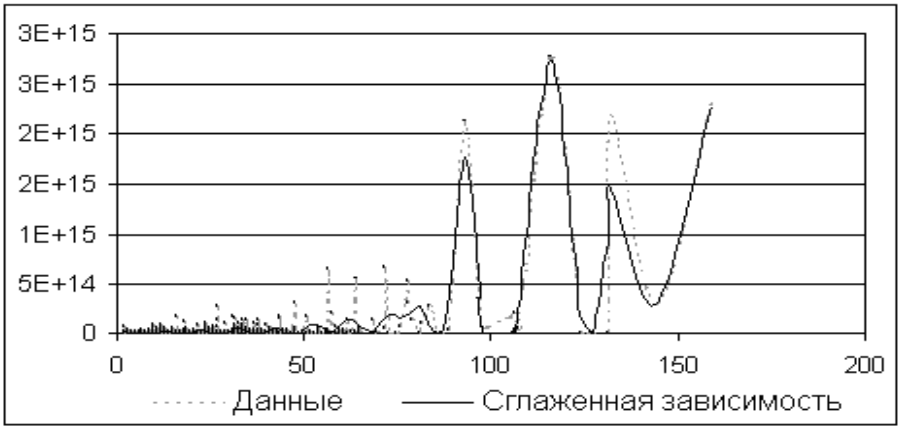

Рис. 5, *в* – зависимость квадрата амплитуды от периода колебаний для Караганского землетрясения

График на рис. 5, *в* построен по средним значениям для амплитуд, соответствующим одним и тем же периодам колебаний. Также на данном рисунке приведен график сглаженной зависимости.

На рис. 5, *г* представлена зависимость периода колебаний от времени в процессе реализации Караганского землетрясения.

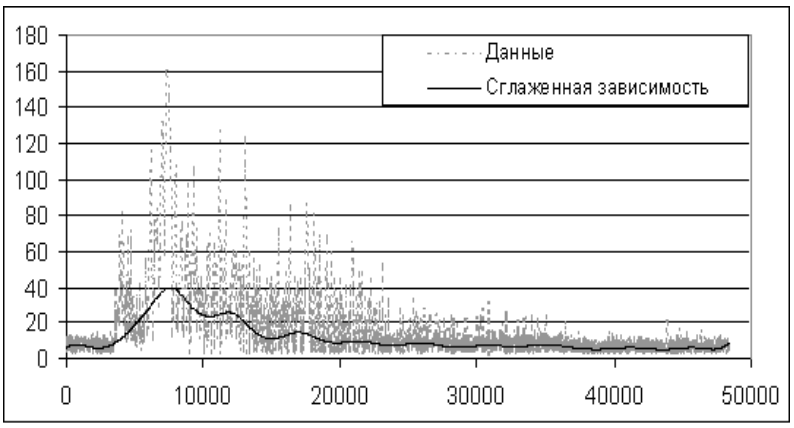

Рис. 5, *г* – зависимость периода колебаний от времени в процессе Караганского землетрясения

Контрольные вопросы и задания

1. Какие основные задачи решаются во время геомониторинга геофизических процессов?
2. Каковы основные требования при геомониторинге очаговых зон землетрясений?
3. Какие предвестники сильных землетрясений наблюдаются? Опишите их.
4. Какие существуют современные каталоги землетрясений?
5. Когда и как образуются сейсмические волны?
6. Чем отличаются объемные волны от поверхностных?
7. На каких принципах основано сейсмическое районирование территории?
8. Перечислите параметры, используемые для оценки геодинамической и геоэкологической опасности.

Литература: раздел 2, основная литература п. 14–16; дополнительная литература п. 13–16, 18–19.

3. ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЕ СЕЙСМИЧЕСКОГО СИГНАЛА

Известно, что термин *вейвлет-анализ* по смыслу аналогичен термину *Фурье-анализ*. В каждом случае речь идет о представлении исследуемого процесса в виде линейной комбинации различных функций, именуемых базисом соответствующего преобразования.

Для вейвлет-анализа данных характерно понятие масштаба, даже графическое представление в виде диаграммы специального вида именуется скейлограмма. Под масштабом следует понимать колебательные процессы различной периодичности.

Полагаем, что низкочастотные колебания имеют более крупный масштаб, а высокочастотные – более мелкий. Вейвлет-анализ данных называют «микроскопом», поскольку он позволяет исследовать каждый масштаб с необходимой и достаточной для него разрешающей способностью.

3.1. Вейвлет-анализ данных для исследования сейсмических колебаний от землетрясений

Для преобразования Фурье (*FT*) базисом являются функции $w_n(t)$, полученные из функции $w(t) = e^{it} = \cos t + i \sin t$ путем масштабного преобразования $w_n(t) = w(nt)$, где $w(t)$ имеет период 2π, т. е. *FT* усредняет спектральные характеристики сигнала.

Вейвлет-преобразование (*WT*) в большой степени позволяет преодолеть перечисленные недостатки *FT*, поскольку базисные функции *WT* обладают свойством временной локализации, т. е. обладают конечной энергией (нормой):

$$E_f = \int_{-\infty}^{\infty} \left| f(t) \right|^2 dt < \infty . \qquad (7)$$

Как и для *FT*, для построения базиса *WT* используется одна функция, именуемая (материнским вейвлетом – $\varphi(t)$). Таким образом,

вейвлет-преобразование сигнала $f(t)$ заключается в разложении функции в виде:

$$f(t) = c_0 \psi(t) + \sum_{j=1}^{N} \sum_{k=0}^{2^j-1} c_{jk} \varphi_{jk}(t), \quad (8)$$

где $\psi(t)$ – называют отцовским вейвлетом, являющимся аналогом константы, а $\{\varphi_{jk}(t)\}$ – базис *WT*, т. е. функции, полученные из материнского вейвлета $\varphi(t)$ путем двух основных операций: сжатия и сдвига, при этом производится временное сжатие в 2^j раз и сдвиг полученной функции на $2^{-j}k$:

$$\varphi_{jk}(t) = 2^{j/2} \varphi(2^j t - k). \quad (9)$$

Вычислительная процедура дискретного преобразования *WT* (*DWT*) состоит в вычислении временной свертки сигнала с $\varphi_{jk}(t)$. Поскольку длина вейвлета для каждого последующего масштаба уменьшается в 2 раза, для *DWT* используется вычислительная процедура, аналогичная быстрому преобразованию Фурье (*БПФ*). Для масштаба $j=1$ определяется 2 коэффициента: c_{10} и c_{11}, для $j=2-4$ коэффициента и т. д., поэтому вычислительный алгоритм *DWT* называют *пирамидальным алгоритмом*.

Для удаления шума производят *DWT* данных наблюдений, обрабатывают полученный образ и производят обратное вейвлет-преобразование (*IDWT*). Алгоритм *IDWT* аналогичен алгоритму *DWT*. Необходимым условием для возможности осуществить восстановление сигнала по его *DWT* путем обратного преобразования является ортогональность базиса. Непрерывное вейвлет-преобразование (*CWT*) имеет вид:

$$CWT_x^{\psi}(\tau,s)=\Psi_x^{\psi}(\tau,s)=\frac{1}{\sqrt{|s|}}\int x(t)\psi^*(\frac{t-\tau}{s})dt, \qquad (10)$$

где s – параметр масштабирования.

Для анализа структуры сигналов удобно пользоваться отображением непрерывного *WT* на скейлограмме. Особенно выразительные скейлограммы получены для базисов, основанных на комплексных вейвлетах. В пространстве действительных функций полная энергия сигнала может быть записана через амплитуды вейвлет-преобразования в виде:

$$E_f=\int f^2(t)dt=C_{\psi}^{-1}\iint W^2(a,b)\frac{dadb}{a^2}. \qquad (11)$$

Плотность энергии сигнала $E_w(a,b)=W^2(a,b)$ характеризует энергетические уровни (уровни возбуждения) исследуемого сигнала $f(t)$ в пространстве (a,b). Одной из основных особенностей вейвлет-преобразования является возможность получать локализованные характеристики и изучать локальные свойства процессов. Так, зная плотность энергии $E_w(a,b)$, можно с помощью окна определить локальную плотность энергии в точке (b_0) или (t_0):

$$E_{\zeta}(a,t_0)=\int E_w(a,b)\zeta\left(\frac{b-t_0}{a}\right)db. \qquad (12)$$

Функция φ поддерживает диапазон около (t_0) и удовлетворяет равенству

$$\int\xi(b)db=1.$$

Если в качестве φ выбрать функцию Дирака, то локальный спектр энергии примет вид:

$$E_\delta(a,t_0) = W^2(a,t_0)\,.$$

Эта характеристика позволяет проанализировать временную динамику передачи энергии процесса по масштабам – обмен энергией между составляющими процесс компонентами разного масштаба в любой заданный момент времени.

Полная энергия распределена по масштабам в соответствии с глобальным спектром энергии коэффициентов вейвлет-преобразования

$$E_w(a) = \int W^2(a,b)db = \int E_w(a,b)db\,. \qquad (13)$$

Его называют также *скейлограммой* дисперсии вейвлет-преобразования. Вейвлет-преобразование, представляющее временную развертку спектра, позволяет получить и более локализованную энергетическую информацию. Скейлограммы, полученные по диапазонам, позволяют проследить эволюцию распределения энергии по масштабам в виде:

$$E_f = C_\psi^{-1} \int \frac{E_w(a)}{a^2} da\,. \qquad (14)$$

Таким образом, величина E_f пропорциональна площади под кривой $\frac{E_w(a)}{a^2}$, а скейлограмма отражает относительный вклад различных масштабов и полную энергию и выявляет распределение энергии процесса по масштабам.

Анализируемая функция имеет конечную энергию, а анализирующий вейвлет – нулевое среднее значение, вследствие чего спектр энергии $E_w(a)$ должен стремиться к 0 на обоих концах шкалы масштабов и должен иметь, по крайней мере, один максимум.

Положение подобных максимумов спектра Фурье $E_F(\overline{w})$ связывается с частотами и соответствующими характерными модами анализируемого сигнала, несущими в себе основную энергию процесса. Максимумы энергетического спектра $E_w(a)$ интерпретируются аналогично – они определяют масштабы процесса, вносящие основной вклад в полную энергию $E_w(a)$.

На практике, если скейлограмма $E_w(a)$ имеет пик при $a = a_0$, то характерный временной масштаб определяется как $\mathrm{d} = \frac{T_0}{2} = \frac{a_0 \pi}{\varpi_\chi}$.

Множитель $\frac{1}{2}$ появился из-за того, что определяется не период, а масштаб элементарного события, детали.

Константа ϖ_χ получена для простой функции. Распространяя результат на произвольный сигнал, даже негармонический, предполагаем, что положение максимума спектра $E_w(a)$ (т. е. выявленный масштаб) можно интерпретировать как среднюю продолжительность элементарного события, вносящего основной вклад в энергию анализируемого процесса.

Выпишем две характеристики, определяемые через плотность энергии, – меру локальной перемежаемости $I_W(a,t)$ и меру контрастности анализируемого сигнала $C_W(a,t)$.

Мера локальной перемежаемости представляется в виде:

$$I_W(a,t) = \frac{E_W(a,t)}{\left\langle E_W(a,t) \right\rangle_t}, \qquad (15)$$

как мера локальных отклонений от среднего поля спектров на каждом масштабе, она позволяет определить степень неравномерности распределения энергии по масштабам. Угловыми скобками здесь обозначено усреднение.

Равенство $I_W(a,t) = 1$ при всех a и t означает, что энергия распределена равномерно на все локальные спектры энергии одинаково; $I_W(a,t) = \alpha$ означает, что вклад компонентов масштаба a в точке t_0 в α раз превосходит усредненный по всем t.

Мера контрастности определяется из следующего соотношения:

$$C_W(a,t) = \frac{E_W(a,t)}{E_W^{'}(a,t)}, E_W^{'}(a,t) = \int_{a'=0}^{a'=a} E_W(a',t)da' \qquad (16)$$

и позволяет определять даже малые изменения в сигнале, когда необходимо, например, выявить структурированность слабого сигнала или слабые вариации на фоне крупной структуры.

Анализируемые с помощью вейвлет-преобразований сейсмические данные представляют собой записи сейсмических волн от зарегистрированных на сейсмических станциях землетрясений, промышленных взрывов и других сейсмических событий на территории Красноярского края.

Массивы данных представляют собой матрицу (3xN), где N – длина записи во временных отсчетах. Каждый из столбцов этого массива содержит информацию – запись сейсмических колебаний по одной из трех компонент сейсмографа (X,Y,Z). Для дальнейшей обработки выделяем каждую компоненту и каждому набору значений ставим в соответствие свой массив данных A_n. Заменяя в приведенных выше соотношениях интегралы суммами, получаем следующий вид расчетных формул:

$$W(a,c) := \left(\sqrt{a}\right)^{-1} \sum_{n=n0}^{n1} \varphi\left(\frac{n-a}{c}\right) \cdot A_n \,. \qquad (17)$$

Для анализа указанных сейсмических сигналов использовался *MHAT* вейвлет:

$$\varphi\left(\frac{n-a}{c}\right) = \left(1 - \left(\frac{n-a}{c}\right)^2\right) \cdot e^{\frac{-1}{2}\left(\frac{n-a}{c}\right)^2} \,. \qquad (18)$$

Из теории следует, что значение амплитуды вейвлет-преобразования в точке (a_0, c_0) тем больше (по абсолютной величине),

чем сильнее корреляция между вейвлетом данного масштаба и поведением сигнала в окрестности $n = c_0$.

Картина коэффициентов демонстрирует, что процесс составляют компоненты разных масштабов: экстремумы $W(a,c)$ наблюдаются на разных масштабах, интенсивность их меняется со временем и масштабом (рис. 6).

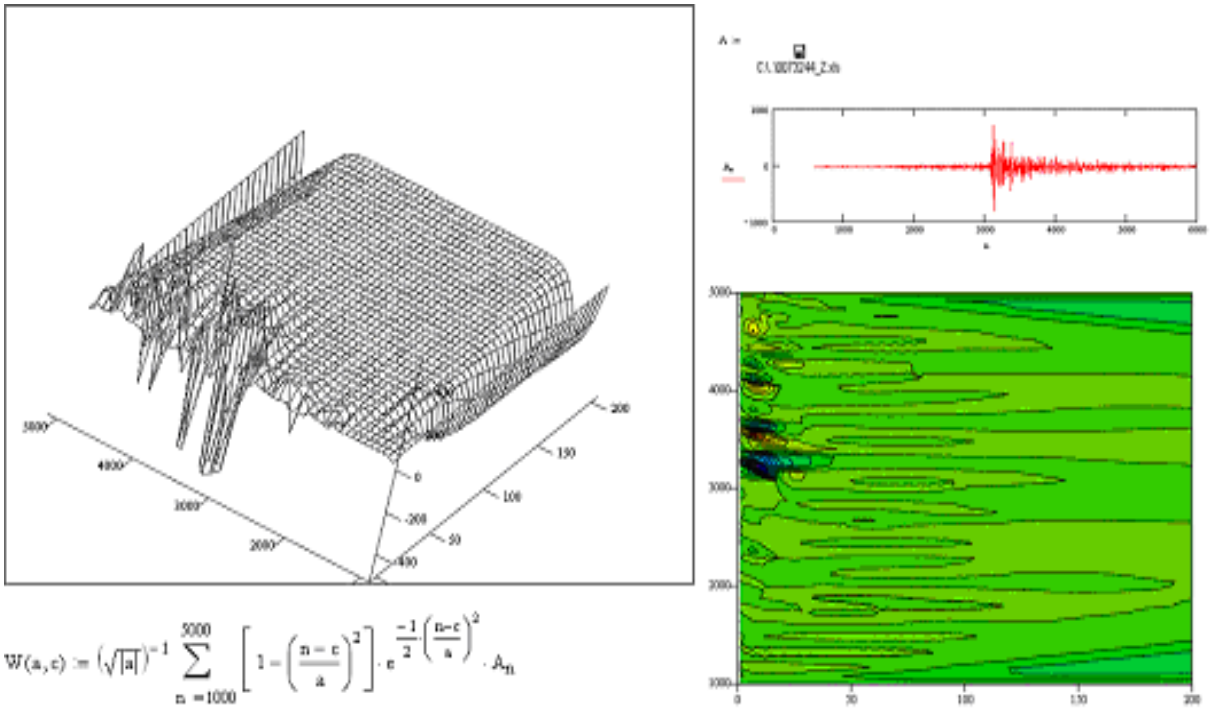

Рис. 6. Спектры вейвлет-преобразования сейсмического сигнала (Z–компонента, сейсмостанция Шира, землетрясение 12.04.2000 г.)

Таким образом, вейвлет-преобразование является более тонким инструментом для исследования непериодических сигналов, чем Фурье-анализ, что не уменьшает значимость последнего. Получение вейвлет-диаграмм для непрерывной функции не представляет проблемы. Задача заключается в подборе анализирующего вейвлета и масштаба, вплоть до которого будет исследоваться сигнал.

3.2. Быстрое вейвлет-преобразование сейсмических сигналов

Программная реализация алгоритма метода «быстрого» получения вейвлет-диаграммы сейсмического сигнала является элементом комплекса программ для обработки сейсмической информации. Запишем интегральное вейвлет-преобразование в виде:

$$\tilde{W}_{a,b} = \tilde{W}_{\phi} f(a,b) = |a|^{-0.5} \int_{-\infty}^{\infty} f(x)\phi\left(\frac{x-a}{b}\right)dx = |a|^{-0.5} \int_{-\infty}^{\infty} f(x)\phi_{a,b}(x)dx, \quad (19)$$

где a – масштабный коэффициент, b – параметр сдвига.

Таким образом, если хотим получить вейвлет-диаграмму функции $f(x)$ на интервале $[x_0, x_1]$, то параметр сдвига b изменяется именно в этом диапазоне. Границы области значений коэффициента масштабирования выбираются в соответствии с требованиями каждой конкретной задачи.

Вычисление на каждом шаге значений функций $f(x)$, $\varphi(x)$ и интегрирование по всей числовой оси делают эту процедуру крайне нерациональной для программно-алгоритмического использования в задачах с реальными данными. Кроме того, в каждой конкретной ситуации требуется своя точность, которая определяется качеством начальных данных или требованиями к качеству и скорости расчетов.

Дискретная цифровая запись какого-либо процесса характеризуется частотой дискретизации, точностью и величиной разрядной сетки сканирующего устройства, и нет смысла решать задачу существенно точнее, чем это диктуется неопределенностью начальных данных. Поэтому рационально использовать приближенные методы вычисления различной степени точности, ориентированные на быстрые вычисления.

Предлагается следующий алгоритм вейвлет-преобразования:

$$
\begin{aligned}
&j = 0, 1, 2, ..., N;\\
&i = 1, 2, ..., M;\\
&\Delta x = \frac{(x_1 - x_0)}{N} = const;\\
&b_j = j \cdot \Delta x \in [0, x_1 - x_0];\\
&a_i \in \{a_1, a_2, ..., a_M\},
\end{aligned}
$$

где N – число дискретных отсчетов на промежутке, M – величина выборки масштабов.

Введем функцию $L(a_i)$, характеризующую окно сходимости базисного вейвлета:

$$
\begin{aligned}
&L(a_i) \in \{1, 2, ..., \infty\};\\
&\left|\varphi\left(\frac{\pm j \cdot \Delta x}{a_i}\right)\right| < \delta, \ \forall\ j > L(a_i)\ ,
\end{aligned}
$$

где δ – порог сходимости, положительная постоянная величина «близкая» по значению к 0, для $j > L(a_i)$ будем считать $\varphi\left(\frac{\pm j \cdot \Delta x}{a_i}\right) \equiv 0$.

Затем введем понятие матрицы базисного вейвлета и матрицы $F_{l,i}$:

$$
\begin{aligned}
&l = 0,1,...,2L(a_M);\\
&\varphi_{i,l} = \varphi\left(\frac{(l - L(a_M)) \cdot \Delta x}{a_i}\right);\\
&F_{l,j} = f\left(x_0 + (j - l + L(a_M)) \cdot \Delta x\right).
\end{aligned}
$$

Теперь вейвлет-преобразование можно переписать в виде интегральной суммы:

$$
W_{a_i, b_j} = W_\varphi f(i, j) = \frac{\Delta x}{2\sqrt{a_i}} \sum_{n=L(a_M)-L(a_i)+1}^{L(a_M)+L(a_i)} \left(F_{n-1,j}\varphi_{i,n-1} + F_{n,j}\varphi_{i,n}\right). \qquad (20)
$$

Так как компоненты обеих матриц и функции $L(a_i)$ постоянные величины, то процедура вейвлет-преобразования сводится к перемножению и сложению некоторого набора констант. При этом матрицы организованы таким образом, чтобы количество операций на вычисление коэффициентов их элементов было минимально. Это существенно уменьшает затраты машинного времени необходимого для вычислений. Если интеграл приближать не методом трапеций, а простой интегральной суммой:

$$W_{a_i,b_j} = W_\varphi f(i,j) = \frac{\Delta x}{\sqrt{a_i}} \sum_{n=L(a_M)-L(a_i)}^{L(a_M)+L(a_i)} F_{n,j} \cdot \varphi_{i,n}, \qquad (21)$$

то скорость вычислений увеличивается в 2 раза, но и погрешность вычислений при этом также существенно увеличивается.

Введем величину, которая характеризует собственную ошибку метода:

$$\Delta\varphi_L(a_i) = \frac{\Delta x}{2} \sum_{n=L(a_M)-L(a_i)+1}^{L(a_M)+L(a_i)} (\varphi_{i,n-1} + \varphi_{i,n}). \qquad (22)$$

По определению материнского вейвлета, каждое значение этой функции должно быть равно нулю. Этого можно добиться, распределив полученные неточности по отрицательным значениям $\varphi_{i,n}$. Так как $L(a_i)$ изначально выбиралась так, чтобы минимизировать ошибку, то изменения не будут превосходить δ.

Теперь абсолютная погрешность будет состоять из безусловной погрешности начальных данных и ошибки метода приближения интегралов.

Абсолютная погрешность:

$$\Delta W_{a_i,b_j} = \left| W_{a_i,b_j} - \tilde{W}_{a_i,b_j} \right|.$$

Относительная погрешность:

$$\delta W_{a_i,b_j} = \frac{\Delta W_{a_i,b_j}}{\left|W_{a_i,b_j}\right|}.$$

Границы абсолютной и относительной погрешностей соответственно:

$$\Delta_W = \max_{i,j}\left|W_{a_i,b_j} - \tilde{W}_{a_i,b_j}\right|;$$

$$\delta_W = \max_{i,j}\left(\frac{\Delta W_{a_i,b_j}}{\left|W_{a_i,b_j}\right|}\right).$$

Очевидно, что чем больше количество дискретных отсчетов N, тем меньше погрешность приближения интеграла. Более точно для каждого конкретного случая эту зависимость можно определить следующим образом.

При получении вейвлет-диаграммы нас интересуют составляющие $f(x)$, лежащие в определенном диапазоне частот, т. е. колебания определенного набора периодов T_i. Очевидно, что ошибка вычисления тем меньше, чем больше дискретных отсчетов попадает в период, но при этом количество операций пропорционально возрастает. Обозначив через T_0 минимальный период из этого диапазона, получим:

$$\delta_W\left(\frac{T_i}{\Delta x}\right) \le \delta_W\left(\frac{T_0}{\Delta x}\right).$$

Можно поставить обратную задачу. Пусть минимальный период колебаний интересующий при обработке T_0, базисный вейвлет $\varphi(x)$ и приемлемая погрешность δ_W. Требуется найти оптимальный шаг дискретизации Δx. Если диапазон доминирующих частот базисного вейвлета не велик, то $\varphi(x)$ вовсе можно исключить из постановки задачи.

Контрольные вопросы и задания

1. Перечислите значимые компьютерные методы обработки геофизических данных.
2. Что такое вейвлет-анализ сигналов?
3. Что называют непрерывным вейвлет-преобразованием данных?
4. Дайте определение дискретного вейвлет-преобразования сигналов.
5. В чем заключается вычислительная процедура дискретного вейвлет-преобразования?
6. В каких случаях используют непрерывное вейвлет-преобразование данных?
7. Что такое скейлограмма?
8. Как происходит интерпретация вейвлет-спектра сейсмического сигнала?

Литература: раздел 3, основная литература п. 2, 7, 9, 12–13; дополнительная литература п. 2–8, 17.

4. ЭТАПЫ ПОЛУЧЕНИЯ ВЕЙВЛЕТ-СПЕКТРА ИССЛЕДУЕМОГО ПРОЦЕССА

Выделим следующие этапы получения вейвлет-спектра исследуемого процесса:

1. Получение начальных данных:

- задание функции $f(x)$ или набора ее значений с постоянной частотой дискретизации $f(x_i)$;
- преобразование исходных данных в формат, пригодный для обработки.

2. Постановка задачи:

- выбор материнского вейвлета;
- задание границ (параметров) вейвлет-преобразования – интервал масштабов и границы обрабатываемого отрезка функции;
- допустимая погрешность вычислений (если функция определена формулой).

3. Проверка корректности условий задачи:

- проверка начальных данных;
- проверка параметров преобразования;
- расчет абсолютной погрешности (если данные задаются набором значений).

4. Предварительные вычисления:

- построение матрицы базисных вейвлетов – $\varphi_{i,j}$;
- переорганизация начальных данных – $F_{i,j}$.

5. Вейвлет-преобразование.

6. Визуализация результата расчетов.

7. Сохранение результата расчетов.

Примеры графической визуализации результатов расчетов приведены на рис. 7.

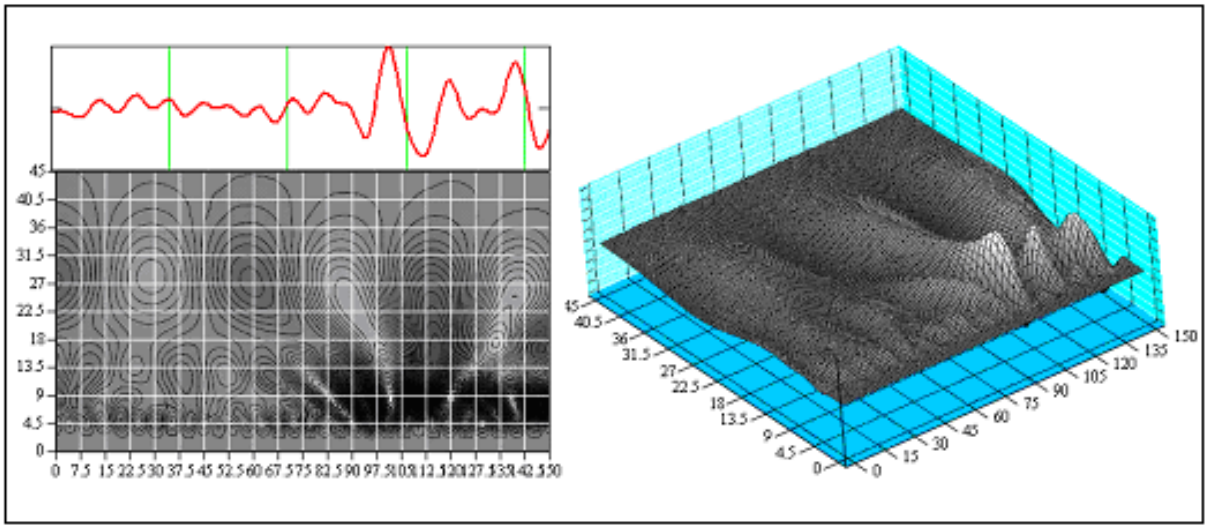

Рис. 7. Примеры графической визуализации вейвлет-диаграммы

Следующим этапом является обработка и интерпретация вейвлет-спектра в зависимости от конкретной задачи и реального геофизического объекта. Более подробное описание данного этапа приведено в п. 4.1.

4.1. Интерпретация вейвлет-спектра сейсмического события, произошедшего в районе г. Дивногорска

На рис. 8, *а–в* поэтапно показана вейвлет-обработка реального сейсмического события.

На рис. 8, *а* показана сейсмограмма сейсмического события, произошедшего в районе г. Дивногорска 7.04.2002 г., записанная сейсмостанцией «Дивногорск» и обработанная оптимальным для данного события полосовым фильтром. Энергетический класс землетрясения 6.2. Эпицентральное расстояние 25 км. Для регистрации сейсмособытия использовалась цифровая регистрирующая станция (ЦС) «Дельта-Геон» и трехкомпонентные сейсмографы, частота дискретизации записи 35.714 Гц.

Отметим, что энергетический класс события и эпицентральное расстояние по меркам сейсмического мониторинга довольно малы. На таких расстояниях *S*-волны только начинают проявляться, и их очень

сложно выделить на сейсмограмме. Если на вейвлет-диаграмме обнулить все точки, кроме экстремальных, и усилить коррелирующие по всем трем компонентам записи, то получаемая картина довольно полно характеризует все событие, и в большинстве случаев позволяет выделить основные фазы.

Основной упор был сделан на определение времени прихода продольных колебаний генерируемых источником (*P*-волна) и поперечных (*S*-волна). Зная разницу во времени этих двух вступлений, по годографу, определяется расстояние от станции до эпицентра события. Чистый сигнал (без примеси сейсмических шумов) должен коррелировать в их окрестностях по всем трем ортогональным компонентам записи. Если использовать вместо компонент сигнала их вейвлет-диаграммы, то проверка на корреляцию проводится не только по времени, но и по частотам.

На рис. 8, *а* хорошо видно вступление *P*-фазы, но довольно нечетко отслеживается момент прихода *S*-волны. В качестве анализирующего вейвлета использовалась восьмая производная функции Гаусса. Исследовано 14 секунд записи для значений масштабного коэффициента от 5 до 60. Получены и затем специальным образом обработаны вейвлет-спектры всех трех компонент записи.

На рис. 8, *б* показана вейвлет-диаграмма *N*-канала записи. Результатом обработки можно считать диаграмму, на которой четко видны не только вступления обеих фаз, но и их частотный состав и его вариации (рис. 8, *в*). Красным выделена полученная таким образом *P*-волна, синим *S*-волна. Кроме того, выделились поверхностные волны, сведения о которых планируется в дальнейшем использовать для классификации событий («взрыв – не взрыв»).

Полное машинное время, затраченное на обработку данного сейсмического события, составило 2 минуты. Также разработаны алгоритмы для выделения и анализа большинства информативных составляющих сейсмического сигнала.

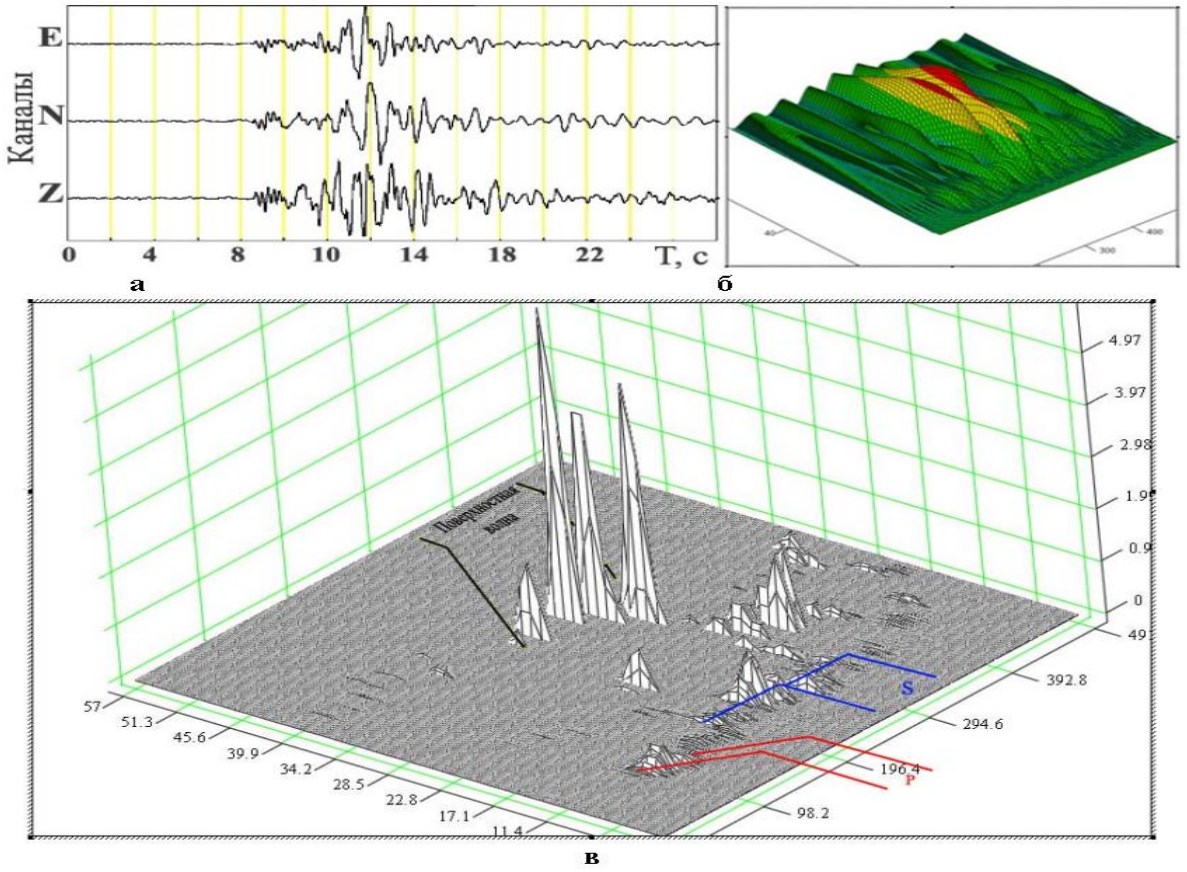

Рис. 8. Обработка сейсмособытия 07.04.2002: сейсмограмма сейсмических колебаний (*а*), вейвлет-диаграмма *N* канал (*б*), обработанная специальным образом вейвлет-спектральная информация о событии (*в*)

Полное машинное время, затраченное на обработку данного сейсмического события, составило 2 минуты. Также разработаны алгоритмы для выделения и анализа большинства информативных составляющих сейсмического сигнала.

На рис. 9 показан результат действия такой процедуры. По частотному составу и динамике его изменения для каждого семейства волн можно исследовать природу и проводить классификацию источника колебаний.

Рис. 9 иллюстрирует разложение сейсмического события на составляющие его семейства волн: объемные (*P* и *S*-волны), поверхностные (Лява и Рэлея). Поверхностные сейсмические волны – это сейсмические волны, распространяющиеся только вдоль поверхности Земли. Скорость поверхностных волн меньше скорости поперечных волн.

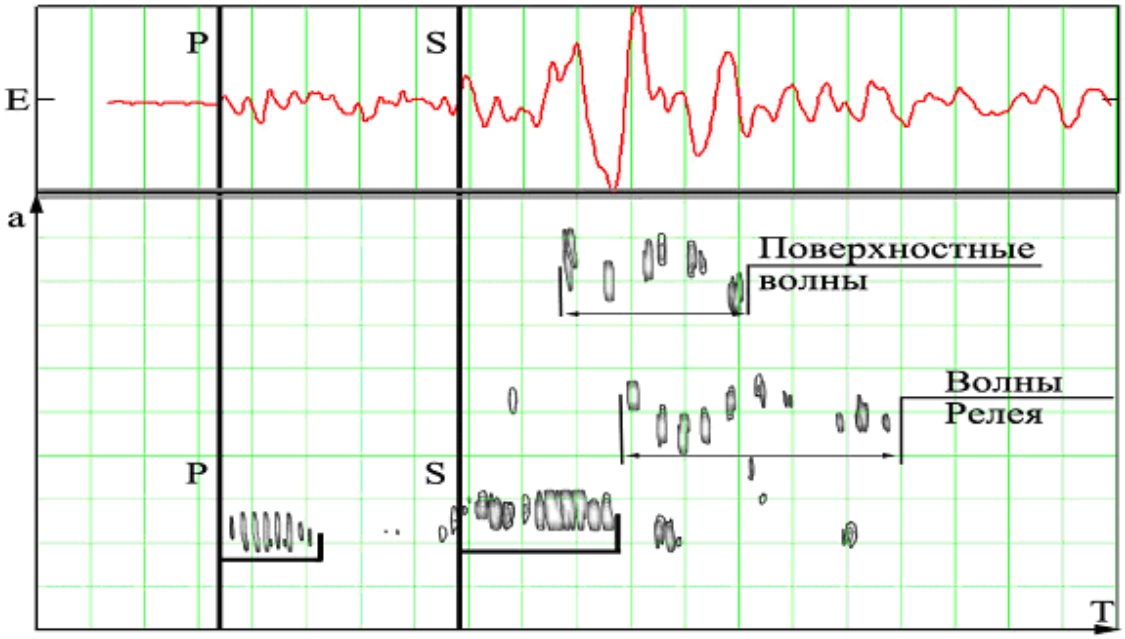

Рис. 9. Разложение сейсмического события на составляющие его семейства волн

В настоящее время выполнены исследования, направленные на создание системы распознавания «промышленный взрыв – землетрясение», применительно к системе сейсмического мониторинга юга Красноярского края. По средствам специальных преобразований вейвлет-диаграмм проводится сравнительный спектральный анализ участков сейсмограмм, содержащих вступления *P*-волн.

На рис. 10, *а–г* приведены сейсмограммы вступления *P*-волн землетрясения и промышленного взрыва, а также их вейвлет-спектрограммы. А именно, на рис. 10, *а* показана сейсмограмма вступления *P*-волн землетрясения.

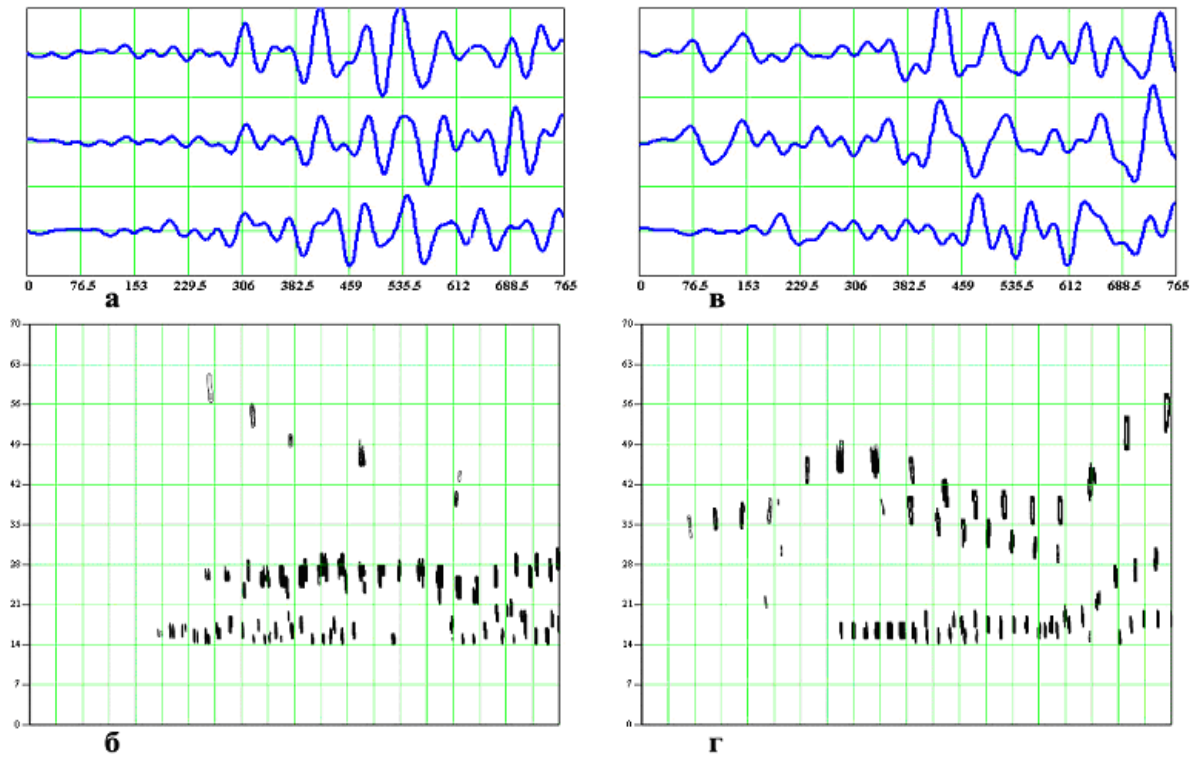

Рис. 10. Сейсмограммы вступления Р-волн землетрясения (*а*)
и промышленного взрыва (*б*), их вейвлет-спектрограммы (*в*, *г*)

Вейвлет-спектрограммы вступления *Р*-волн землетрясения и промышленного взрыва представлены на рис. 10, *б* и рис. 10, *г* соответственно. На рис. 10, *в* – сейсмограмма промышленного взрыва. Существенные различия между спектрами промышленного взрыва и землетрясения хорошо видны на рис. 10, *г*.

Надо заметить, что благодаря использованию вейвлет-преобразования, а именно его свойству пространственно-временной локализации, стало возможным исследование развития частотного состава со временем. Основываясь на этом, был предложен критерий, который в совокупности с рядом признаков показал высокую эффективность при решении задач распознавания.

Описанные выше методы и алгоритмы в комплексе с поляризационным анализом и пространственным вейвлет-анализом используются в качестве инструментов в системе автоматического сбора и обработки сейсмической информации.

Создание и развитие в настоящее время системы сейсмического мониторинга юга Красноярского края на основе цифровых станций является важным этапом в повышении эффективности геофизических

исследований. Это необходимый этап для создания службы прогноза землетрясений в Красноярской промышленной агломерации. Основными требованиями для такой системы являются высокая надежность и оперативность обработки и, как следствие, высокий уровень автоматизации на всех этапах – от получения данных на станциях до окончательной обработки их в центре.

Результаты обработки с помощью программных средств разработанных алгоритмов реального сейсмического сигнала, полученного экспериментальной системой геомониторинга юга Красноярского края, приведены на рис. 11, *а* и 11, *б*.

На рис. 11, *а* приведены сейсмические записи промышленного взрыва (1–3). На рис. 11, *б* показана визуализация расчетов координат эпицентра землетрясения и местоположения промышленного взрыва.

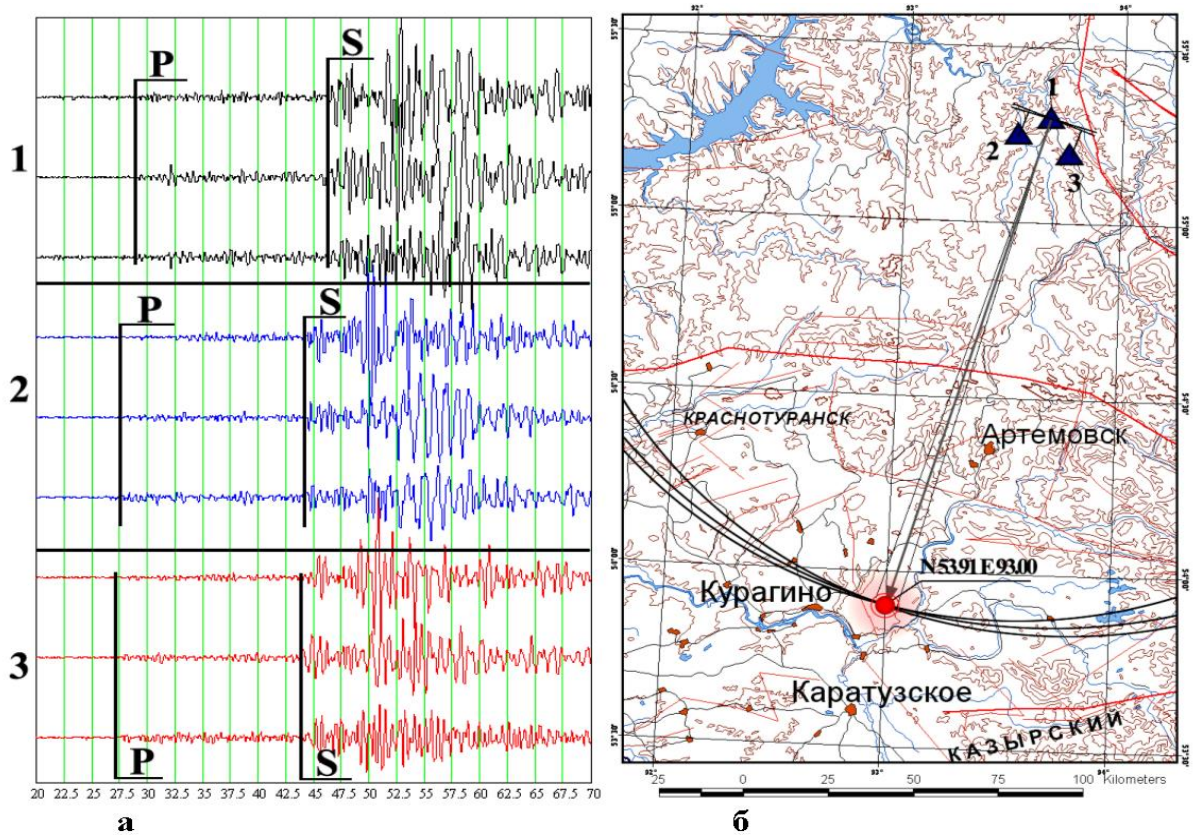

Рис. 11. Обработка по средствам разработанных алгоритмов реального сейсмического сигнала, полученного экспериментальной системой геомониторинга: сейсмические записи промышленного взрыва (*а*); визуализация результата расчетов координат эпицентра землетрясения и местоположения промышленного взрыва (*б*)

Контрольные вопросы и задания

1. Проанализируйте этапы получения вейвлет-спектра исследуемого геофизического процесса.

2. Что понимается под вейвлет-диаграммой?

3. Дайте определения поверхностных и объемных волн.

4. Какие условия необходимо задать на вейвлет-диаграмме, чтобы выделить основные фазы сейсмического события?

5. Что характеризует годограф скоростей?

6. Какие параметры используются для классификации источника колебаний?

7. Сформулируйте основное свойство вейвлет-преобразования.

8. В чем заключается сущность метода поляризационного анализа.

Литература: раздел 4, основная литература п. 2, 7, 9, 12–13; дополнительная литература п. 2–8, 17.

5. АЛГОРИТМЫ ВЫДЕЛЕНИЯ И ОБРАБОТКИ СЛАБЫХ СЕЙСМИЧЕСКИХ СИГНАЛОВ НА БАЗЕ *ЛСГ*

Представлены технология и программные алгоритмы для создания автоматической системы выделения и обработки слабых сейсмических сигналов на базе *локальной сейсмической группы* (*ЛСГ*). В ходе исследований были разработаны и проверены на реальных записях сейсмических событий приемы обработки и методы решения, как отдельных вопросов, так и задачи в целом.

В предложенных алгоритмах предполагается, что анализ данных проводится по идентичным трехкомпонентным каналам станций *ЛСГ*. Для того чтобы совокупность сейсмических станций и каналов передачи данных считалась *ЛСГ*, должны выполняться следующие условия:

- станции наблюдения расположены на одной плоскости, т. е. не учитывается кривизна поверхности Земли и разность высот пунктов;
- фронт сейсмической волны плоский, т. е. эпицентр достаточно удален от системы наблюдений;
- среда, в которой распространяются сейсмические волны, однородна;
- амплитудно-частотные и фазовые характеристики каналов на всех пунктах идентичны.

В настоящее время за рубежом опубликован ряд работ, в которых рассматриваются задачи обнаружения и обработки сейсмических процессов, на основании *ЛСГ*. Отличительной особенностью предложенных алгоритмов является их ориентированность на обработку именно слабых сигналов, регистрируемых в заданном регионе. Немаловажным фактом является возможность использования и комбинирования методов в зависимости от каждой конкретной ситуации.

5.1. Применение методов фильтрации для выделения и обработки сейсмического сигнала

Для выделения и последующей обработки сигнала используются два вида фильтрации: частотная и поляризационная, а также Фурье и вейвлет-анализ.

На основе наблюдений, проведенных при обработке записей региональной сейсмической сети, выяснилось, что интересующая нас информация содержится в полосе частот от 0.5 до 3 Гц. Полосовая частотная фильтрация применяется для подавления микросейсмических шумов на частотах, находящихся вне диапазона частот сигналов. Затем проводится вейвлет-анализ спектрального состава отфильтрованной полосы, с целью выделения из нее наиболее информативных частот, ценных для обработки.

Из отфильтрованного сигнала выбирается участок записи, предположительно содержащий вступления фаз события, который разлагается в вейвлет-спектр. В качестве материнского вейвлета используется

$$\varphi(t) = (105 - 420t^2 + 210t^4 - 28t^6 + t^8) \cdot e^{\frac{-t^2}{2}}.$$

Сейсмическая запись дискретна и, следовательно, используется также дискретное вейвлет-преобразование. Заменим интеграл суммой, это не приводит к большой погрешности, чтобы это показать, возьмем функцию в виде:

$$f(t) = \sin\frac{t}{\pi} + \cos 2.5\frac{t}{\pi},$$

и получим ее вейвлет-диаграмму двумя способами. Максимальная ошибка 0.37 % (граничный эффект), при этом количество отсчетов на период колебания в несколько раз меньше, чем в реально обрабатываемых сигналах.

Необходимо отметить также, что на данной стадии разработок допустимая погрешность вычисления 5 %. Намного большее значение имеет то, что скорость вычисления, при помощи такой замены, можно увеличить до 8 раз, не выйдя за указанный предел ошибки. На получаемой вейвлет-диаграмме выделяются области скопления точек экстремумов (предположительно вступления фаз события), а точнее наборы этих точек. Значения $W_{\varphi} f(a,b)$ в них должны превышать определенный порог и отвечать еще ряду условий.

В каждой группе методом наименьших квадратов усредняется значение компоненты характеризующей частотную принадлежность элементов. Набор частот, получаемый таким образом, для каждого события уникален и впоследствии с учетом статистики может указать на природу очага и особенности геолого-тектонического строения среды.

5.2. Обработка геофизических данных сейсмического события

1. Выделение фаз сейсмического события. Если на вейвлет-диаграмме обнулить все точки кроме экстремальных, и усилить коррелирующие по всем трем компонентам записи, то получаемая картина довольно полно характеризует все событие, и в большинстве случаев позволяет выделить основные фазы.

Основное внимание уделялось на определение времени прихода продольных колебаний генерируемых источником (*P*-волна) и поперечных (*S*-волна). Зная разницу во времени этих двух вступлений, по годографу, определяется расстояние от станции до эпицентра события. Под годографом сейсмических волн понимается график зависимости времени пробега волны от источника до приемника волны (регистрирующего устройства) от эпицентрального расстояния.

Чистый сигнал (без примеси сейсмических шумов) должен коррелировать в их окрестностях по всем трем ортогональным компонентам записи. Если использовать вместо компонент сигнала их вейвлет-диаграммы, то проверка на корреляцию проводится не только по времени, но и по частотам.

Результаты вейвлет-обработки реального сейсмического события показаны на рис. 9. Были выделены объемные волны: *P* и *S*-волны. Кроме того, выделились поверхностные волны, сведения о которых важно использовать для классификации событий (промышленный взрыв – землетрясение).

2. Определение азимута на эпицентр сейсмического события. Если качество записи позволяет точно определить время вступления хотя бы одной из фаз одновременно по трем или более станциям группы, то вычисление азимута сводится к решению несложной геометрической задачи. В противном случае приходится прибегать к более сложным методам.

Пространственно-временная запись сейсмического волнового поля, зарегистрированного группой сейсмоприемников, позволяет провести анализ структуры волнового поля и оценить его параметры на основе частотно-волнового спектра мощности (рис. 12).

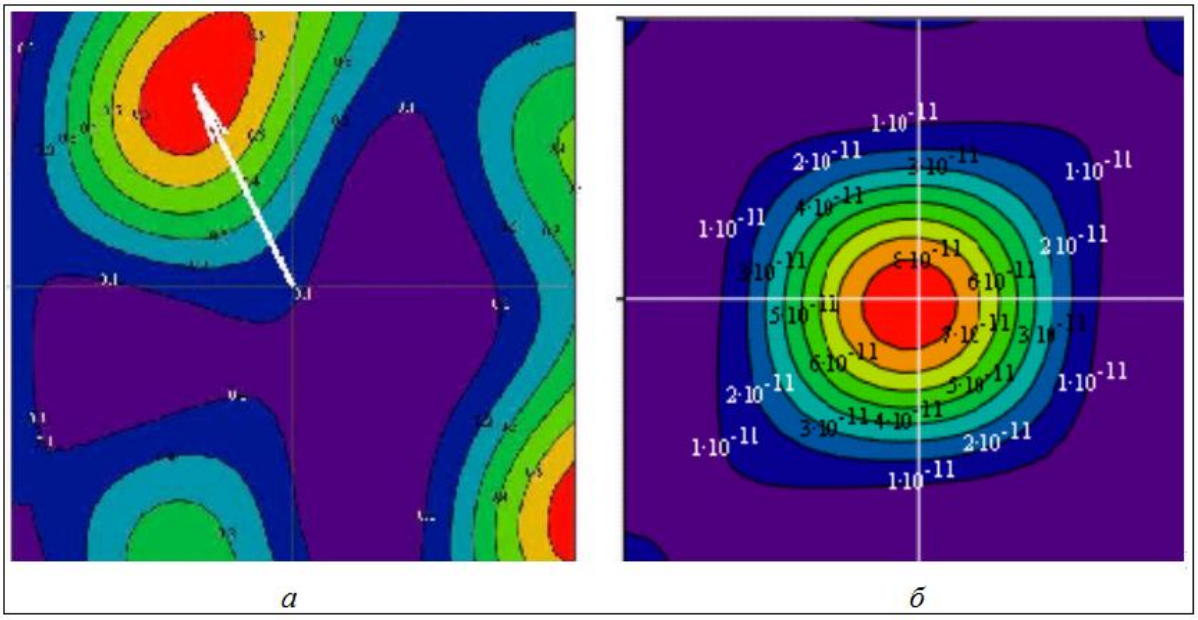

Рис. 12. Частотно-волновой спектр мощности сигнала (*а*) и шума (*б*)

Если в сейсмическом волновом поле присутствует регулярная компонента, то максимумы спектров в пространстве волновых чисел дают представление о пространственном положении сейсмического источника:

$$P(\bar{k}) = \frac{1}{m}\sum_{j=1}^{m} F(\bar{k}, f_j), \qquad F(\bar{k}, f_j) = \sum_{j=1}^{n-1}\sum_{l=1}^{n} F_{jl} e^{i\frac{2\pi f \bar{n}}{c}(\bar{r_j} - \bar{r_l})},$$

где f – набор частот, полученных на предыдущем этапе; m – размер частотной выборки; F_{jl} – взаимный спектр Фурье сигналов в точках j и l. Параметры f_0 и $f_1 = f_0 + m\Delta f$ – соответственно, верхняя и нижняя границы полосы пропускания; m – число дискретных отсчетов в выбранной полосе; F_{jl} – взаимный спектр сигналов в точках j и l.

Зададим волновое число в виде $\bar{k} = (k_x, k_y)$. Пусть $f_0 = f_1 = f$ и $w = 2\pi f$, тогда

$$P(kx, ky) = \sum_{j=1}^{n-1}\sum_{l=1}^{n} F_{jl} e^{i(kx(x_j - x_l) + ky(y_j - y_i) + w(\tau_j - \tau_l))}, \qquad (23)$$

здесь (x_i, y_i) – координаты i-й станции.

Время прихода волны на i-ю станцию определяется по формуле:

$$t_i = t_0 + \frac{\cos\varphi}{c}(x_i - x_0) + \frac{\sin\varphi}{c}(x_i - x_0) + \tau_i, \qquad (24)$$

где t_0 время прихода в данную точку (x_0, y_0) и τ_i – остаточный член (станционная поправка), c – скорость волны, φ – азимутальный угол, а вектор:

$$\bar{k} = (k_x, k_y) = \left(\frac{w \cdot \cos\varphi}{c}, \frac{w \cdot \sin\varphi}{c}\right). \qquad (25)$$

3. Пространственный волновой спектр мощности дает нам прямой доступ к физической структуре волнового поля, поскольку он указывает на величину мощности, относящейся к волнам с различной скоростью или к волнам, подходящим с разных направлений.

Вычисления проводятся в два этапа. Во время первой итерации, количество узлов сетки (30x30), и расчет ведется по волновым числам,

модуль компонентов k_x и k_y не превышает 1,5. На втором этапе вокруг найденного значения выставляется область в 15 раз меньше начальной и берется сетка в 16 раз мельче.

После расчета по формулам 21–23 получаем пространственный спектр мощности волнового поля. По смешению максимума спектра от нуля определяется азимут на эпицентр. Зная эпицентральное расстояние, вычисляем координаты события. На рис. 13 показана физическая структура волнового поля. Здесь на рис. 13, *а* представлен спектр мощности сейсмического поля в пространстве волновых чисел, на рис. 13, *б* – метод определения координат эпицентра по азимуту и эпицентральному расстоянию.

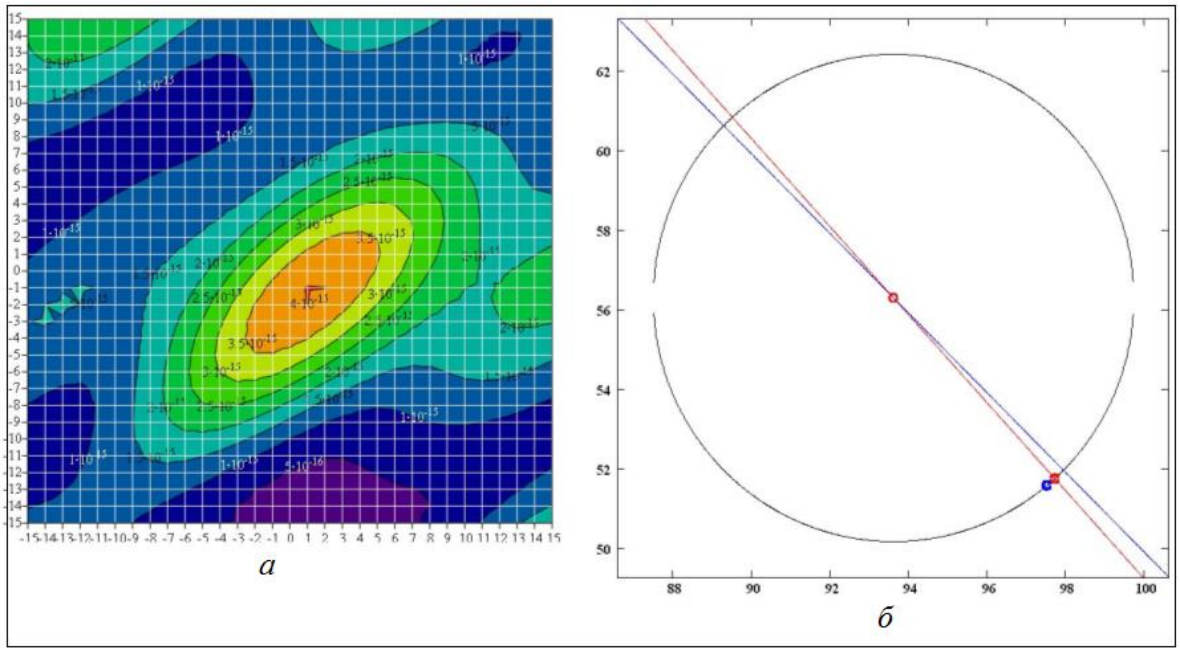

Рис. 13. Физическая структура волнового поля: спектр мощности сейсмического поля в пространстве волновых чисел (*а*); определение координат эпицентра по азимуту и эпицентральному расстоянию (*б*)

4. Поляризационный фильтр. Зная азимут прихода волны к сейсмостанции, можем перейти к поляризационному фильтру. Продольные и поперечные волны характеризуются высокой степенью поляризации. При распространении продольной волны смещение частиц среды происходит в направлении распространения волны, при распространении поперечной волны смещение частиц среды происходит в направлении,

перпендикулярном распространению волны. У продольных волн движение частиц среды совпадает с азимутом распространения волны.

Поверхностные волны релеевского типа обычно поляризованы эллиптически в вертикально-радиальной плоскости. Поскольку фон микросейсм образован преимущественно волнами релеевского типа, его поляризация должна быть эллиптической, но без выраженной направленности.

Помехи, порождаемые сигналом, могут быть также поляризованными, но направление поляризации по природе будет случайной. На основании указанных характеристик траекторий поляризации частиц среды, можно спроектировать такие фильтры, которые сохраняют движение, удовлетворяющее условиям поляризации в определенном направлении, и ослабляют волновое движение, если оно не удовлетворяет заданным критериям.

Частный случай поляризационной фильтрации, принимающий во внимание только прямолинейность в радиальном и вертикальном направлениях, а не в трехмерном пространстве, называется детектором прямолинейного движения, или *ДПД*-фильтром.

Принцип действия *ДПД*-фильтра заключается в повороте вертикальной Z и радиальной R составляющих сейсмического сигнала до такого положения, когда ожидаемое направление распространения анализируемой волны будет делить угол между двумя ортогональными компонентами. В этом случае имеем:

$$\begin{aligned}\vec{Z} &= Z\cos(\frac{\pi}{4}-\theta)+N\cos(\alpha)\sin(\frac{\pi}{4}-\theta)+E\sin(\alpha)\sin(\frac{\pi}{4}-\theta);\\ \vec{R} &= Z\sin(\frac{\pi}{4}-\theta)-N\cos(\alpha)\cos(\frac{\pi}{4}-\theta)-E\sin(\alpha)\sin(\frac{\pi}{4}-\theta)\end{aligned},$$

где θ – угол подхода, азимут по большому кругу Z, N, E – вертикальная, северная и восточная составляющая вектора движения.

Оператор фильтра получается из функции взаимной корреляции между $\vec{Z}(t)$ и $\vec{R}(t)$, вычисляемой по отрезку записи W с центром на времени t:

$$C(+T) = \sum_{t-W/2}^{t+W/2} \vec{Z}(t)\vec{R}(t+T).$$

Чтобы оператор был четным и не вносил фазовых искажений, отрицательные сдвиги образованы из положительных значений:

$$C(-T) = C(+T).$$

Если поляризация в плоскости *R-Z* преимущественно прямолинейная, функция *C(T)* принимает большие значения. Если движение эллиптично или случайно, значения *C(T)* будут малыми.

В результате свертки функции *C(T)* с первоначальным временным рядом те движения, которые характеризуются высокой степенью прямолинейности, будут усилены, а эллиптически поляризованные движения – ослаблены. Сигнал на выходе такого фильтра имеет вид:

$$Y_R(t) = P_P K \sum_{T=-L}^{L} \vec{R}(t-T)C(T);$$

$$Y_Z(t) = P_P K \sum_{T=-L}^{L} \vec{Z}(t-T)C(T),$$

где P_P – оператор поляризации; K – нормирующий множитель.

Оператор $C(T)$ для каждого момента времени t должен быть свой, поэтому на обработку уходит много машинного времени. Рекомендуется выбирать максимальный сдвиг T_{max} равный половине интервала анализа, т. е. $\frac{L}{2}$. Нормирующий множитель K определяется по функции автокорреляции входного сигнала; его можно использовать для усиления слабых сигналов с высокой степенью прямолинейности:

$$K(t) = \left[\sum_{t-W/2}^{t+W/2} \vec{R}^2(t) \sum_{t-W/2}^{t+W/2} \vec{Z}^2(t) \right]^{\frac{1}{2}}.$$

Отметим, что в случае продольных волн имеем:

$$\vec{R}(t)\vec{Z}(t) > 0,$$

а для вертикально поляризованных поперечных волн:

$$\vec{R}(t)\vec{Z}(t) < 0.$$

Оператор поляризации имеет вид:

$$P_p = \begin{cases} 1, \vec{R}(t)\vec{Z}(t) > 0 \\ 0, \vec{R}(t)\vec{Z}(t) < 0 \end{cases},$$

в случае продольной волн P и

$$P_s = \begin{cases} 1, \vec{R}(t)\vec{Z}(t) < 0 \\ 0, \vec{R}(t)\vec{Z}(t) > 0 \end{cases},$$

в случае вертикально поляризованной поперечной волны *SV*.

Подобные фильтры являются очень мощным и удобным инструментом для выделения из сейсмического сигнала отдельных составляющих, и в совокупности с вейвлет-анализом позволяют автоматизировать обработку слабых сейсмических записей, отношение сигнал/помеха в которых меньше единицы.

Рассмотренный фильтр подобен описанному ранее методу выделения фаз за счет корреляции компонент сигнала, но в данном случае для каждого масштаба вейвлет-диаграммы применяется *ДПД*.

Скорость действия можно существенно увеличить, если обрабатывать не весь сигнал, а лишь наиболее интересующие нас части. Кроме того, параметры вейвлет-преобразования также могут варьироваться в зависимости от того, какую фазу мы собираемся выделить. На рис. 14 приведены результаты выделения фаз с помощью фильтра, созданного на базе вейвлет-анализа и *ДПД*.

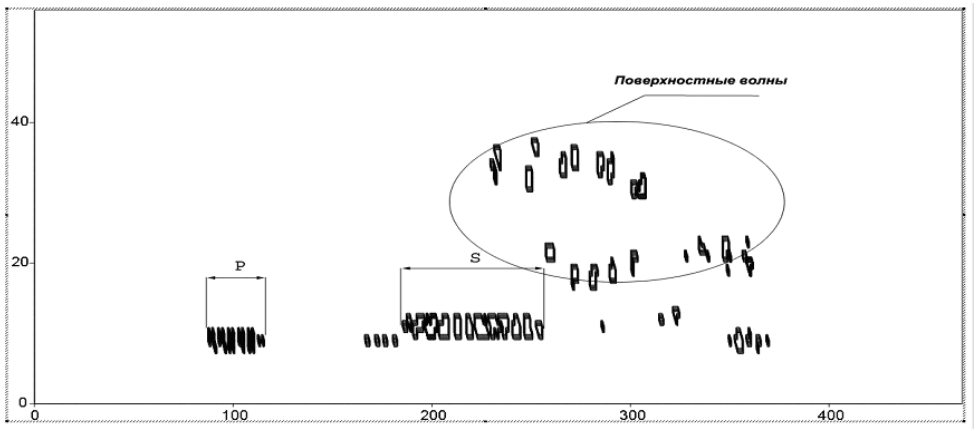

Рис. 14. Выделение фаз с помощью фильтра созданного
на базе вейвлет-анализа и *ДПД*

Времена вступлений фаз определяются по диаграмме. Исходя из них, по годографу, получаем расстояние от станции до события.

Таким образом, рассмотрен полный комплекс математических приемов для проведения обработки сейсмической информации. В зависимости от качества записи и предъявляемых требований к точности, выбираются комбинации методов, обеспечивающие оптимальные по качеству и времени получения результаты.

Испытания показали, что для получения качественных данных лучше всего изначально ориентировать *ЛСГ* на мониторинг определенного района (сектора).

Желательные параметры *ЛСГ*:

- четыре сейсмостанции (три основных и одна вспомогательная);
- апертура группы 8–10 км;
- расстояние до зоны наблюдения не менее 100 км.

Подобная группа в состоянии покрывать область в 200х200 км^2, а этого более чем достаточно, чтобы исследовать очаговые зоны землетрясений и прилегающие сейсмоактивные разломы.

Контрольные вопросы и задания

1. Что понимается под термином *ЛСГ*.
2. Каковы особенности мониторинга на основе локальной сейсмической группы?
3. Перечислите основные параметры *ЛСГ*.
4. Какие виды фильтрации используются для выделения и последующей обработки слабого сигнала?
5. Каким образом выделяются информативные частоты, ценные для обработки?
6. В чем отличие методов обработки для случая слабых сейсмических сигналов?
7. Когда используются алгоритмы частотной и поляризационной фильтрации сигналов?
8. Что называют детектором прямолинейного движения (*ДПД*-фильтр)

Литература: раздел 5, основная литература п. 9, 15; дополнительная литература п. 2–3, 6, 10–11, 14–15.

6. ИСХОДНЫЕ ДАННЫЕ СЕЙСМИЧЕСКИХ СИГНАЛОВ ЛСГ, СХЕМА ИХ ОБРАБОТКИ И МОДЕЛИРОВАНИЯ

6.1. Исходная информация для моделирования сигналов локальной сейсмической группы

В качестве исходных данных для моделирования сейсмических сигналов использовалась геофизическая информация о двух временных сейсмических группах в районах г. Красноярска и п. Агинское по данным ГС СО РАН г. Новосибирск.

Временная сейсмическая группа в районе г. Красноярска. Для отработки методов и выяснения особенностей алгоритмов использовались записи двух сильных землетрясений в Республике Тыва. Параметры сейсмических событий:

18.10.2001	***21.10.2001***
Время в очаге: ***20:14:47***	*Время в очаге:* ***20:14:47***
Координаты ***51.6*** *N* ***97.5*** *E*	*Координаты* ***51.43*** *N* ***90.7*** *E*
Энергетический класс K = ***9***	*Энергетический класс K=****10.7***

Для эксперимента были взяты данные шести станций, выставленных для сбора сейсмической информации на технической площадке, как показано на рис. 15.

Сейсмическая группа (рис. 15) отвечает не всем требованиям и тем самым не может считаться *ЛСГ*, но для решения задач геофизического эксперимента вполне приемлема. Расстояние между станциями группы не превышало несколько сот метров.

Выяснилось, что на таких малых апертурах невозможно воспользоваться ни одним из методов определения азимута. Волны просто не успевают достаточно разбежаться по времени.

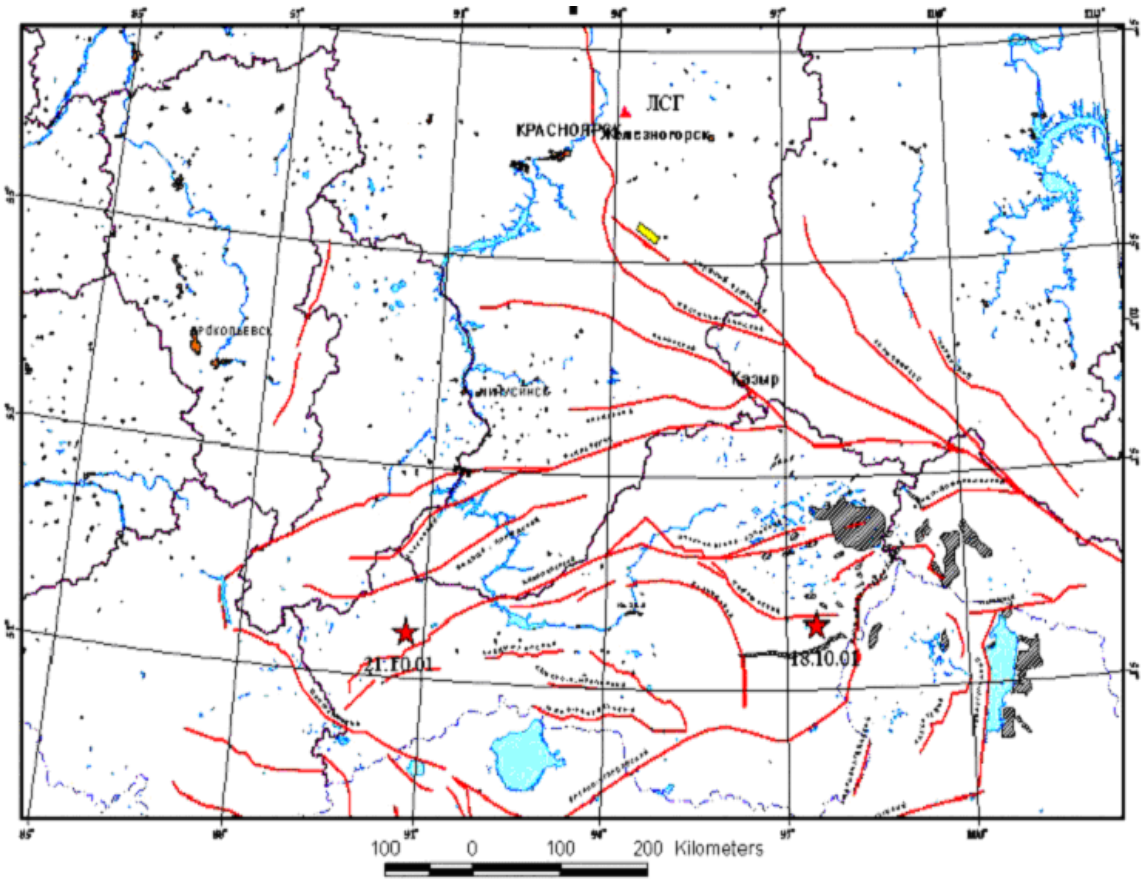

Рис. 15. Расположение временной *ЛСГ* и эпицентров землетрясений

Записи со станций группы довольно сильно различались по качеству, это позволило испытать возможности алгоритмов выделения фаз. Из-за большого эпицентрального расстояния и сильной зашумленности сигнала, вступления фаз, после применения обычного полосового фильтра, и даже *ДПД*, не выделялись достаточно четко для качественной обработки.

Фильтр, созданный на базе вейвлет-преобразования, позволил определить эти величины с точность до 0.3 секунды, при частоте дискретизации датчиков 35.714 отсчетов/секунду. Установка станций с более высокими характеристиками позволит повысить точность расчетов, но отрицательно скажется на скорости обработки.

Временная сейсмическая группа в районе п. Агинское. Исходя из результатов предыдущего эксперимента, была подобрана группа сейсмостанций большей апертуры в поселках: «Зеленино», «Карлык», «Малиновка».

Для обработки использовались две записи: землетрясение и промышленный взрыв:

10.07.2001
Время в очаге: ***06:13:50***
Координаты ***54.80*** *N* ***94.87*** *E*
Энергетический класс K = ***8.5***

25.06.2001
Время в очаге: ***07:37:03.5***
Координаты ***55.92*** *N* ***94.75*** *E*
Энергетический класс K = ***8.1***

Эпицентральные расстояния порядка 40 и 85 км до ближайших станций (рис. 16). Фронт сейсмических волн на таких расстояниях нельзя считать прямым, следовательно, методы определения азимута будут давать ошибку пропорционально кривизне фронта.

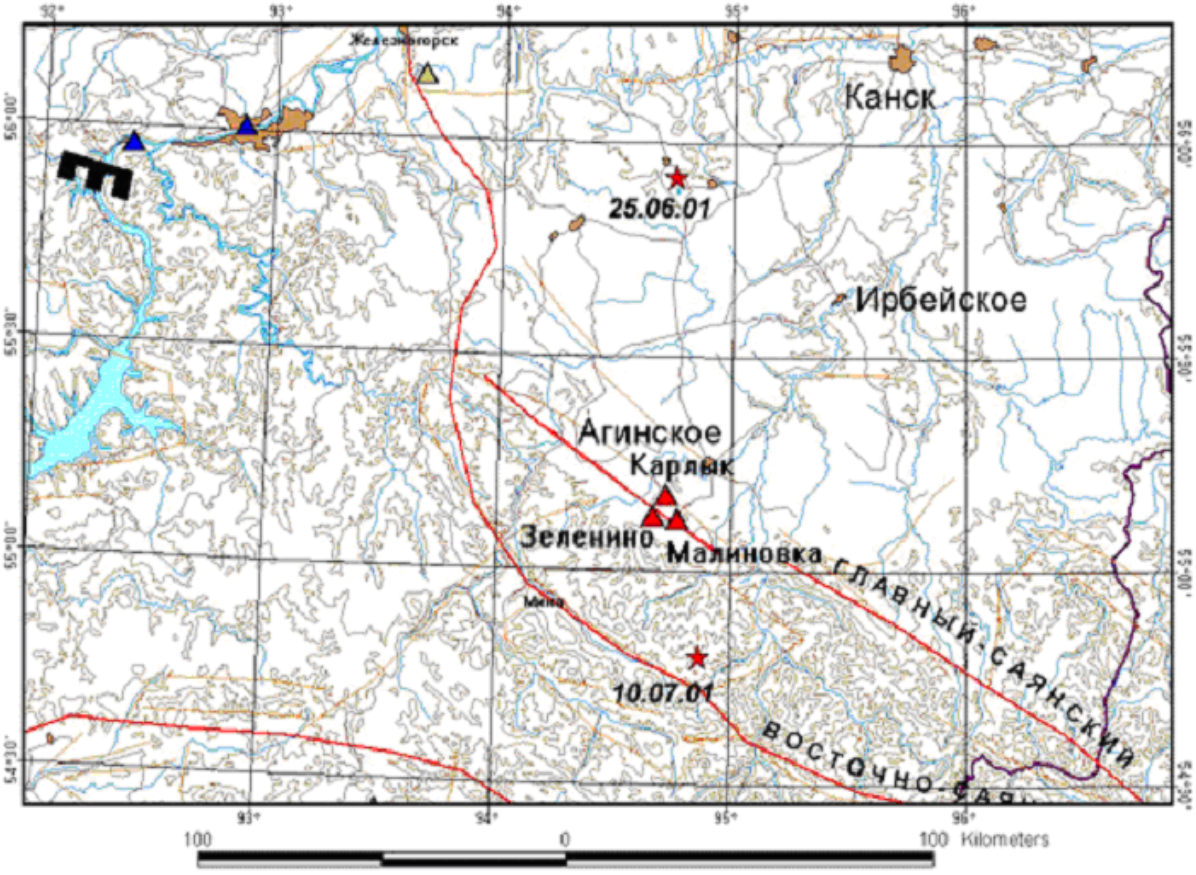

Рис. 16. Расположение временной *ЛСГ* и эпицентров сейсмических событий

В данном случае ошибки составили не больше 2–3 км. Под ошибкой имеется в виду различия в координатах эпицентра, полученных экспериментальным методом и стандартными средствами обработки, которые в свою очередь тоже не могут считаться эталонными.

Набрав достаточное количество статистических данных, можно добиться при помощи системы поправочных коэффициентов, чтобы погрешность составляла не более 3–4 % от эпицентрального расстояния.

6.2. Обработка и анализ сейсмических сигналов на основе ЛСГ

Обработка и анализ сейсмических сигналов по данным локальной сейсмической группы (*ЛСГ*) проводилась согласно этапам, приведенным на схеме рис. 17.

Рис. 17. Схема обработки и анализа сейсмических сигналов на основе *ЛСГ*

Контрольные вопросы и задания

1. Дайте определения землетрясения и очаговой области.
2. Какие локальные сейсмические группы использовались для мониторинга?
3. Что означает термин «эпицентральное расстояние»?
4. Какие виды фильтрации используются для выделения слабого сигнала?
5. Приведите оценки ошибки в случае применения метода определения азимута.
6. В чем отличие методов обработки для случая слабых сейсмических сигналов?
7. Что входит в комплекс математических приемов для проведения обработки сейсмической информации?
8. Опишите схему обработки и анализа сейсмических сигналов на основе ЛСГ.

Литература: раздел 6, основная литература п. 9, 15; дополнительная литература п. 2–3, 6, 10–11, 14–15.

7. МОДЕЛИРОВАНИЕ ДАННЫХ КАТАЛОГА ЗЕМЛЕТРЯСЕНИЙ НЕЙРОСЕТЕВЫМИ ТЕХНОЛОГИЯМИ

Нейросетевое компьютерное моделирование используется для разработки новых эффективных алгоритмических и программных средств по созданию регрессионных моделей (аппроксимирующих функций) для сложных объектов, заданных трудно формализуемой системой наблюдений с пробелами и неточностями. Неточно заданные, т. е. содержащие пропуски и ошибки данные, а также заданные приближено, составляют в настоящее время основной объем численной научной и прикладной информации, в том числе и получаемой при сейсмическом районировании.

7.1. Анализ предвестников землетрясений

Важной задачей является предсказание землетрясений по совокупности предвестников, измеряемых в период подготовки сильного сейсмического события.

Построенная нейросетевая модель должна предсказать оставшееся количество дней до землетрясения по данным с нескольких датчиков, измеряющих различные физические показатели на полигоне, предназначенном для изучения динамики поведения предвестников перед сильным землетрясением в Китае.

Период наблюдений составлял несколько месяцев, а выборка содержала момент возникновения сильного землетрясения 1976 г. и его форшока. Построенный по этим данным интегральный предвестник показал весьма обнадеживающие результаты.

На рис. 18 показана карта расстановки регистрирующих станций в провинции Хебей (Китай). Здесь расположение станций, регистрирующих кажущееся сопротивление, отмечено символом ⭘, станций, регистрирующих уровень грунтовых вод – символом ■, место землетрясения отмечено символом ✸.

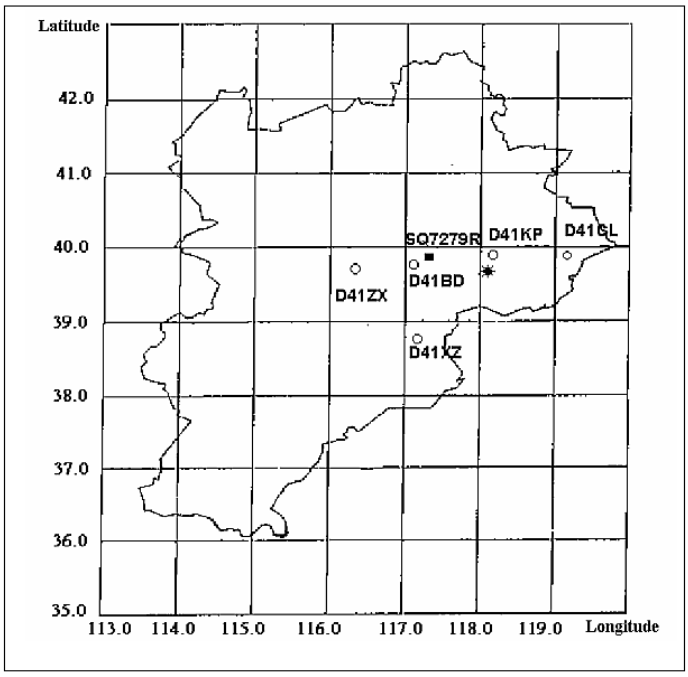

Рис. 18. Карта расстановки регистрирующих станций в провинции Хебей (Китай)

На рис. 19, *а* приведены результаты нейросетевого моделирования информационной значимости сконструированного интегрального показателя на основе данных измерений двух предвестников, регистрирующих свойства среды, с целью прогноза времени сильного землетрясения на территории Китая (1976 г.). По горизонтали отложено истинное время до землетрясения (в днях), по вертикали – прогноз времени землетрясения по двум датчикам.

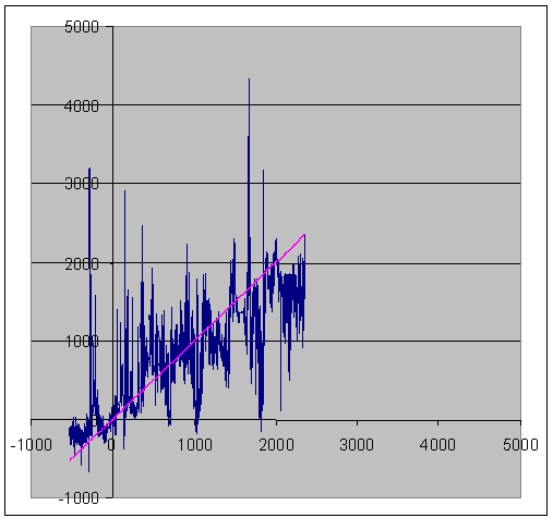

Рис. 19, *а* – возможности прогноза времени землетрясений
по предвестникам различной природы

По указанному интегральному показателю можно уловить тенденцию приближения к моменту главного толчка изучаемого землетрясения (начало координат), но при долгосрочном прогнозе возможны и ложные предупреждения, которые связаны со сложными взаимосвязанными сейсмическими процессами в регионе. Рис. 19, *б* иллюстрирует возможности среднесрочного прогноза времени сильного землетрясения. В первую очередь результаты нейросетевого моделирования оценки информативности интегрального показателя.

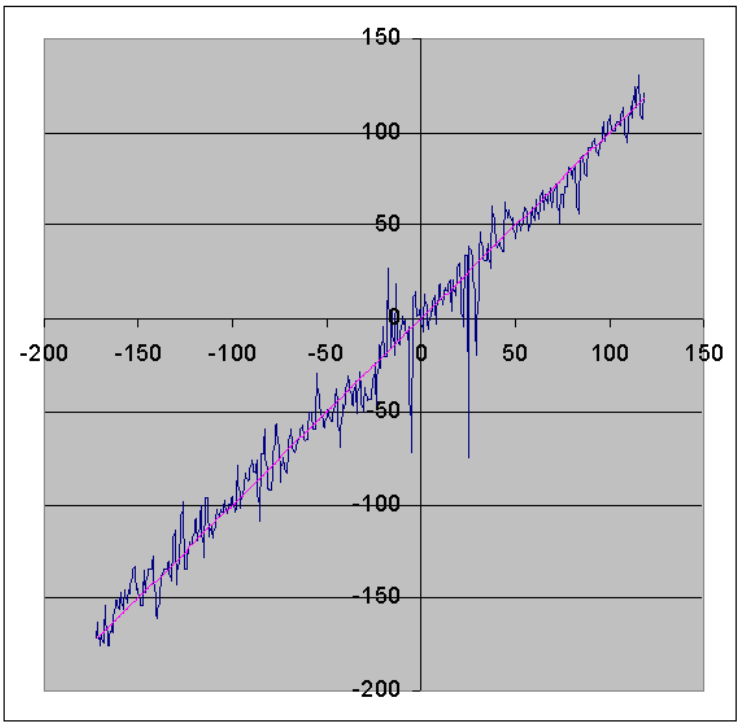

Рис. 19, *б* – результаты нейросетевого моделирования оценки информативности интегрального показателя

График на рис. 19, *б*, построен на основе показаний предвестников землетрясения с 7 датчиков, регистрирующих различные свойства геодинамической среды за несколько месяцев до землетрясения на территории Китая (1976 г.). Здесь по горизонтали отложено истинное время до землетрясения (в днях), а по вертикали – прогноз времени наступления основного толчка этого землетрясения (начало координат). При наличии нескольких датчиков и среднесрочном прогнозе времени основного землетрясения возможен более точный прогноз, чем при долгосрочном прогнозе, но форшоки главного землетрясения способны осложнить процесс интерпретации данных наблюдений.

7.2. Моделирование каталога землетрясений

Известно, что задача статистического анализа сейсмического режима не представляет принципиальной сложности, если количество сейсмических событий на рассматриваемой площади достаточно велико – в этом случае можно определить плотности распределения по территории различных типов сейсмических событий, просто подсчитывая количество имевшихся событий в ячейках, на которые разбивается территория.

Если же количество зарегистрированных землетрясений невелико, решение этой задачи оказывается возможным только с использованием крайне грубого разбиения на ячейки. При этом характеристики сейсмичности определяются весьма приближенно, результаты оказываются чувствительны к случайному характеру происхождения данных.

В рамках проводившегося исследования апробировался новый способ решения указанной проблемы. А именно, ставилась задача создания *специализированного датчика случайных величин*, генерирующего параметры сейсмических событий, воспроизводящий «природный датчик случайных параметров землетрясений».

После синтеза модельного датчика сейсмических событий, настроенного по имеющимся данным, проблема дефицита данных исчезает – можно генерировать их в любых количествах. Принципиально возможно также вовсе не генерировать данные, а просто изучать синтезированный датчик случайных величин, параметры которого однозначно задают интересующие нас в конечном итоге плотности распределения землетрясений.

7.3. Общий принцип построения самонастраивающегося датчика случайных величин

Общий принцип построения самонастраивающегося датчика случайных величин состоит в том, что вводится некоторое вспомогательное многообразие, для которого «хорошо определено» понятие «более плотного» распределения.

Например, если в качестве такого многообразия взять квадрат, разбитый на клеточки одинакового размера, с регулярной решеткой

из центров точек, то «в два раза более плотное, а в остальном такое же» распределение «хорошо определено», а именно, представляет собой объединение множества центров ячеек и множества угловых точек ячеек. Фактически наибольший интерес представляет случай, когда в пределах многообразия распределение равномерное – именно тогда легко вводятся представления о «редком» равномерном распределении и «более плотном» равномерном распределении.

Датчик случайных величин строится как нейросетевое отображение из вспомогательного многообразия в многообразие реальных данных. При этом сначала берется «редкое» распределение точек многообразия, с общим количеством точек соответствующим количеству зарегистрированных сейсмических событий.

Это редкое распределение должно, однако, быть равномерным, если нет оснований приписывать одним событиям априори большую значимость, чем другим. Затем производится адаптация нейросетевого отображения, в сочетании с заменами в разбиениях на пары «точка вспомогательного многообразия – параметры реального события». И адаптация, и перестановки в парах проводятся так, чтобы отклонение «нейросетевых образов» от реальных параметров сейсмических событий стало минимальным.

Когда построено итоговое нейросетевое отображение, производится замена исходного многообразия на «более» плотное. Нейросетевые образы «уплотненного» многообразия и образуют модельное облако «уплотненных», «пополненных» данных.

С принципиальной точки зрения наиболее сложным является вопрос о выборе вспомогательного многообразия. В этой связи отметим то обстоятельство, что размерности «естественных» многообразий данных чаще всего вообще дробные, что довольно сложно воспроизводить на современных компьютерах.

Таким образом, в работе реализуется идея «равноправия» прообразов реальных данных на вспомогательном многообразии и если об области реальных данных известно лишь, что она односвязна, проще всего рассмотреть в качестве кандидатов на роль многообразий-прообразов семейство сфер различной целочисленной размерности.

7.4. Нейросетевой анализ каталога землетрясений

Исследование базировалось на анализе каталога реальных сейсмических событий для территории юга Красноярского края за периоды: 1761–1991 гг., 1992–1997 гг. и по настоящее время.

Для 315 землетрясений на территории края построена их приближенная нейросетевая аппроксимация. По долготе достигнут коэффициент корреляции реальных и аппроксимированных данных R_1=0.976, по широте – R_2=0.984, и по энергетическому коэффициенту – R_3=0.98.

Видимо, можно констатировать близость (но не тождество) плотностей распределения реальных и модельных событий. На рис. 20 указаны реальные данные с разбиением их на три группы, равные по численности и величине энергетического класса: слабые, средние и сильные.

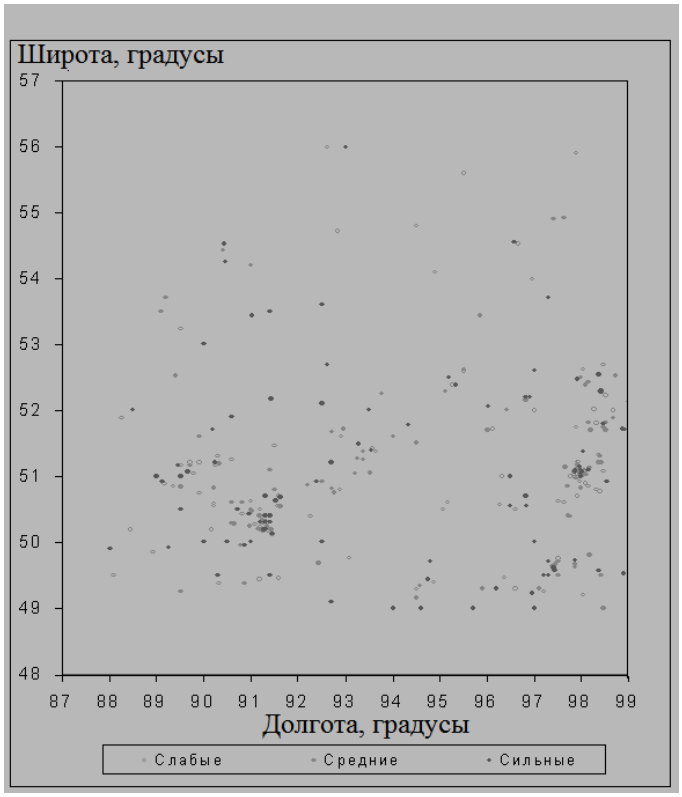

Рис. 20. Распределение сейсмических событий
для трех групп энергетических классов на основе реальных данных

На рис. 21 показаны соответствующие аппроксимированные данные для этих трех типов энергетических классов. В целом по территории среднеквадратичное отклонение эпицентров аппроксимированных и реальных сейсмических событий составляет 56.8 км, в масштабе приведенной схемы (около 900 км * 2200 км) отклонения не чрезмерно велики. Как можно видеть, по сравнению с реальными данными эпицентры модельных сейсмических событий оказываются несколько «размытыми» для некоторых точек, в которых происходят повторные землетрясения, хотя «сглаженная» часть статистики моделируется достаточно хорошо.

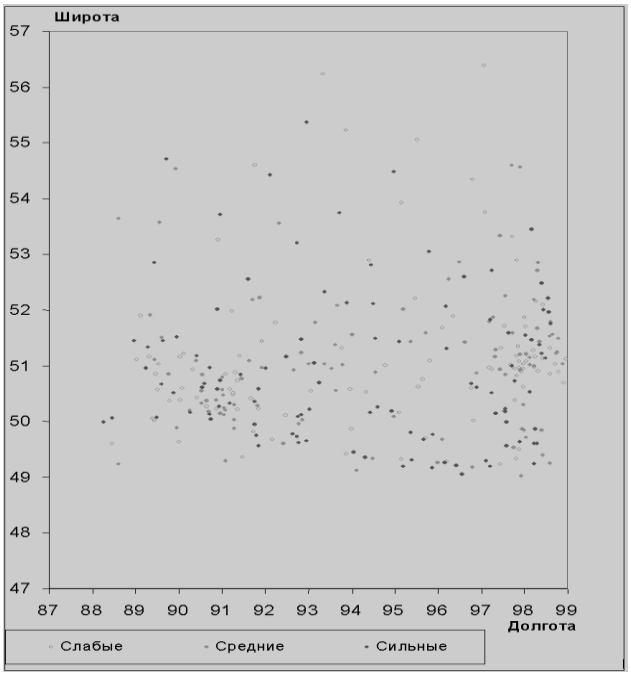

Рис. 21. Распределение сейсмических событий для трех групп энергетических классов: слабые, средние, сильные (аппроксимация данных)

Здесь могут быть также разработаны и применены методы «модельного контрастирования» получаемых сглаженных карт-схем плотности распределения сейсмических событий. Вместе с тем, сглаженная карта-схема распределения сейсмособытий достаточно обоснованно предсказывает средние количества эпицентров сейсмических событий на площадках с размерами порядка 50 км * 50 км.

Традиционные методы сейсморайонирования могут согласовываться с картами-схемами, построенными на основе

нейросетевого моделирования, по данному показателю. На рис. 22 приведены «размноженные» до 3000 модельных событий данные.

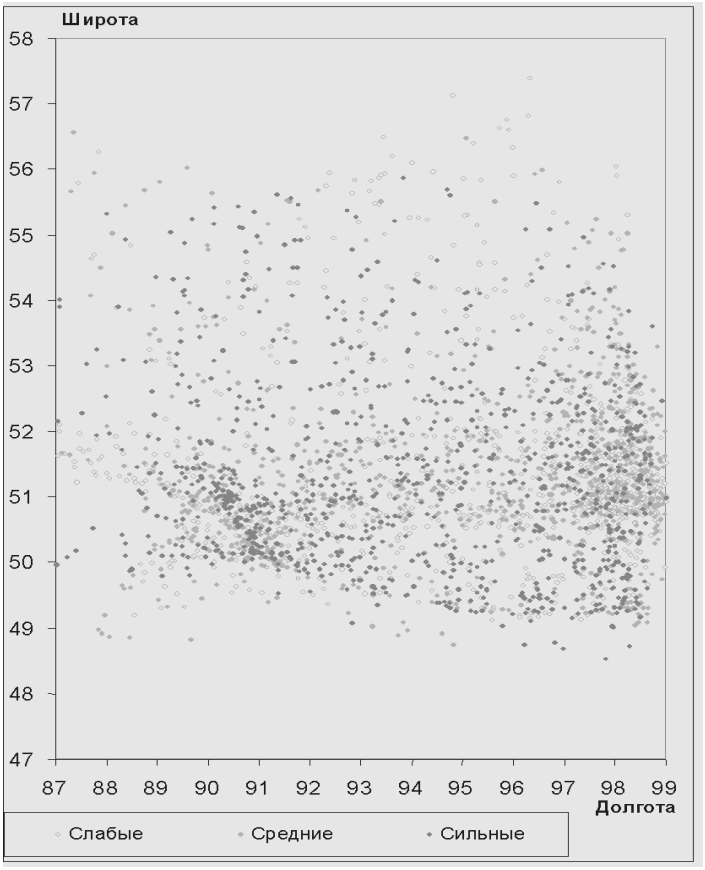

Рис. 22. Распределение сейсмических событий для трех групп энергетических классов («размноженные» данные)

Всего размноженных данных имеется 30000 (в виде файлов), что составляет около одного сейсмического события на 60 кв. км рассматриваемой территории. Этих данных достаточно для построения карты-схемы для анализа сейсмичности с разрешением порядка 20–50 км стандартными методами двумерных гистограмм. При необходимости могут быть рассчитаны дополнительные файлы модельных данных.

На рис. 23 приведено распределение 10 % модельных данных, соответствующих наиболее сильным землетрясениям. Точность данной модели не очень велика, в частности из-за того, что она «обобщает» ограниченную выборку реальных данных (порядка 30 событий). Тем не менее данная модельная карта-схема может качественно визуализировать зоны, в которых возможны относительно сильные землетрясения.

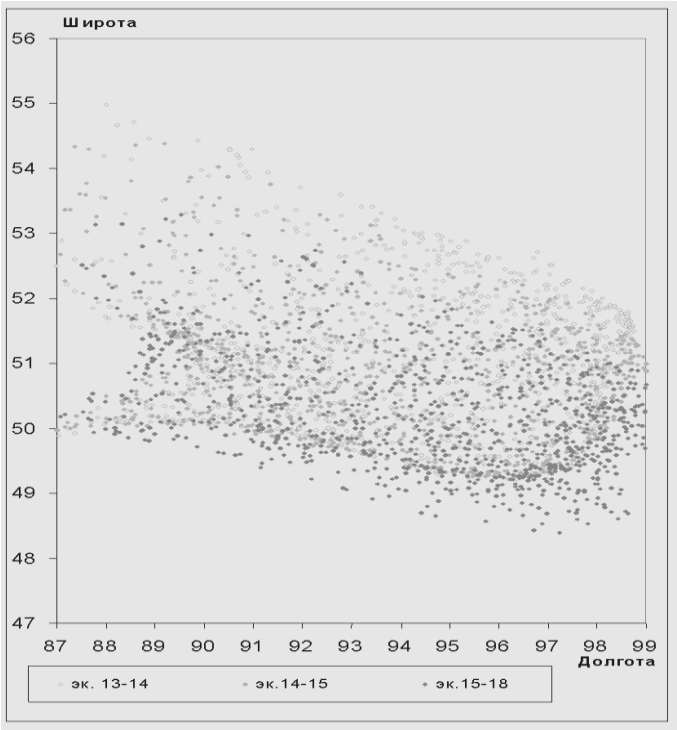

Рис. 23. Распределение сейсмических событий для трех подгрупп энергетических классов (10 % наиболее сильных землетрясений, модель)

7.5. Статистические характеристики модельных распределений сейсмических событий

Отметим, что гистограммы распределения модельных и реальных сейсмических событий на рис. 24 обнаруживают большое сходство. Для распределения магнитуд землетрясений на рис. 25 достигнутое соответствие несколько хуже.

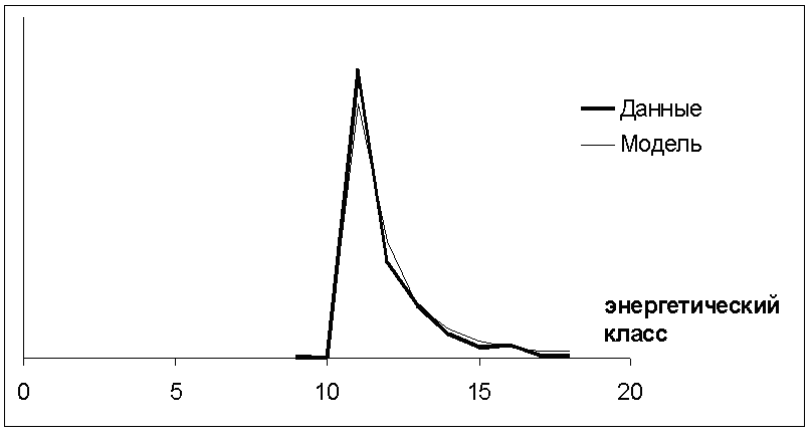

Рис. 24. Сопоставление гистограмм модельных и реальных распределений сейсмических событий по энергетическому классу

Рис. 25. Сопоставление гистограмм модельных и реальных распределений сейсмических событий по магнитуде

Худшее соответствие, в частности, для магнитуды 3, видимо, обусловлено тем, что в исходном каталоге нет землетрясений с этой магнитудой. Наконец, для магнитуд теоретически должно быть распределение, подобное распределению для энергетических классов, поскольку эти две величины пропорциональны друг другу.

В то же время данная пропорциональность в каталоге реальных данных не вполне выдерживается, как это можно видеть на рис. 26.

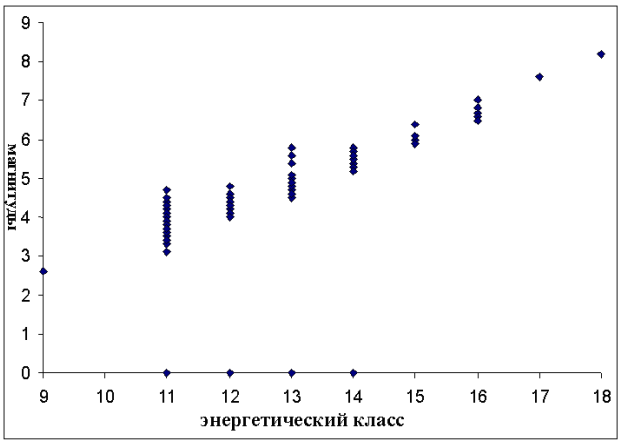

Рис. 26. Соотношение между магнитудой и энергетическим классом для изучаемого каталога землетрясений

Как видно из рис. 26, пропорциональность магнитуд и энергетических классов для части данных выдерживается в каталоге неточно, и даже имеются записи, в которых при различных энергетических классах указывается одна и та же магнитуда. По этой причине нами всюду используются данные по энергетическим классам, видимо, более надежные, чем данные по магнитудам.

Оценка сейсмических сотрясений. Если известны энергии и эпицентры землетрясений, то можно, зная характер затухания сейсмических волн, рассчитать интенсивность проявлений землетрясений на заданном расстоянии от эпицентров. К сожалению, характер распространения сейсмических волн на территории края мало изучен.

Однако, для оценки можно использовать стандартные приближенные формулы, описывающие затухание сейсмических волн. Одна из такого рода формул, ориентировочно описывающая распространение сейсмических волн в регионах с преобладанием скальных грунтов, имеет вид:

$$I = E * 0.97 - 3.12 - 2.48 * LN(R),$$

где I – интенсивность по шкале MSK-64, E – энергетический класс землетрясения, R – расстояние от рассматриваемой точки до эпицентра землетрясения (в км).

При расчетах необходимо учитывать не только широту и долготу эпицентра, но и глубину, на которой располагается гипоцентр землетрясения. В проведенных расчетах предполагалось, что все землетрясения происходят на стандартной глубине 40 км.

По расширенному модельному каталогу (30 000 событий) были рассчитаны частоты сейсмических событий для района расположения ГХК (г. Железногорск) и для района Саяно-Шушенской ГЭС. Результаты по ГХК приводятся на рис. 27, а по Саяно-Шушенской ГЭС – на рис. 28.

Рис. 27. Функция повторяемости для района ГХК (модель)

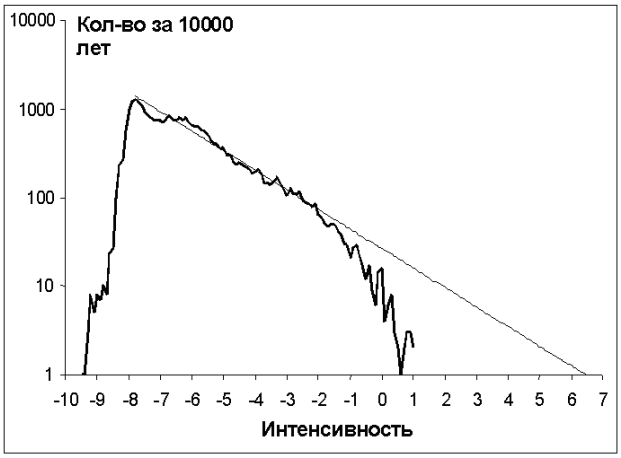

Рис. 28. Функция повторяемости для района Саяно-Шушенской ГЭС (модель)

Как видно на рис. 27 и 28 большинство событий модельного каталога сосредоточено в области крайне малой интенсивности сотрясений. При этом для слабых сейсмических событий прослеживается линейное падение логарифма частоты событий с интенсивностью. Такой линейный характер зависимости хорошо известен в сейсмологии. Для использованного модельного каталога линейность нарушается для редких событий.

Такое нарушение естественно связать с тем, что за период наблюдений, за который были набраны исходные данные, сильные (редкие) сейсмические события для большей части территории не проявляли себя ни разу, поэтому они отсутствуют и в «размноженных» модельных данных.

Следовательно, можно предполагать, что в зоне нарушения линейности модельные данные неточны (занижены). Исправить их можно, продолжив линейную зависимость в область редких событий. Результат линейной экстраполяции показан на рис. 27–28.

Таким образом, согласно итоговому прогнозу, раз в 5–10 тыс. лет в районе ГХК можно ожидать 3–5 балльные сотрясения от землетрясения (по шкале MSK-64), а для района Саяно-Шушенской ГЭС – 6–7 балльных сотрясений. Этот результат согласуется с данными карты ОСР-97 А.

Контрольные вопросы и задания

1. Как можно обосновать суть нейросетевого моделирования пространственно-временных данных?
2. Сформулируйте принципы нейросетевого моделирования случайных данных.
3. Какие существуют способы моделирования данных каталога землетрясений нейросетями?
4. При каких условиях строится нейросетевая модель предвестников землетрясений?
5. Каким образом проводится оценка сейсмических сотрясений?
6. Как определяется плотность распределения по территории различных типов сейсмических событий?
7. Каков принцип построения самонастраивающегося датчика случайных величин?
8. Что означает энергетический класс землетрясения?

Литература: раздел 7, основная литература п. 11, 15, 17; дополнительная литература п. 13–16, 18–19.

8. МОДЕЛИРОВАНИЕ СВЯЗЕЙ ГЕОЛОГО-ГЕОФИЗИЧЕСКИХ ФАКТОРОВ С СЕЙСМИЧНОСТЬЮ

8.1. Модель и алгоритм метода упругих сеток

В этом разделе предложен подход визуализации связей геолого-геофизических факторов с сейсмичностью на основе построения упругой сетки.

Рассмотрим двумерную прямоугольную сетку узлов, в которой p узлов по горизонтали, q узлов по вертикали. Перенумеруем узлы этой сетки с помощью двух индексов – $y^{ij}, i = 1...p, j = 1...q$. Сетка должна обладать следующими свойствами (рис. 29):

• *свойство близости к точкам данных*. Сетка должна быть в каком-то смысле аналогичной плоскости двух первых главных компонент – оптимальной в смысле минимума среднего квадрата расстояния от точек данных до ближайшего узла при определенных ограничениях на свойства сетки;

• *свойство упругости по отношению к растяжению*. Это свойство до некоторой степени обеспечивает *равномерность* сетки;

• *свойство упругости по отношению к изгибу*. Это свойство до некоторой степени обеспечивает *гладкость* результирующего многообразия.

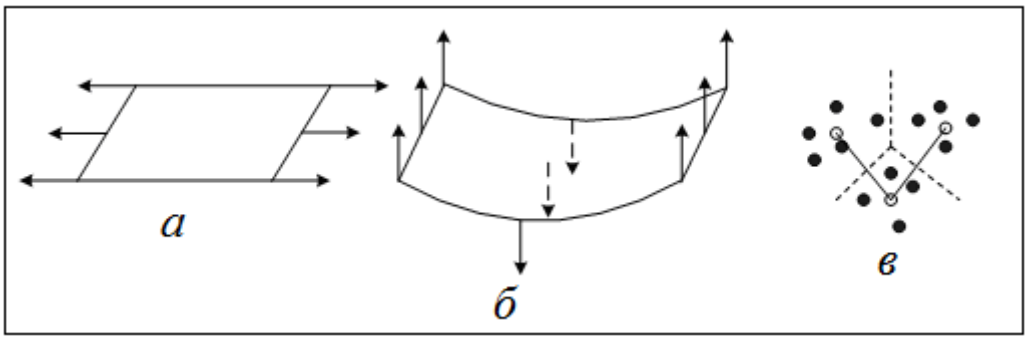

Рис. 29. Иллюстрация к методу упругих карт: силы, растягивающие упругую пластину (*а*); силы, изгибающие упругую пластину (*б*); одномерный случай абсолютно мягкой карты (*в*), каждый узел карты располагается в центре таксона, пунктиром изображены границы зон, в пределах которых точки являются ближайшими к данному узлу – так называемых *ячеек Вороного*

Разобьем все множество данных X на $p*q$ подмножеств $K_{ij}(i=1...p, j=1...q)$ (таксонов), в пределах каждого из которых точки подмножества окажутся ближе к узлу сетки y^{ij}, чем к какому-нибудь другому узлу. Обозначим это обстоятельство следующим образом:

$$K_{ij} = \left\{ x \in X, \left\| y^{ij} - x \right\| \to \min_{i,j} \right\}.$$

В качестве меры близости сетки к точкам данных выберем величину (на одну точку данных) квадрата расстояния от точки до ближайшего узла сетки. Каждый узел сетки (кроме граничных) имеет четырех соседей, с которыми он соединяется «ребром». Чем больше средняя (на один узел) длина ребра, тем сильнее сетка «растянута», поэтому мы должны по возможности минимизировать эту величину.

Таким образом, в минимизируемый функционал должны войти разности между положениями соседних узлов. Степень изогнутости определим с помощью точечной оценки величины второй производной (с помощью так называемых *вторых разностей*). В результате получим следующий функционал «качества» построения сетки:

$$D = \frac{D_1}{|X|} + \lambda \frac{D_2}{pq} + \mu \frac{D_3}{pq} \to \min,$$

где $|X|$ – число точек в X; λ, μ – коэффициенты упругости, отвечающие за растяжение и изогнутость сетки, соответственно; D_1, D_2, D_3 – слагаемые, отвечающие за свойства сетки:

$$D_1 = \sum_{i,j} \sum_{X_k \in K_{ij}} \left\| X_k - y^{ij} \right\|^2,$$

определяется как мера близости расположения узлов сетки к данным.

Здесь K_{ij} – подмножества точек из X, для которых узел сетки y^{ij} является ближайшим (таксоны):

$$D_2 = \sum_{i=1}^{p}\sum_{j=1}^{q-1}\left\|y^{ij} - y^{i,j+1}\right\|^2 + \sum_{i=1}^{p-1}\sum_{j=1}^{q}\left\|y^{ij} - y^{i+1,j}\right\|^2,$$

- мера растянутости сетки;

$$D_3 = \sum_{i=1}^{p}\sum_{j=2}^{q-1}\left\|2y^{ij} - y^{i,j-1} - y^{i,j+1}\right\|^2 + \sum_{i=2}^{p-1}\sum_{j=1}^{q}\left\|2y^{ij} - y^{i-1,j} - y^{i+1,j}\right\|^2,$$

- мера изогнутости (кривизны) сетки.

Отметим, что границы суммирования выбраны так, чтобы в функционале D_2 ребро не входило в сумму дважды.

Пусть метрика является евклидовой. В этом случае функционал D становится квадратичным по положению узлов y^{ij}. Это значит, что при заданном разбиении множества точек данных на таксоны для его минимизации потребуется решить систему линейных уравнений размерами $pq * pq$.

Следовательно, эффективным методом минимизации функционала D оказывается такой алгоритм:

Шаг 0. Узлы сетки, так или иначе, располагаются в пространстве данных.

Шаг 1. При заданных положениях узлов сетки производится разбиение на таксоны – подмножества K_{ij}.

Шаг 2. При заданном разбиении множества точек данных на таксоны производится минимизация функционала D.

Шаги 1 и 2 повторяются до тех пор, пока функционал D не перестанет изменяться (в пределах заданной точности).

Процесс сходится, поскольку на каждом этапе минимизации величина D, очевидно, будет уменьшаться; вместе с тем она ограничена снизу нулем. Более того, он сходится за конечное число шагов, поскольку число вариантов разбиения точек данных на таксоны конечно (рис. 30).

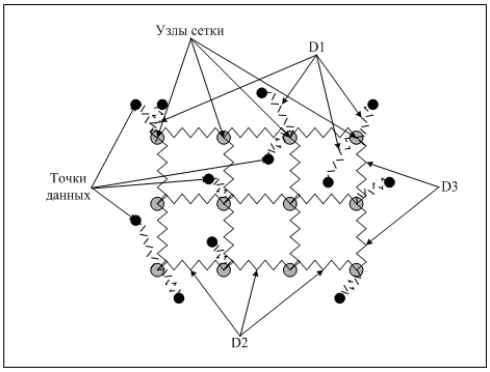

Рис. 30. Упругая сетка

Ниже приведена блок-схема алгоритма построения упругих карт и последующая визуализация данных (рис. 31).

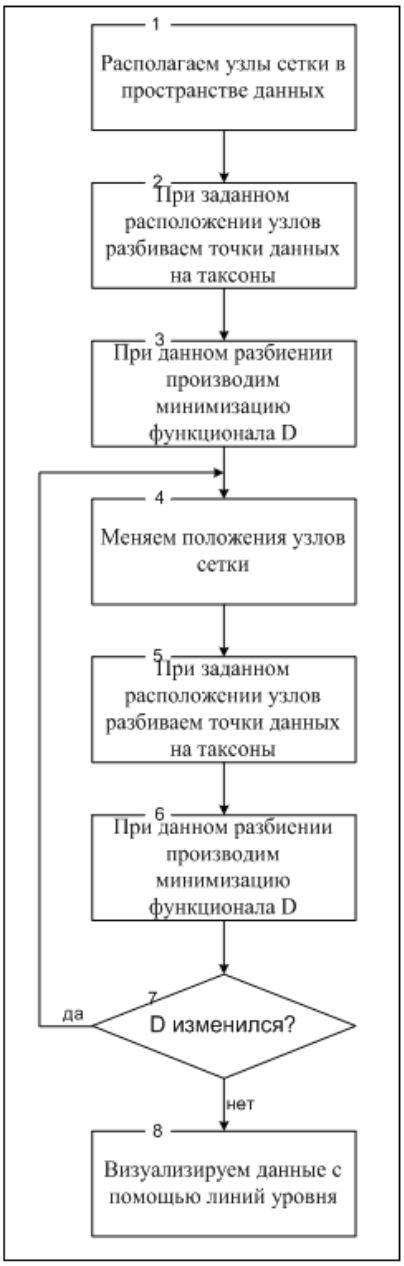

Рис. 31. Алгоритм визуализации сейсмогеологических данных

Приведем краткое описание данных и результаты расчетов. Таблица геологических критериев представляет собой набор геологических параметров, описывающих свойства земной коры. Следовательно, в каждом узле сетки имеется совокупность критериев с различными значениями.

Составление расчетной таблицы проводилось в следующем порядке: выделение сейсмогеологических критериев, их оцифровка; разграфка карт на стандартные ячейки; оцифровка ячеек и разноска сейсмогеологических признаков на карты (всего 14 критериев).

Выделение сейсмогеологических признаков (критериев) производится на основе анализа соответствующих фондовых материалов и публикаций по региону, обобщения полевых материалов, дешифрирование аэрофото и космоснимков.

8.2. Составление карт сейсморайонирования

Составление карт сейсморайонирования предполагает выделение наиболее рационального комплекса геологических признаков. Для удобства построения карт выделяются мелкие блоки, согласно стандартного расчленения площадей по соответствующим градусам, минутам, секундам восточной долготы и северной широты. Территория охватывает площадь между 52° до 57° с.ш. и 90° – 98° в.д. Площадь каждой стандартной ячейки – 412,5 км^{2}. Всего ячеек 32х20=640; общая площадь карты – 264 тыс. км^{2}.

Основополагающими геологическими документами, по которым производился отбор геологических признаков (критериев) для создания карты сейсморайонирования, являются:

- Геологическая карта масштаба 1:500 000, которая составлялась на основе геологической карты масштаба 1:1000 000 (лист №-46 (47), новая серия); геологических карт масштаба 1:200 000 (новые серии); геологических карт масштаба 1:50 000. Составленная геологическая карта масштаба 1:500 000 близка по детальности к картам масштаба 1:200 000. При этом наиболее важные и сложные узлы составлялись преимущественно по съёмкам масштаба 1:50 000.

• Карта – схема неотектоники масштаба 1:500 000 составлена на основе геологической карты масштаба 1:500 000, сети современных разломов, палеосейсмодислокаций, современного вулканизма и других признаков и критериев.

Базовыми геологическими признаками для составления карты сейсморайонирования являются современные разломы, при этом важны: возраст разломов, их параметры, насыщенность (плотность), данные о пересечениях разноориентированных разломов, кольцевых и дуговых разломах и др. (всего пять признаков).

На схеме также приводятся сведения по размещению палеосейсмодислокаций, тектонических уступов, подпорных осадков, современных базальтов, минеральных и термальных источников (всего пять признаков).

Перечисленные геологические и сейсмогеологические критерии являются наиболее важными индикаторами сейсмических событий. Схема отражает блоковое строение территории. Весьма важно установить и отразить на карте стыки различных структур региона. В этом случае блоки ограничиваются разломами различных размеров и значения (генеральные, региональные, локальные).

Исключительно важны сведения о местах стыка (сочленения) разноориентированных сейсмоконтролирующих разломов, а также данные о соседних блоках, имеющих контрастный характер движений, сведения о подобных блоках и разломах исключительно важны, так как чаще всего они контролируют особо опасные землетрясения.

Весьма важны также места стыков различных типов рельефа, узлы смыкания разноориентированных хребтов (всего четыре признака). Таким образом, всего составлено 14 геологических признаков.

Следующим этапом в работе стал сравнительный анализ геологических и сейсмических данных. Сначала была построена модель всех геологических критериев, а затем для оптимального числа признаков (рис. 32, 33).

При сравнении моделей сейсмической активности и геологических критериев значения в соответствующих таксонах нормировались. Затем строилась среднеквадратическая ошибка для каждого критерия.

В результате были выбраны три критерия с наименьшей ошибкой, график ошибки приведен на рис. 34.

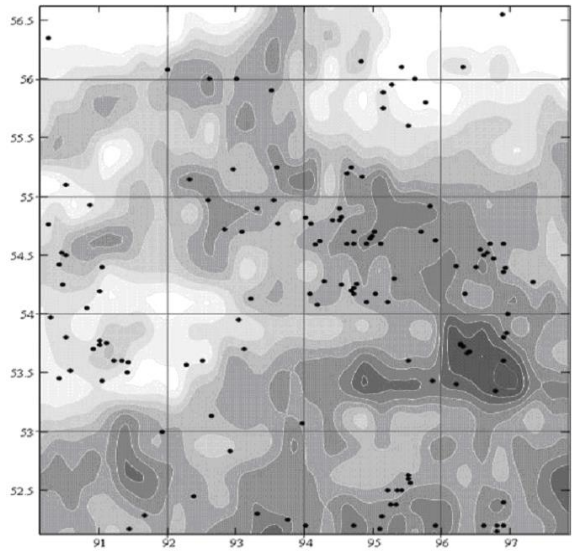

Рис. 32. Визуализация всех геологических критериев

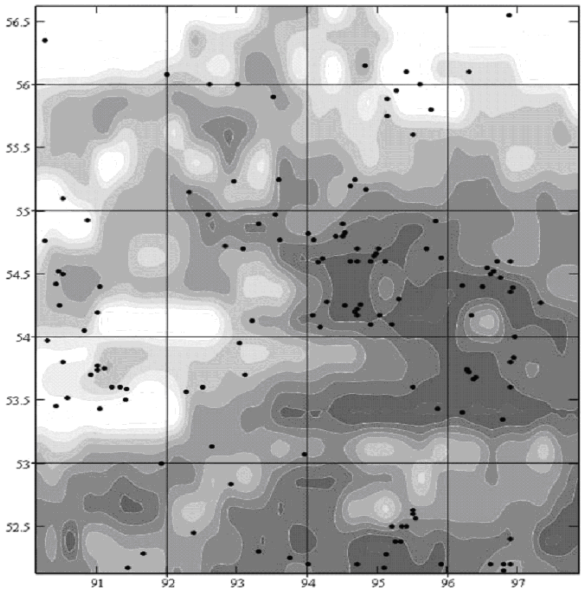

Рис. 33. Визуализация 2, 3, 13 критериев

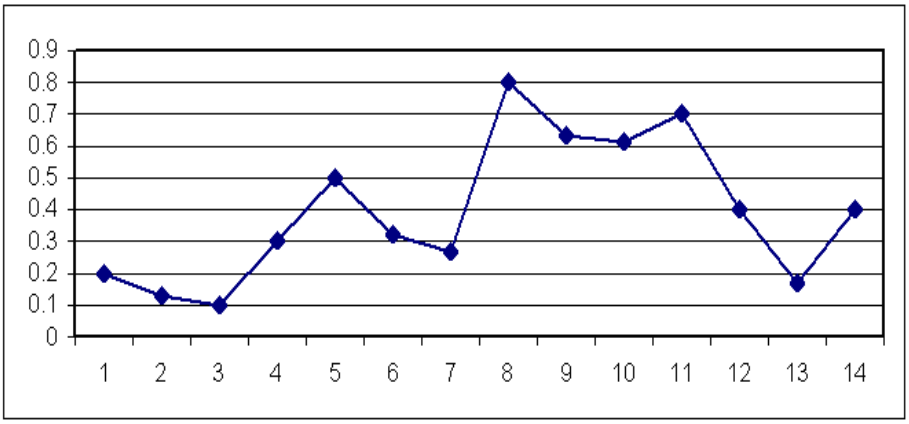

Рис. 34. График среднеквадратической ошибки

Результат визуализации трех выбранных критериев показывает приуроченность некоторых сейсмоактивных зон к областям с определенным геологическим строением. В результате визуализации выявлены несколько сейсмоактивных зон. Для каждой зоны необходимо проводить отдельное исследование, т. е. определить максимально возможную магнитуду события, повторяемость землетрясений, а также геологические критерии характерные для данной зоны.

Следующим этапом стал сравнительный анализ геолого-геофизических и сейсмических данных. На рис. 35 показана модель для всех геолого-геофизических критериев с указанием положения основных сейсмоактивных разломов и эпицентров сильных землетрясений.

Далее, используя геолого-геофизические факторы (их визуализация представлена на рис. 35) и результат решения первого этапа (моделирование сейсмических данных на основе регрессионного подхода), построена прогнозная модель сейсмической активности на основе регрессионной модели связи геолого-геофизических факторов и данных каталога землетрясений. Результат представлен на рис. 36.

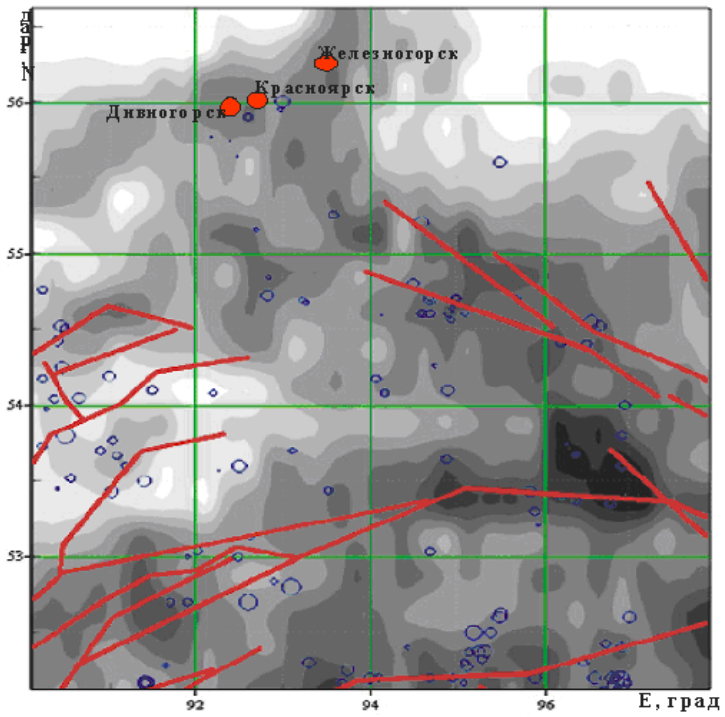

Рис. 35. Модель (визуализация) для геолого-геофизических критериев с указанием положения основных сейсмоактивных разломов и эпицентров сильных землетрясений

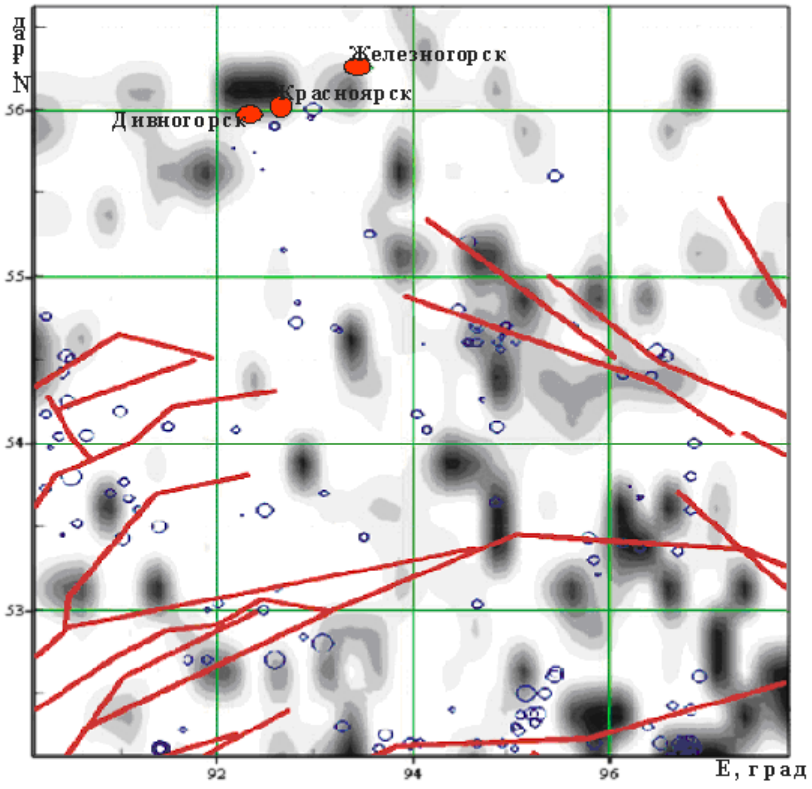

Рис. 36. Визуализация суммарной выделившейся сейсмической энергии. Оценка сейсмической активности (опасности) на основе регрессионной модели связи геолого-геофизических факторов и данных каталога землетрясений

Контрольные вопросы и задания

1. Привести последствия проявлений сильных землетрясений с позиции геоэкологии.
2. Как взаимосвязаны геолого-геофизические факторы и сейсмичность?
3. Каковы особенности геологических и сейсмических данных?
4. Опишите метод построения упругой карты.
5. Какова алгоритмическая схема построения упругой карты?
6. Визуализация карт сейсмического районирования.
7. Какие геологические критерии выбираются при составлении карты сейсморайонирования?
8. Что понимается под сейсмоактивными зонами?

Литература: раздел 8, основная литература п. 10, 14–15; дополнительная литература п. 13–16, 19.

ЗАКЛЮЧЕНИЕ

Представлены алгоритмы и вычислительные методики (способы) анализа сейсмических сигналов, регистрируемых набором датчиков, а также данных комплексного геомониторинга очаговых зон и каталогов землетрясений:

- многопараметрическая нелинейная регрессия сигнала с использованием нейросетевой вычислительной технологии обработки данных, который позволяет выделять сигнал на фоне помех во временной области без каких-либо априорных предположений о частотном составе анализируемого сигнала;
- вейвлет-преобразование сигналов, масштабно-временная разверстка, получающаяся в результате вейвлет-преобразования сигнала, позволяет выявлять не только осцилляции с хорошо фиксированными частотами, но и нестационарные колебания, а также локализованные периодичности и др.;
- спектрально-временной анализ сигналов, который позволяет исследовать дисперсионные характеристики распространения изучаемых колебаний в среде;
- способ построения нейросетевой модели, которая предсказывает оставшееся количество дней до землетрясения по данным с нескольких датчиков, измеряющих различные физические показатели на полигоне, предназначенном для изучения динамики поведения предвестников перед сильным землетрясением, а также методика нейросетевого анализа каталога землетрясений, которая позволяет моделировать данные, восполняя их и выделяя скрытые закономерности;
- способ визуализации связей геолого-геофизических факторов с сейсмичностью на основе построения упругой сетки, с помощью этого подхода: построена модель распределения сейсмической активности на изучаемой территории (юг Красноярского края) и выделены сейсмоактивные зоны.

Результаты комплексного анализа данных геомониторинга представляются обнадеживающими для достижения поставленной цели – изучение сейсмического режима в очаговых зонах и оценки степени сейсмической опасности.

ГЛОССАРИЙ

Абсолютная погрешность – это модуль разности точного значения и приближенного значения.

Аномалия – отклонение от нормы или общей закономерности; неправильность.

Афтершок – подземный повторный толчок, следующий за главным сейсмическим ударом из одной с ним очаговой области, возникший при сейсмических процессах.

Вейвлеты – это обобщённое название математических функций определенной формы, которые локальны во времени и по частоте и в которых все функции получаются из одной базовой, изменяя её (сдвигая, растягивая).

Вейвлет-анализ – представление исследуемого процесса в виде линейной комбинации различных функций, именуемых базисом соответствующего преобразования.

Вейвлет-преобразование (Wavelet transform, WT) – интегральное преобразование, которое представляет собой свертку вейвлет-функции с сигналом.

Восстановление данных – процедура извлечения информации с запоминающего устройства в случае, когда она не может быть прочитана обычным способом.

Волновое поле – область упругой среды, вовлеченная в волновой процесс.

Вычислительный эксперимент – метод изучения геофизических процессов (объектов) с помощью математического моделирования и вычислительных средств.

Геодинамика – раздел геофизики, изучающий природу глубинных сил и процессов, возникающих в результате планетарной эволюции Земли, и обуславливающих движение вещества внутри планеты.

Геодинамическая безопасность – защищённость природно-технических систем и объектов, которые функционируют в тех районах, где человек ведет активную промышленную деятельность.

Геомониторинг – система регулярных целенаправленных наблюдений, сбора, накопления, обработки, обобщения и анализа информации для оценки состояния окружающей среды.

Геофизические поля – физические поля, образованные планетой Земля.

Гипоцентр – **фокус** землетрясения, условный центр очага землетрясения.

Годограф сейсмических волн – график зависимости времени пробега волны от источника до приемника волны (регистрирующего устройства) от эпицентрального расстояния.

Детектор прямолинейного движения – частный случай поляризационной фильтрации, принимающий во внимание только прямолинейность в радиальном и вертикальном направлениях.

Длина сейсмической волны – расстояние между соседними гребнями или впадинами волны.

Землетрясение – подземные толчки и колебания земной поверхности, вызванные образованием очага высвобождения энергии деформации в некоторой зоне земной коры.

Изосейсты – условные линии, соединяющие территории с одинаковой интенсивностью землетрясения.

Интенсивность землетрясения – сейсмический эффект, мера воздействия землетрясения на объекты, выражаемая баллами сейсмической шкалы и его кинематическими параметрами (например, ускорением).

Интерполяция – метод нахождения неизвестных промежуточных значений некоторой функции по имеющемуся дискретному набору ее известных значений.

Корреляционный анализ – метод, позволяющий обнаружить зависимость между несколькими случайными величинами.

Коэффициент корреляции – параметр, характеризующий связь между двумя переменными.

Магнитуда – величина, представляющая собой десятичный логарифм максимальной амплитуды сейсмической волны (в тысячных долях миллиметра), записанной стандартным сейсмографом на расстоянии 100 км от эпицентра землетрясения.

Математическая модель – представление изучаемого явления или процесса с помощью математических понятий.

Математическое моделирование – процесс построения и изучения математических моделей реальных процессов и явлений.

Мера локальной перемежаемости – мера локальных отклонений от среднего поля спектров на каждом масштабе, позволяет определить степень неравномерности распределения энергии по масштабам.

Мода – значение во множестве наблюдений, которое встречается наиболее часто.

Микросейсмы (микросейсмические колебания) – относительно слабые колебания Земли, вызванные суммарным воздействием множества сейсмических источников, характеристики которых являются случайными величинами.

Натурный эксперимент – способ познания объекта природы, заключающийся в изучении природных явлений в специально созданных условиях.

Нелинейная регрессия – вид регрессионного анализа, в котором экспериментальные данные моделируются функцией, являющейся нелинейной комбинацией параметров модели и зависящей от одной и более независимых переменных.

Нейросетевая модель – математическая модель, а также её программное или аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей.

Носитель информации – любой материальный объект или среда, содержащий (несущий) информацию, способный достаточно длительное время сохранять в своей структуре занесённую в(на) него информацию.

Относительная погрешность – это частное от деления абсолютной погрешности на модуль измеряемой величины.

Очаг землетрясения – область зарождения землетрясения или область внутри Земли, где происходит образование разломов и возникновение сейсмических волн.

Период сейсмической волны – интервал времени между двумя соседними гребнями или впадинами; величина, обратная частоте волны. Обычно период сейсмических волн составляет 0.1–10 сек.

Плейстосейстовая область – территория, где землетрясение проявилось с максимальной силой.

Поверхностные сейсмические волны – сейсмические волны, распространяющиеся только вдоль поверхности Земли. Скорость поверхностных волн меньше скорости поперечных волн. Различают два типа поверхностных волн: волны Лява и Рэлея.

Поперечная волна – это волна, при распространении которой смещение частиц среды происходит в направлении, перпендикулярном распространению волны.

Преобразование Фурье (*FT*) – операция, сопоставляющая одной функции вещественной переменной другую функцию вещественной переменной.

Продольная волна – это волна, при распространении которой смещение частиц среды происходит в направлении распространения волны.

Регрессия – односторонняя стохастическая зависимость, устанавливающая соответствие между случайными переменными, т. е. математическое выражение, отражающее связь между зависимой переменной y и независимыми переменными x при условии, что это выражение будет иметь статистическую значимость.

Регрессионный анализ – статистический метод исследования взаимосвязи переменных.

Сейсмическое воздействие – колебательное движение грунта, вызванное природными или техногенными факторами и вызывающее движение или деформации зданий и сооружений. Сейсмические воздействия выражаются в баллах шкалы сейсмической интенсивности, в ускорениях, в амплитудах колебаний грунта или в иных характеристиках, использующихся при проектировании зданий и сооружений.

Сейсмические волны – упругие (механические) колебания, распространяющиеся в Земле от очагов землетрясений, взрывов и др. источников и переносящие энергию упругих колебаний в горных породах.

Сейсмичность – статистическое распределение интенсивности землетрясения на выделенной территории в зависимости от его повторяемости и наличия возможных очагов.

Сейсмическая опасность – максимальные сейсмические воздействия на заданной территории, превышение которых возможно с определенной вероятностью в течение заданного интервала времени.

Сейсмическое событие – кратковременный процесс, результатом которого являются сейсмические сигналы, регистрируемые сейсмографом.

Сейсморайонирование – это направление геофизических работ, основная задача которого заключается в уточнении данных общего сейсмического микрорайонирования и определении степени сейсмической опасности застраиваемых территорий.

Сейсмическое микрорайонирование – часть инженерно-геологических и инженерно-геотехнических изысканий на площадках строительства особо опасных объектов.

Сейсмовоздействие – колебательное принудительное движение условной платформы (основания), сообщающей закреплённому на ней объекту переносное (во внешней неподвижной системе отсчета) ускорение, заданное акселерограммами в общем случае в трех ортогональных направлениях движения.

Сейсмограмма – совокупность сейсмических трасс, сгруппированных по общему признаку.

Сжатие данных – алгоритмическое преобразование данных, производимое с целью уменьшения занимаемого ими объёма.

Сигнал – материальное воплощение сообщения для использования при передаче, переработке и хранении информации.

Скейлограмма – графическое представление данных в виде диаграммы специального вида для вейвлет-анализа.

Форшок – предварительный более слабый толчок, предшествующий главному сейсмическому удару, связанный с началом образования разломов.

Экстремум – максимальное или минимальное значение функции на заданном множестве.

Энергия землетрясений – потенциальная энергия среды, которая при землетрясении идет на разрушение материала среды, трение при подвижке и образование сейсмических волн.

Энергетический класс – количественная мера величины землетрясений; десятичный логарифм высвободившейся сейсмической энергии, измеренной в Джоулях.

Эпицентр землетрясения – проекция гипоцентра на земную поверхность.

Эпицентральное расстояние – расстояние между эпицентром землетрясения и рассматриваемой точкой, измеренное вдоль поверхности Земли. Выражается в линейных (км) или угловых (градус) величинах.

СПИСОК ЛИТЕРАТУРЫ

Основная литература

1. Самарский А. А., Михайлов А. П. Математическое моделирование. Идеи. Методы. Примеры. – М.: Физматлит, 2001. – 320 с.

2. Малинецкий Г. Г., Курдюмов С. П. Нелинейная динамика и проблемы прогноза // Вестник РАН. – 2001. – Т 71. – № 3. – С. 210–232.

3. Моисеев Н. Н. Математические задачи системного анализа. – М.: Наука, 1981.

4. Марчук Г. И. Математическое моделирование в проблеме окружающей среды – М.: Гидрометеоиздат, 1982. – 316 с.

5. Пентл Р. Методы системного анализа окружающей среды. – М.: Мир, 1981.

6. Джефферс Дж. Введение в системный анализ: применение в экологии. – М.: Мир, 1981. – 256 с.

7. Белолипецкий В. М., Шокин Ю. И. Математическое моделирование в задачах охраны окружающей среды. – Новосибирск: Инфолио-пресс, 1997. – 240 с.

8. Джексон П. Введение в экспертные системы. – М.: «Вильямс». – 2001.

9. Канасевич Э. Р. Анализ временных последовательностей в геофизике. – М.: Недра, 1985. – 400 с.

10. Зиновьев А. Ю. Визуализация многомерных данных. – Красноярск: ИПЦ КГТУ, 2000.

11. Горбань А. Н., Россиев Д. А. Нейронные сети на персональном компьютере. – Новосибирск: Наука, 1996. – 276 с.

12. Григорьев Ал. А., Кондратьев К. Я. Спутниковый мониторинг природных и антропогенных катастроф // Исследование Земли из космоса. – 1996. – № 3. – С. 68–79.

13. Григорьев Ал. А., Кондратьев К. Я. Природные и антропогенные экологические катастрофы // Исследование Земли из космоса. – 2000. – № 2. – С. 72–82.

14. Природные опасности России / Под общей ред. В. И. Осипова, С. К. Шойгу, Т. 1–6. – М.: Изд-во «КРУК», 2003.

15. Сейсмический риск и инженерные решения / Под ред. Ц. Ломнитца и Э. Розенблюта. – М.: Наука, 1981. – 375 с.

16. Акимов В. А., Новиков В. Д., Радаев Н. Н. Природные и техногенные ЧС: опасности угрозы, риски. – М.: ЗАО ФИД «Деловой экспресс», 2001. – 344 с.

17. Управление риском: Риск. Устойчивое развитие. Синергетика. – М.: Наука, 2000. – 431 с.

Дополнительная литература

1. Афанасьев К. Е., Стуколов С. В. Многопроцессорные вычислительные системы и параллельное программирование. – Кемерово: Кузбассвузиздат, 2003. – 233 с.

2. Бат М. Спектральный анализ в геофизике. – М.: Недра, 1980.

3. Бендат Д., Пирсол А. Прикладной анализ случайных данных. – М.: Мир, 1989. – 540 с.

4. Блаттер А. Вейвлет-анализ. Основы теории – М.: Техносфера, 2004. – 273 с.

5. Воробьев В. И., Грибунин В. Г. Теория и практика вейвлет-преобразования. – СПб.: Изд–во ВУС, 1999.

6. Девис Дж. Статистика и анализ геологических данных. – М.: Мир, 1977.

7. Добеши И. Десять лекций по вейвлетам. – М.: «РХД», 2001.

8. Кашкин В. Б. Автоматизированная обработка изображений. Космические средства контроля окружающей среды. – Красноярск: ИПЦ КГТУ, 2000.

9. Кравцов Ю. А. Случайность, детерминированность, предсказуемость // Успехи физических наук. – 1989. – Т 158. – № 1. – С. 93–122.

10. Лбов Е. С. Методы обработки разнотипных экспериментальных данных. – Новосибирск, 1981. – 156 с.

11. Марпл.-мл. С. А. Цифровой спектральный анализ и его приложения. – М.: Мир, 1990. – 584 с.

12. Пределы предсказуемости / Под ред. Ю. А. Кравцова – М.: Центрком, 1997.

13. Садовский М. А., Писаренко В. Ф. Сейсмический процесс в блоковой среде. – М.: Наука, 1991. – 96 с.

14. Соболев Г. А. Основы прогноза землетрясений. – М.: Наука, 1993. – 313 с.

15. Соболев Г. А., Пономарев А. В. Физика землетрясений и предвестники. – М.: Наука, 2003. – 270 с.

16. Федотов С. А. Долгосрочный сейсмический прогноз для Курило-Камчатской дуги. – М.: Наука, 2005. – 302 с.

17. Чуи К. Введение в вейвлеты. – М.: Мир, 2001. – 412 с.

18. Шахраманьян М. К. Оценка сейсмического риска и прогноз последствия землетрясений в задачах спасения населения (Теория и практика). – М.: ВНИИ ГУЧС, 2000. – 247 с.

19. Шебалин Н. В. Сильные землетрясения. – М.: Изд-во академии горных наук, Избр. труды, 1997. – 547 с.

Учебное издание

ТЕОРИЯ ОБРАБОТКИ ГЕОФИЗИЧЕСКИХ ДАННЫХ

Учебно-методическое пособие

Составители:
Симонов Константин Васильевич
Перетокин Сергей Анатольевич
Рублева Татьяна Васильевна

Корректор *А. В. Прохоренко*
Компьютерная верстка *И. В. Владимировой*

Подписано в печать 16.04.2021. Печать плоская. Формат 60×84/16
Бумага офсетная. Усл. печ. л. 5,8. Тираж 100 экз. Заказ № 12921

Библиотечно-издательский комплекс
Сибирского федерального университета
660041, Красноярск, пр. Свободный, 82а
Тел. (391) 206-26-16; http://bik.sfu-kras.ru,
E-mail: publishing_house@sfu-kras.ru