

**Московский государственный университет им. М.В. Ломоносова
Российский государственный геологоразведочный университет
имени Серго Орджоникидзе**

А.А. Никитин, А.А. Булычев

**КОМПЛЕКСНЫЙ АНАЛИЗ
И КОМПЛЕКСНАЯ ИНТЕРПРЕТАЦИЯ
ГЕОФИЗИЧЕСКИХ ПОЛЕЙ**

Учебное пособие

Для студентов и магистров МГУ, МГРИ-РГГРУ,
обучающихся по специальностям «Геофизика» и «Технология
геологической разведки»

Печатается по решению Ученого совета геологического факультета
МГУ имени М.В. Ломоносова (протокол № 8 от 17 декабря 2015 г.)

**Москва
2015**

УДК 550.83(075)

ББК 26.324.34

Н624

Рецензенты:

*Галуев В.И., доктор технических наук, главный научный сотрудник
ФГУП ГНЦ РФ «ВНИИгеосистем»*

Шевнин В.А., доктор физико-математических наук, профессор кафедры геофизических методов исследования земной коры геологического факультета МГУ имени М.В. Ломоносова

Никитин А.А., Булычев А.А.

Н624 Комплексный анализ и комплексная интерпретация геофизических полей : учебное пособие / А.А. Никитин, А.А. Булычев. – М. : ВНИИгеосистем, 2015. – 94 с. : ил.

ISBN 978-5-8481-0205-5

Для студентов и магистров МГУ, МГРИ-РГГРУ, обучающихся по специальностям «Геофизика» и «Технология геологической разведки».

Допущено УМС геологического факультета
МГУ имени М.В. Ломоносова в качестве учебного пособия для студентов, обучающихся по направлению «Геология», профиль «Геофизика».

ISBN 978-5-8481-0205-5

© Никитин А.А., Булычев А.А., 2015

© Оформление. ВНИИгеосистем, 2015

ОГЛАВЛЕНИЕ

Введение	4
Глава I. Основные определения и процедуры комплексного анализа	6
Глава II. Оценка информативности атрибутов и их комплекса	9
Глава III. Алгоритмы комплексного анализа при наличии эталонных объектов	13
Глава IV. Алгоритмы комплексного анализа при отсутствии эталонных объектов	20
Глава V. Развитие метода главных компонент при наличии корреляционных связей между полями по площади наблюдений	25
Глава VI. Математическая постановка задач количественной комплексной интерпретации	32
Глава VII. Критерии адекватности физико-геологических моделей реальной среде	37
Глава VIII. Комплексная интерпретация геофизических полей при построении глубинных моделей земной коры	43
Глава IX. Технология «прямого» прогноза углеводородов на основе комплексного анализа атрибутов геофизических полей	56
Глава X. Технологии построения согласованных ФГМ осадочного чехла и фундамента при интерпретации данных наземных и скважинных методов	61
Глава XI. Геоинформационные системы и технологии комплексного анализа и интерпретации геофизических данных	72
Контрольные вопросы по комплексному анализу и комплексной интерпретации геофизических данных	87
Литература.....	88
Приложение.....	89

ВВЕДЕНИЕ

Комплексный анализ и комплексная интерпретация геофизических полей, наряду с непосредственным комплексированием геофизических методов, представляет наиболее эффективное средство повышения однозначности решения обратных задач геофизики.

Если комплексный анализ по набору предварительно рассчитанных атрибутов полей обеспечивает решение задач структурно – тектонического районирования изучаемых площадей, то комплексная интерпретация геополей ориентирована на количественную оценку геометрических (глубины, формы и элементов залегания) и физических (физические свойства) параметров искомых объектов (аномалиеобразующих источников).

Решение задач комплексного анализа реализуется как при наличии, так и при отсутствии эталонных объектов.

При наличии эталонных объектов используются методы математической логики, регрессионного анализа, проверки статистических гипотез, нейронные сети. Для этих приемов наиболее широко применяется термин «распознавание образов», хотя в последнее время под этим термином понимаются и приемы комплексного анализа при отсутствии эталонных объектов. Среди последних нашли применение метод главных компонент и метод *K*-средних и их различные модификации, кроме того, все чаще используются генетические алгоритмы. Основной задачей количественной комплексной интерпретации является построение согласованных (по геометрии и физическим параметрам) физико-геологических моделей (ФГМ) в пределах изучаемого геологического разреза для 2D и 3D моделей.

Математическая постановка задач количественной комплексной интерпретации реализуется как для детерминированного подхода Г.Я. Голиздры [2], так и для статистического подхода по Ф.М. Гольцману [3]. На этой основе обеспечивается построение согласованных по скорости и плотности, по скорости и проводимости, по плотности и магнитной восприимчивости моделей. Привлечение имеющихся данных по геофизическим исследованиям скважин (ГИС) реализует возможность привязку к конкретным геологическим породам и распространение моделей в межскважинное пространство.

По построению согласованных ФГМ особое значение приобретает оценка адекватности моделей реальной среде, в частности, по величине так называемого обобщенного расстояния (по А.А. Никитину). Для решения задач комплексного анализа нашли применение также автоматизированные и геоинформационные системы, как ГИС ПАРК (М.А. Белобородов), ГИС

ИНТЕГРО (М.Я. Финкельштейн, Е.Н. Черемисина), КОСКАД-3Д (А.В. Петров), ПАНГЕЯ (В.В. Колесов и др.), Multalt (Ф.М. Гольцман, Д.Ф. Калинин).

Для практической реализации построения согласованной ФГМ разработаны автоматизированные и геоинформационные системы: ГИС ИНТЕГРО ГЕОФИЗИКА (В.И. Галуев), GCIS (А.И. Кобрунов, А.П. Петровский), СЭВР и СОМІNTER (Н.П. Смилевец), СИГМА-3D (Ю.И.Блох), КОСКАД-3Д (А.В. Петров, А.А. Никитин), GeoFrame, Petrel (Schlumberge) и др.

ГЛАВА I. ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ И ПРОЦЕДУРЫ КОМПЛЕКСНОГО АНАЛИЗА ГЕОДАННЫХ

Под комплексным анализом понимается обработка данных нескольких геофизических полей и их атрибутов (признаков, характеристик). При этом решаются задачи поисков геокартирования и районирования площади наблюдений на несколько классов, охарактеризованных заданным набором атрибутов. Целью комплексного анализа является построение структурно-тектонической карты, а также выделение слабоконтрастных аномалиеобразующих объектов.

Комплексный анализ реализуется при двух подходах: при наличии эталонных объектов и при их отсутствии. При наличии эталонных объектов, т.е. объектов с установленной геологической природой нашли применение методы математической логики, регрессионного анализа в виде регрессионных нейронных сетей и проверки статистических гипотез. Для этих методов утвердился термин распознавание образов.

Ввиду изменения геологического строения, особенно в пределах больших территорий, методы комплексного анализа эффективны при достаточно детальных геофизических съемках.

При отсутствии эталонных объектов (распознавание на принципах самообучения) обеспечивается разделение пунктов наблюдения на некоторое, обычно наперед заданное число классов, геологическая природа которых неизвестна. Среди методов самообучения наиболее популярны метод главных компонент и метод K -средних и их модификации.

При распознавании образов с обучением на эталонных объектах все объекты делятся на объекты обучения, распознаваемые объекты и объекты экзамена. Для последних часть эталонных объектов не входит в обучение, а оставляется для оценки качества распознавания.

Основные процедуры комплексного анализа сводятся к следующему:

- задание математической модели комплексного анализа;
- расчеты атрибутов геофизических полей в скользящих окнах и корреляционных связей между атрибутами;
- оценка информативности атрибутов и их сочетаний;
- выбор алгоритма комплексного анализа в зависимости от наличия априорной информации, корреляционных связей между атрибутами;
- проведение непосредственного комплексного анализа по выбранному алгоритму;
- оценка качества комплексного анализа.

Математическая модель комплексного анализа для L -полей $\vec{F}_{li}$ или атрибутов при разбиении площади исследований на M -классов объектов для i -й точки наблюдений может быть определена для M -гипотез H_κ как

$$\vec{F}_i = \begin{cases} \text{либо } \vec{S}_{li} + \vec{n}_{li} \rightarrow H_1 \\ \dots\dots\dots \\ \text{либо } \vec{S}_{Mi} + \vec{n}_{Mi} \rightarrow H_M \end{cases}, \kappa=1, \dots, M, l=1, \dots, L, I=1, \dots, n,$$

где $\vec{S}_{Mi}$ и $\vec{n}_{Mi}$ – соответственно аномальная составляющая поля и случайная помеха. Знак вектора относится к числу полей $l=1, \dots, L$.

При разделении объектов на рудные (H_1) и безрудные (H_2) достаточно ограничиться двумя классами (гипотезами), т.е. $M=2$, а в задачах районирования и геокартирования $M>2$.

Расчет атрибутов как по исходным полям, так и по их составляющим обычно реализуется в скользящих окнах заданного размера. Для состоятельности и эффективности оценок атрибутов необходимое число точек в окне должно быть не менее 25-30.

В качестве атрибутов используется достаточно большое число статистических оценок, градиентных, корреляционных и спектральных характеристик.

Статистические оценки – это среднее значение поля (атрибута) в окне $\bar{x}$, дисперсия D , асимметрия A ; эксцесс E ; коэффициент вариации $V = \bar{x}/\sigma$, где $\sigma = \pm\sqrt{D}$, энтропия $H = -\sum_1^n p_i \log p_i$, p_i – частота встречаемости i -го значения в окне.

Градиентные атрибуты: градиент поля по оси x $\partial f/\partial x$; по оси y $\partial f/\partial y$, полный градиент $\Delta = \sqrt{(\partial f/\partial x)^2 + (\partial f/\partial y)^2}$, направление полного градиента $\alpha = \arctg(\Delta y/\Delta x)$.

При этом следует отметить, что градиенты рассчитываются по огибающей поля для трех–пяти значений поля, поскольку при резкой изменчивости поля оценки градиентов по соседним значениям поля не устойчивы.

Для оценки корреляционных связей между полями и атрибутами строится корреляционная матрица R по парным коэффициентам корреляции

$$r_{lj} = \frac{1}{n\sigma_l\sigma_j} \sum_1^n (x_{li} - \bar{x}_l)(x_{ji} - \bar{x}_j), l, j = 1, \dots, L; i = 1, \dots, n$$

При наличии эталонных объектов такая матрица строится для каждого из K -классов $k = 1, \dots, M$.

Оценка информативности атрибутов и их комплекса, а также сами алгоритмы комплексного анализа с оценками качества комплексного анализа изложены в последующих трех главах.

ГЛАВА II. ОЦЕНКА ИНФОРМАТИВНОСТИ АТТРИБУТОВ И ИХ КОМПЛЕКСА

Наиболее универсальной оценкой информативности как отдельно взятого признака (поля), так и сочетаний (комплекса) признаков (полей) следует считать *надежность* разделения объектов двух классов γ , определяемую по разности площадей под гистограммами распределений признака S_1 и S_2 , построенными для объектов классов H_1 и H_2 :

$$\gamma = 1 - q, \text{ где } q = \frac{S_{\text{пер.}}}{S_1 + S_2}. \quad (2.1)$$

Здесь q – вероятность ошибки разделения классов H_1 и H_2 , равная $q = p_1\alpha + p_2\beta$, где p_1 и p_2 – априорные вероятности появления объектов классов H_1 и H_2 соответственно. Обычно эти априорные вероятности полагают равными, т.е. $p_1 = p_2 = 0,5$.

В случае разделения объектов двух классов, что соответствует решению задачи поисков, надежность может быть определена через интеграл вероятности, в котором верхним пределом является отношение аномалия/помеха (сигнал/помеха) ρ [8]:

$$\gamma = \Phi(\sqrt{\rho}/2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\sqrt{\rho}/2} e^{-x^2/2} dx. \quad (2.2)$$

Зависимость (2) приведена на рис. 2.1.

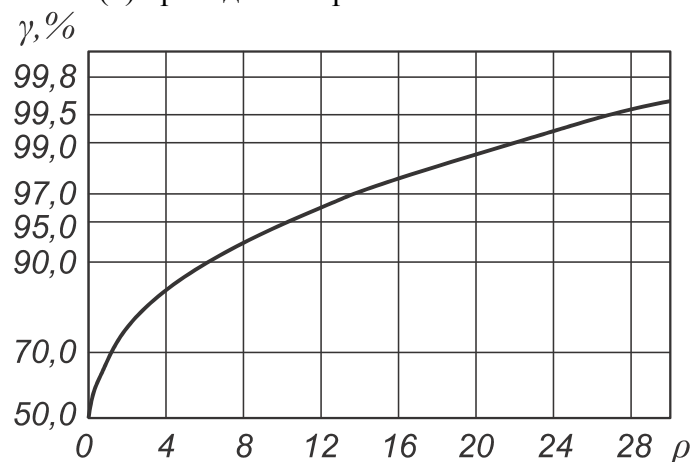


Рис. 2.1. Зависимость надежности обнаружения аномалии γ
от энергетического отношения аномалия/помеха ρ

Величина α представляет вероятность ошибки I рода, которая состоит в том, что принимается решение о наличии объекта класса H_2 , в то время как

на самом деле имеется объект класса H_1 . Величина β – есть вероятность ошибки II рода, состоящая в том, что принимается решение о наличии объекта класса H_1 , в то время как на самом деле имеет место объект класса H_2 . Вероятности ошибок α и β определяются величинами площадей пересечения $S_{пер.}$ гистограмм значений признаков, рассчитанных соответственно для объектов классов H_1 и H_2 (рис. 2.2, а). Сумма вероятностей ошибок I и II рода характеризует степень расхождения гистограмм значений признака X_i : $p(X_i/H_1)$ и $p(X_i/H_2)$. При полном их совпадении признак X_i не обладает информативностью, поскольку в этом случае $\gamma=0,5$ при $p_1=p_2=0,5$. Наоборот, при отсутствии пересечения гистограмм решение о наличии объектов H_1 и H_2 принимается однозначно и величина надежности $\gamma=100\%$. Признаки, для которых наблюдается совпадение гистограмм в пределах доверительных интервалов частот, можно считать неинформативными и при комплексном анализе в дальнейшем их не использовать.

В целом один класс объектов рассматривается как помеха для другого. Следовательно, определив величину γ по формуле (2.1), можно, используя зависимость (2.2), найти «аналог» отношения сигнал/помеха ρ для задачи поисков или геокартирования на два класса. Далее определяется величина такого сигнала ρ для каждого признака после чего, при условии независимости признаков, рассчитывается отношение сигнал/помеха по комплексу признаков, как сумма

$$\rho_{комп.} = \sum_1^L \rho_l, L - \text{общее число признаков, } \rho_l.$$

Надежность разделения объектов двух классов по комплексу l -признаков определяется как

$$\gamma_{комп.} = \Phi \left(\frac{\sqrt{\sum \rho_l}}{2} \right). \quad (2.3)$$

Выражение (2.3) может быть использовано для оценки информативности, т.е. надежности разделения объектов, по любому сочетанию признаков. По максимальной величине $\gamma_{комп.}$ устанавливается информативная совокупность признаков. Такой совокупностью обеспечивается выбор наиболее эффективного в смысле оценки геологической природы источников аномалий комплекса геофизических методов.

При решении задачи геокартирования на число классов, большее двух, величина ошибки разделения классов q определяется суммой площадей попарного пересечения гистограмм. Например, для классов H_1 , H_2 и H_3 , $q=q_{12}+q_{23}+q_{13}$ (см. рис. 2.2, в).

Соответственно, надежность разделения будет равна $\gamma=1-q$, а по формуле (2.2) находится величина ρ_1 для первого признака. Аналогично вычисляются значения ρ для других признаков, а затем определяется надежность разделения по их комплексу по формуле (2.3). Этот прием легко реализуется для любого числа признаков, любого их сочетания и любого числа классов.

В качестве примера рассмотрим расчет величины γ по гистограммам, приведенным на рис. 2.2, а и б. Эти гистограммы характеризуются следующими значениями:

5, 10, 15, 20, 20, 15, 10, 5..... для $P(x/H_1)$
 5, 10, 15, 20, 20, 15, 10, 5..... для $P(x/H_2)$
 10, 20, 40, 20, 10..... для $P(x/H_1)$
 5, 8, 12, 15, 20, 15, 12, 8, 5..... для $P(x/H_2)$.

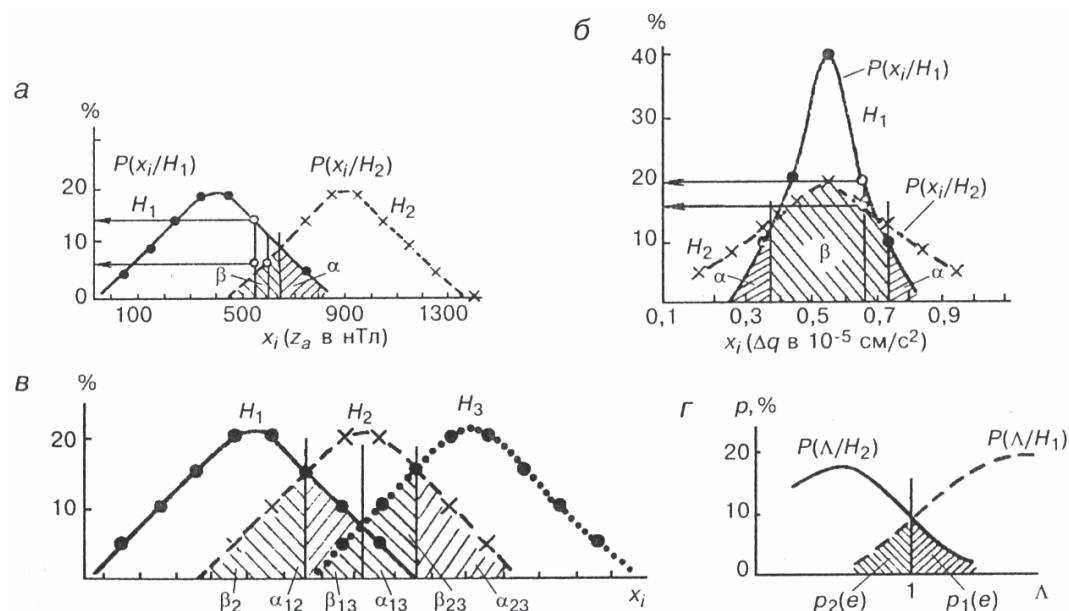


Рис. 2.2. Плотности распределения признака x_i (а-в) и коэффициента правдоподобия Λ (г)

В первом случае (рис. 2.2, а) при равных значениях ρ_1 и ρ_2 величина $q=0,5(5+10+5)=10\%$. Соответственно надежность разделения $\gamma=90\%$.

Во втором случае (рис. 2.2, б) $q=0,5(10+15+20+15+10)=35\%$ и соответственно $\gamma=65\%$. Ясно, что разделение признаков в первом случае более эффективно, чем во втором. Количественно это разделение оценивается ве-

личины γ . Для оценки информативности совокупности двух рассмотренных признаков, используя зависимость рис. 2.1, найдем величину $\rho_1=6$ по $\gamma_1=90\%$ и $\rho_2=1$ по $\gamma_2=65\%$. Отсюда $\rho_{\text{комп}}=7$, $\gamma_{\text{комп}}=91\%$.

Определение информативности признака, связанное с надежностью разделения, нетрудно обобщить на случай любого числа классов [4, 8]. Так, например, для трех классов, гистограммы которых приведены на рис. 2.2, в, получаем

$$\gamma_{\text{комп}} = 1 - (p_1\alpha_{12} + p_2\beta_{23}) - (p_2\alpha_{23} + p_3\beta_{23}) - (p_1\alpha_{13} + p_3\beta_{13}) = 1 - q_{12} - q_{23} - q_{13},$$

где p_1-p_3 – априорные вероятности появления объектов классов H_1-H_3 , полагаемые обычно равными между собой; α_{12} , α_{23} и α_{13} – вероятности ошибок первого рода при разделении объектов соответственно первого и второго, второго и третьего, первого и третьего классов, β_{12} , β_{23} и β_{13} – аналогичные значения вероятностей для ошибок II рода.

Для случая, приведенного на рис. 2.2, в, гистограммы представлены следующими значениями:

5, 10, 15, 20, 20, 15, 10, 5..... для $P(x/H_1)$

5, 10, 15, 20, 20, 15, 10, 5..... для $P(x/H_2)$

5, 10, 15, 20, 20, 15, 10, 5..... для $P(x/H_3)$.

При $p_1=p_2=p_3=0,33$ получаем $\gamma_{\text{комп}}=69\%$.

Использование оценок информативности в виде (2.3) позволяет сравнивать информативность тех признаков, для которых разделение осуществляется на разное число классов, например для признака x_1 на два класса, а для признака x_2 на три класса. С целью сравнения информативности по величине γ в этом случае следует считать, что какие-то два класса по признаку x_1 не разделяются (либо H_1 и H_3 , либо H_2 и H_3).

ГЛАВА III. АЛГОРИТМЫ КОМПЛЕКСНОГО АНАЛИЗА ПРИ НАЛИЧИИ ЭТАЛОННЫХ ОБЪЕКТОВ

Комплексный анализ геофизических данных при наличии эталонных объектов

В настоящее время используется большое число алгоритмов комплексного анализа признаков с *предварительным обучением на эталонных объектах*, основные из них ориентированы на методы теории статистических решений, математической логики и так называемых потенциальных функций.

Комплексный анализ на основе статистических решений. *Алгоритм по проверке статистических гипотез* базируется на критериях принятия статистических решений, среди которых чаще всего применяются критерии максимального правдоподобия и максимума апостериорной вероятности.

Для их использования строятся оценки плотностей распределения значений признаков (гистограмм) и корреляционных матриц по эталонным объектам разных классов.

Рассмотрим алгоритм обработки для случая двух классов, который соответствует сравнению двух гипотез H_1 и H_2 на основе вычисления коэффициента правдоподобия. Для независимых признаков коэффициент правдоподобия равен отношению функций правдоподобия $P(x_l/H_k)$, вычисляемых для каждого l -го признака и по каждому k -му классу:

$$\Lambda(\vec{X}) = \frac{P(x_1 / H_1)P(x_2 / H_1)...P(x_L / H_1)}{P(x_1 / H_2)P(x_2 / H_2)...P(x_L / H_2)}. \quad (3.1)$$

Значения частных коэффициентов правдоподобия

$$\Lambda(x_l) = P(x_l / H_1) / P(x_l / H_2)$$

характеризуют относительный вклад каждого признака в общую величину $\Lambda(\vec{X})$. Эти значения можно рассматривать как количественную оценку информативности признаков. В (3.1) вектор $\vec{X}$ – это вектор L -полей F_i .

Распознавание исследуемых объектов сводится к нахождению в i -й точке (ячейке) значений $P(x_{li} / H_1)$ и $P(x_{li} / H_2)$, которые снимаются с гистограмм, полученных на эталонных объектах. После этого рассчитываются величины коэффициента правдоподобия (3.1). Следует отметить, что ограничения на законы распределения признаков при этом отсутствуют. Решение о принадлежности i -го искомого объекта (точки, ячейки)

к классу H_1 принимается согласно критерию максимального правдоподобия при $\Lambda(\vec{X}_i) > 1$.

Если $\Lambda(\vec{X}_i) < 1$, справедлива гипотеза H_2 , т.е. объект принадлежит ко второму классу.

На основании формулы Байеса по коэффициенту правдоподобия находится апостериорная вероятность гипотезы H_1 :

$$P(H_1 / \vec{X}_i) = \frac{(p_1 / p_2) \Lambda(\vec{X}_i)}{(p_1 / p_2) \Lambda(\vec{X}_i) + 1}, \quad (3.2)$$

где p_1 и p_2 – априорные вероятности появления объектов соответственно классов H_1 и H_2 .

Приняв $p_1 = p_2$, получаем правило решения для апостериорной вероятности в виде $p(H_1 / \vec{X}_i) > 0,5$, когда справедлива гипотеза H_1 , при $p(H_1 / \vec{X}_i) < 0,5$ – справедлива гипотеза H_2 .

Поясним расчет коэффициента правдоподобия (3.1) и апостериорной вероятности (3.2) на примере.

Пусть гистограммы значений двух признаков (Z_a и Δg) над эталонными объектами двух классов H_1 и H_2 имеют вид, аналогичный приведенному соответственно на рис. 2.2, а и б. Требуется установить принадлежность точки наблюдений (искомого объекта) со значениями $x_1 = Z_a = 550$ нТл и $x_2 = \Delta g = 0,65$ мГал, к одному из классов.

По гистограмме рис. 2.2, а получаем $P(x_1/H_1)=0,15$ (15%), $P(x_2/H_1)=0,05$ (5%), а по гистограмме рис. 2.2, б $P(x_2/H_1)=0,2$ (20%), $P(x_2/H_2)=0,17$ (17%). Тогда согласно формуле (3.1) имеем $\Lambda(x_1, x_2) = (15/5) \cdot (20/17) = 3 \cdot 1,2 = 3,6 > 1 = \Lambda_{\text{пор}}$, т.е. объект со значениями $Z_a=550$ нТл и $\Delta g=0,65$ мГал принадлежит к классу H_1 .

Значение апостериорной вероятности из (3.2) при $p_1=p_2=0,5$ будет $P(H_1/x_1, x_2) = 3,6 / (3,6 + 1) = 0,76 > 0,5$, что соответствует гипотезе H_1 .

При условии, что изучаемые (распознаваемые) объекты отнесены к первому классу H_1 , т.е. являются теми точками наблюдений, которые оконтуриваются изолиниями со значениями $\Lambda(\vec{X}) > 1$, коэффициент правдоподобия в силу своей случайной природы в оконтуренной области будет иметь некоторое вполне определенное распределение. Обозначим плотность такого распределения через $P(\Lambda(\vec{X}) / H_1)$. При условии, что объекты относятся ко второму классу H_2 , плотность распределения коэффициента правдоподобия в этой области будет равна $P(\Lambda(\vec{X}) / H_2)$. Очевидно, что плотность $P(\Lambda(\vec{X}) / H_1)$ не равна тождественно $P(\Lambda(\vec{X}) / H_2)$.

Таким образом, получаем, два одномерных распределения $P(\Lambda / H_1)$ и $P(\Lambda / H_2)$ (см. рис. 2.2, г). Интегрируя эти распределения, можно оценить вероятностью ошибки проведенной обработки и ее надежность.

Так, вероятности ошибок первого и второго родов (обозначим их через $p_1(e)$ и $p_2(e)$) в этом случае равны соответственно

$$p_1(e) = \int_{-\infty}^{\Lambda_{пор.}} P(\Lambda(\vec{X}) / H_1) d\Lambda(\vec{X}); \quad p_2(e) = \int_{\Lambda_{пор.}}^{\infty} P(\Lambda(\vec{X}) / H_2) d\Lambda(\vec{X}).$$

В частности, $\Lambda_{пор.}$ может быть равно единице.

Общая безусловная вероятность ошибки

$$p(e) = \sum_{k=1}^2 p(H_k) p_k(e).$$

Отсюда получаем надежность классификации после распознавания $\gamma = 1 - p(e)$.

В случае необходимости распознавания объектов на число классов, большее двух ($M > 2$), вычисления производятся непосредственно по формуле Байеса

$$p(H_k / \vec{X}_i) = \frac{p_k P(\vec{X}_i / H_k)}{\sum_{k=1}^M p_k P(\vec{X}_i / H_k)}. \quad (3.3)$$

При этом для каждого класса последовательно находят значения $p(H_1 / \vec{X}_i), \dots, p(H_M / \vec{X}_i)$ и по максимальному из них принимают решение о принадлежности искомой i -й точки к тому или иному классу.

При построении алгоритмов распознавания в случае зависимых признаков исходят из предположения об их нормальном распределении и оценивают корреляционные матрицы R_k , составленные из коэффициентов корреляции r_{lj} ($l, j = 1, \dots, L$) по каждому классу в отдельности. Многомерная плотность распределения вектора признаков $\vec{X} = (X_1, \dots, X_L)$ для класса H_k определяется выражением

$$P(\vec{X}_i / H_k) = \frac{1}{(2\pi)^{L/2} |R_k|^{1/2}} \exp \left[-\frac{1}{2} (\vec{X}_i - \vec{\mu}_k)' R_k^{-1} (\vec{X}_i - \vec{\mu}_k) \right], \quad (3.4)$$

где R_k^{-1} – матрица, обратная матрице R_k ; $|R_k|$ – определитель корреляционной матрицы; $\vec{\mu}_k$ – вектор средних значений признаков k -го класса; $(\vec{X}_i - \vec{\mu}_k)'$ и $(\vec{X}_i - \vec{\mu}_k)$ – вектор-строка и вектор-столбец.

Используя опять формулу Байеса (3.3), решение о принадлежности объекта к тому или иному классу принимают по максимальной величине апостериорной вероятности из M чисел $p(H_k / \vec{X}_i)$, $k = 1, \dots, M$. Реализация такого алгоритма весьма громоздка, поэтому с целью упрощения используются различные допущения, одним из которых является условие равенства матриц $R_1 = R_2 = \dots = R_M = R$. При этом выражение для коэффициента правдоподобия принимает вид

$$\Lambda_i = \exp \left[-\frac{1}{2} (\vec{X}_i - \vec{\mu}_1)' R^{-1} (\vec{X}_i - \vec{\mu}_1) + \frac{1}{2} (\vec{X}_i - \vec{\mu}_2)' R^{-1} (\vec{X}_i - \vec{\mu}_2) \right]. \quad (3.5)$$

При $\Lambda_i > 1$ принимается решение о наличии объекта класса H_1 . Логарифмирование последнего выражения приводит к алгоритму

$$\vec{X}_i - [(\vec{\mu}_1 - \vec{\mu}_2) / 2]' R^{-1} (\vec{\mu}_1 - \vec{\mu}_2) > 0. \quad (3.6)$$

Левая часть (3.6) носит название *дискриминантной функции*.

Поскольку вектор признаков $\vec{X}_i$ входит в (3.6) линейно, то при равенстве нулю это выражение описывает гиперплоскость в L -мерном пространстве признаков, разделяющую это пространство на две области $\vec{X}_i \in H_1$ и $\vec{X}_i \in H_2$. Алгоритм построения дискриминантной функции можно использовать и для распознавания объектов при $M > 2$, производя последовательное деление каждого класса на два подкласса.

Наконец, в тех случаях, когда имеется лишь один класс эталонных объектов и решается задача о принадлежности искомого объекта по комплексу признаков именно к этому классу, можно использовать критерий χ^2 . В указанной ситуации возможна проверка следующих гипотез:

1) о равенстве средних векторов признаков эталонного и распознаваемого объектов. Для проверки этой гипотезы используется правило

$$\sum \frac{(x_{i1} - x_{i2})^2}{\sigma_{i1}^2 + \sigma_{i2}^2} \leq \chi_{\beta}^2(L), \quad (3.7)$$

где σ_{i1}^2 и σ_{i2}^2 оценки дисперсий i -го признака x_i для эталонного (1) и распознаваемого (2) объектов; $\chi_{\beta}^2(L)$ – квантиль вероятности β распределения χ^2 с L степенями свободы;

2) о равенстве средних значений и среднеквадратических отклонений. Эта гипотеза проверяется по правилу

$$\frac{n_1 n_2}{n_1 + n_2} \left[\frac{(\bar{x}_1 - \bar{x}_2)^2}{\sigma^2} + 2 \left(\ln \frac{\sigma_1}{\sigma_2} \right)^2 \right] \leq \chi_{\beta}^2(L), \quad (3.8)$$

где $\sigma^2 = (n_1\sigma_1^2 + n_2\sigma_2^2) / (n_1 + n_2)$; n_1 и n_2 – объемы выборок соответственно для эталонного (1) и распознаваемого (2) объектов $\bar{x}_1$ и $\bar{x}_2$, σ_1^2 и σ_2^2 – оценки соответственно средних векторов и дисперсий признаков эталонного и распознаваемого объектов;

3) о равенстве корреляционных матриц при условии зависимости признаков. Эта гипотеза имеет решающее правило в виде

$$(\bar{X}_1 - \bar{X}_2)' \left[\frac{R_1}{n_1} + \frac{R_2}{n} \right]^{-1} (\bar{X}_1 - \bar{X}_2) \leq \chi_\beta^2(L), \quad (3.9)$$

где R_1 и R_2 – оценки корреляционных матриц для эталонных и распознаваемых объектов.

Для рассмотренных выше алгоритмов решение о принадлежности объекта к тому или иному классу принимается на основе фиксированного порога $\Lambda_{\text{пор.}} = 1$; $\chi_\beta^2 = \chi_{0,9}^2$ и т.д. Эти алгоритмы для зависимых признаков строятся в предположении об их нормальном распределении. На преодоление последнего ограничения направлена разработка алгоритмов непараметрической классификации, для которых выбор порога принятия решения происходит в процессе обучения на эталонных объектах.

Комплексный анализ на основе математической логики. *Алгоритмы математической логики* основаны на установлении меры сходства (меры аналогии) исследуемых объектов с эталонными по величине суммарной информативности признаков. Одним из первых в прикладной геофизике был предложен алгоритм «Кора-3» (Ш.А. Губерман), с помощью которого решалась задача выделения нефтегазоносных пластов по комплексу методов геофизических исследований скважин. Кодирование признаков при этом осуществляется в двоичной системе, когда объект задается набором признаков $\vec{F}$, причем $f_{kl} = 0$, если k -й объект не обладает l -м признаком, и $f_{kl} = 1$, если k -й объект обладает l -м признаком. Все эталонные и исследуемые объекты характеризуются единым набором признаков и единым порядком их расположения. Этап обучения сводится к перебору всех возможных сочетаний признаков по три (отсюда цифра «3» в названии алгоритма и соответствующей программы) для каждого класса.

Если появилось сочетание признаков, которое не менее p раз встречается среди эталонов первого класса и ни разу не встречается среди эталонов других классов, такое сочетание выделяется как сложный признак первого класса. Пороговое значение p задается эмпирически. Если один из двух сложных признаков характеризует большее число эталонных объектов, то он будет информативнее другого. На следующем этапе осуществляется распо-

знание исследуемых объектов. При этом проверяется, сколько сложных признаков первого класса встретилось в искомом объекте. Если сложных признаков первого класса больше, чем сложных признаков других классов, объект относится к первому классу. С увеличением мощности и памяти современных ЭВМ указанный алгоритм позволяет осуществлять перебор сочетаний признаков по любому их числу и не ограничивается кодированием признака двумя состояниями.

Другим распространенным алгоритмом является *алгоритм тупиковых тестов* (А.Н. Дмитриев, Ю.И. Журавлев), основанный на поиске матрицы T , содержащей описание объектов обучения (объекты, как правило, кодируются в двоичной системе) специальных подмножеств – так называемых тупиковых тестов. Подмножество столбцов $(i_1, i_2, \dots, i_l)$ матрицы T называется тестом, если после удаления из T всех столбцов, за $(i_1, i_2, \dots, i_l)$, получается матрица, все строки которой различны. Тест называется тупиковым, если после исключения любого столбца он перестает быть тестом.

Для каждого из признаков $x_1, \dots, x_L$, характеризующих объекты обучения, вычисляется информационный вес признака $h_i = n_i / N$ ($i=1, \dots, L$), где n_i – число вхождений признака x_i в тупиковые тесты; N – общее число тупиковых тестов для матрицы T . Информационный вес исследуемого объекта определяется как сумма информационных весов, которые присутствуют на данном объекте, т.е. принимают значение «1».

Для каждого объекта обучения $a_k (x_{k1}, \dots, x_{kL})$ и исследуемого объекта $a_x (x_1, \dots, x_L)$ находится вероятность:

$$p_{li} = \begin{cases} 0 & \text{при } x_{kl} \neq x_l \\ 1 & \text{при } x_{kl} = x_l \end{cases},$$

$p_{li} = 1/2$, если значение x_{kl} или x_l соответствует прочерку (отсутствие данных о признаке), где x_{kl} – значение l -го признака объекта k -го класса.

Затем для каждого из M классов оценивают величину комплексного параметра (меры сходства)

$$y_{\text{комп.}k} = \frac{1}{n_k} \sum_{i=1}^{n_k} \sum_{l=1}^L p_{li} h_l, \quad k = 1, \dots, M,$$

где n_k – число объектов обучения k -го класса.

Объект a_x относится к тому классу (образу), которому соответствует максимальная величина $y_{\text{комп.}k}$.

Достаточно широкое применение получили алгоритмы распознавания, основанные на построении потенциальной функции, выражаемой через обобщенное расстояние для изучаемых признаков в L -мерном пространстве.

Алгоритм «Потенциал-2» (И.Д. Савинский) позволяет для каждого исследуемого объекта найти в L -мерном пространстве признаков «потенциал» вида

$$V = \frac{1}{n_k} \sum_{k=1}^{n_k} \frac{\alpha_k}{\left[\sum_{i=1}^L h_i (x_{kl} - x_l)^2 \right]},$$

где x_{kl} , x_l – координаты l -го признака k -го объекта обучения и исследуемого объекта; n_k – число объектов обучения k -го класса; α_k , h_l – «веса» k -го объекта обучения и l -го признака, соответственно обусловленные совокупностью объектов различных классов.

Исследуемый объект относится к тому классу, «потенциалы» (меры сходства) которого будут большими, чем для других классов.

С целью оценки качества обработки по этому алгоритму на самих объектах обучения предусмотрена процедура поочередного исключения объектов из обучающей выборки и их контроль на основе остающихся классов. Число ε ошибочно распознаваемых объектов обучения при этом служит одним из показателей эффективности распознавания.

ГЛАВА IV. АЛГОРИТМЫ КОМПЛЕКСНОГО АНАЛИЗА ПРИ ОТСУТСТВИИ ЭТАЛОННЫХ ОБЪЕКТОВ

Комплексный анализ геофизических данных при отсутствии эталонных объектов

Отсутствие эталонных объектов, на которых проводится оценка статистических характеристик и информативности изучаемых признаков, приводит к необходимости разработки алгоритмов обработки данных комплекса на принципах самообучения. *Распознавание образов без обучения на эталонных объектах называется самообучением.* При этом число классов и статистические характеристики (распределение) признаков по каждому классу должны быть получены в процессе обработки исходных данных. При распознавании образов без обучения термин «классификация» больше отвечает сущности задачи, поскольку в конечном итоге площадь исследования разделяется на конкретное число однородных по комплексу признаков классов. Геологическая природа таких классов остается неизвестной из-за отсутствия возможности их сравнения с некоторыми эталонами, для ее установления требуется проведение горнобуровых работ. Однако необходимость разбуривания по сети здесь отпадает, так как достаточно бурения одной скважины для каждого класса. Классификация геологических объектов без обучения приобретает особое значение при обработке данных рудной геологии и геофизики в сложно построенных в геологическом отношении районах, когда сложный и непостоянный характер геофизических полей обычно исключает возможность использования эталонных объектов.

Среди существующих алгоритмов классификации выделяются: *эвристические, использующие различные приемы суммирования (или произведения) признаков; корреляционные*, базирующиеся на аппарате компонентного и факторного анализа; статистические, основанные на проверке статистических гипотез.

Эвристические алгоритмы классификации. *Эвристические алгоритмы классификации* строятся в предположении независимости признаков и одинаковой их информативности. В одном из алгоритмов суммирования диапазон значений каждого признака разбивается на заданное и одинаковое по каждому признаку число градаций, например четыре. Результатом суммирования является число

$$y_i = \sum_{j=1}^L x_{li}^{(j)},$$

где $x_{li}^{(j)}$ – представляет j -й разряд l -го признака в i -й точке.

В тех же предположениях о признаках проводится суммирование полных нормированных градиентов, предварительно рассчитанных по методике Березкина для каждого признака в отдельности. Помимо равновесного суммирования признаков, когда равновесность, т.е. равная информативность, обеспечивается делением признаков на одно и то же число разрядов, используются алгоритмы неравновесного суммирования типа

$$y_l = \sum_{i=1}^L \frac{(x_{li} - \bar{x}_l)^2}{\sigma_l^2} \text{ или } y_l = \sum_{i=1}^L \frac{|x_{li} - \bar{x}_l|}{\sigma_l},$$

где $\bar{x}_l$ и σ_l^2 – соответственно среднее значение и дисперсия l -го признака. Для определения порога принятия решения при этом можно использовать критерий χ^2 .

При решении частных задач оказываются эффективными алгоритмы типа $y_i = x_{1i}x_{2i} / x_{3i}$, для которых не существует математического обоснования и соответственно невозможно оценить величину порога принятия решения.

Метод главных компонент. *Корреляционный алгоритм классификации* (метод главных компонент) позволяет провести классификацию геологических объектов по зависимым друг от друга признакам, а с помощью дисперсионного анализа оценить значимость вклада (веса) коэффициента) каждого признака. Чтобы свести задачу обработки данных комплекса признаков для случая разделения объектов на два класса к задаче компонентного анализа, сформулируем принцип классификации следующим образом. Пусть i -й объект относится к классу H_1 если выполняется условие

$$\sum_{l=1}^L h_l x_{li} = y_i > y_{\text{нор.}}, \quad (4.1)$$

где y_i – комплексный параметр, равный сумме взвешенных значений каждого признака для i -го объекта; h_l – весовой коэффициент l -го признака; $y_{\text{нор.}}$ – порог принятия решения.

Если в рассматриваемом множестве объектов существуют два класса и веса h_l выбраны так, что обеспечивают различные значения y_i для этих классов, то вычисление h_l можно свести к процедуре нахождения собственного вектора корреляционной матрицы R исходных признаков, соответствующего ее максимальному собственному числу λ_{max} .

Порог для принятия решения о наличии объекта искомого класса определяется на основе центральной предельной теоремы, согласно которой сумма случайных величин распределена нормально. Поэтому величина комплексного параметра y_i распределена нормально и порог $y_{\text{нор.}}$ определяется заданием вероятности γ .

$$\gamma = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y_{\text{пор}}} e^{-x^2/2} dx, \text{ где } x = (y_{\text{пор}} - \bar{y}) / \sigma_y \sqrt{2}.$$

Для вероятности $\gamma=95\%$ $y_{\text{пор}}=1,65$. Поскольку дисперсия величины y_i равна λ_{max} , для вероятности 95% порог будет $s_{\text{пор}}=1,65\sqrt{\lambda_{\text{max}}}$. Соответственно, для вероятности $\gamma=99\%$ $y_{\text{пор}}=2,4\sqrt{\lambda_{\text{max}}}$.

Последовательное вовлечение в обработку новых признаков дает возможность по скорости возрастания комплексного параметра y_i судить об их информативности.

Выражение (4.17) представляет собой первую главную компоненту, поэтому сам алгоритм реализует вычисление по методу главных компонент. В дальнейшем применение метода главных компонент может идти: путем последовательного нахождения первой главной компоненты отдельно для объектов со значениями $y_i < y_{\text{пор}}$ и $y_i > y_{\text{пор}}$ и т.д.; путем нахождения нескольких (двух-трех) главных компонент, определяющих основную часть дисперсии в пространстве признаков, и их геологического истолкования. Первая главная компонента выделяет объекты, характеризующиеся наибольшей энергией (дисперсией) в пространстве всех признаков. С целью геологического истолкования нередко оказывается полезной процедура получения нового набора главных компонент (факторов), т.е. вращения факторов.

Метод K -средних. Алгоритм классификации методом K -средних предложен Д. Мак-Куином и реализован А.Н. Кленчиным, его суть состоит в следующем. Пусть множество объектов n следует разбить на заданное число классов $M \ll n$, однородных в смысле заданной метрики классов. Каждый i -й объект описан вектором изучаемых признаков x_{li} ($l=1, \dots, n$). На первом этапе случайно выбирается k объектов из общей совокупности n объектов (точек наблюдений или ячеек) и в дальнейшем ведется последовательное уточнение этих выбранных, полагаемых эталонными, объектов $E^v(e_1^v, e_2^v, \dots, e_k^v)$ с соответствующим пересчетом приписываемых им весов $h_1^v, \dots, h_k^v$, где v – номер итерации, $v=0, 1, 2, \dots$. Первые случайно расположенные k -объекты являются нулевым приближением, т.е. $e_i^{(0)} = x_i$; $h_i^{(0)} = 1$; $i=1, \dots, k$. Затем извлекается объект (точка) x_{k+1} и выясняется, к какому из эталонов $e_i^{(0)}$ он оказался ближе всего. Этот эталон заменяется новым, определяемым как «центр тяжести» старого эталона и присоединенного к нему объекта x_{k+1} с увеличением на единицу соответствующего ему веса. Пересчет эталонов и соответствующих им весов на v -м шаге проводится согласно формулам

$$e_i^{(\nu)} = \frac{h_i^{(\nu)} e_i^{(\nu-1)} + x_{k+\nu}}{h_i^{(\nu)} + 1}, \quad h_i^{(\nu)} = h_i^{(\nu-1)} + 1$$

при условии, что среднеквадратическое расстояние между $x_{k+\nu}$ и $e_i^{(\nu-1)}$ минимально, т.е.

$$\rho(x_{k+\nu}, e_i^{(\nu-1)}) = \min \rho(x_{k+\nu}, e_j^{(\nu-1)}), j = 1, \dots, k,$$

и по формулам $e_i^{(\nu)} = h_i e^{(\nu-1)}$, $h_i^{(\nu)} = h_i^{(\nu-1)}$, если это условие не выполняется.

После уточнения эталонной совокупности проводится окончательная классификация множества объектов n по комплексу признаков с целью выделения M классов в соответствии с правилом минимального дистанционного разбиения относительно «центров тяжести» (эталонных объектов) классов, т.е.

$$S_l(E) = [\bar{X} : \rho(x, E_l) < \rho(x, E_j)], j = 1, \dots, k, j \neq 0.$$

Это разбиение – наилучшее в смысле функционала качества

$$Q(E) = \sum_{l=1}^L \sum_{x_l \in S} \rho^2(x_l, \bar{x}(l)),$$

который является взвешенной суммой внутриклассовых дисперсий.

Результаты классификации четырех геофизических полей, приведенных на рис. 4.1, по методу K -средних, иллюстрируются на рис. 4.2.

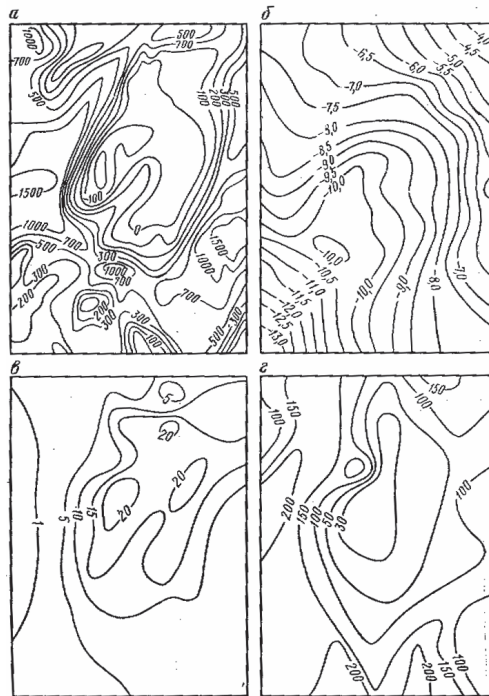


Рис. 4.1. Геофизические поля над Бенкалинским меднопорфировым месторождением: а) Z_a в nT , б) Δg в mG , в) η_k в %, г) ρ_k в Ом·м

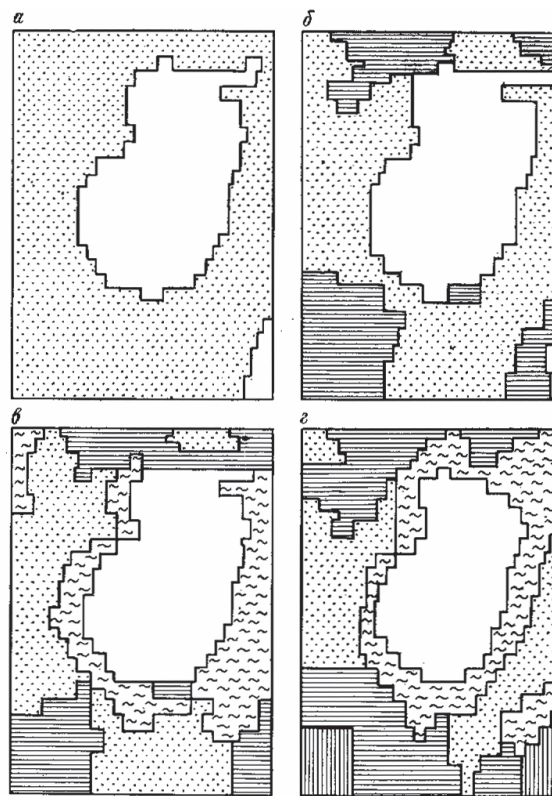


Рис. 4.2. Результат классификации геофизических полей методом K -средних на два: а) три, б) четыре, в) и пять, г) классов

В настоящее время существуют различные модификации метода K -средних, в частности, рассмотренные в учебном пособии [8].

ГЛАВА V. РАЗВИТИЕ МЕТОДА ГЛАВНЫХ КОМПОНЕНТ ПРИ НАЛИЧИИ КОРРЕЛЯЦИОННЫХ СВЯЗЕЙ МЕЖДУ ПОЛЯМИ ПО ПЛОЩАДИ НАБЛЮДЕНИЙ

Алгоритм многопризнаковой фильтрации геофизических полей

Комплексный анализ геополей и их атрибутов методом главных компонент используется практически во всех современных геоинформационных системах и технологиях [10]. Однако его реализация ограничивается расчетом первой главной компоненты по комплексу полей (признаков атрибутов) для каждого профиля сети наблюдений в отдельности. На рис. 5.1 приведен схематический пример определения первой главной компоненты для трех полей по профилю.

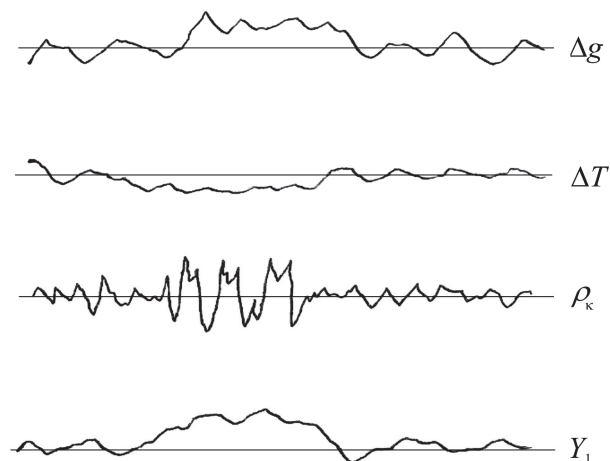


Рис. 5.1. Схематический пример определения первой главной компоненты для трех полей

Здесь гравитационное поле Δg характеризуется повышенными значениями, магнитное поле ΔT – пониженными значениями, а поле сопротивлений ρ_k отличается повышенной дисперсией над искомым объектом. Первая главная компонента Y_1 при этом выделяет область повышенной энергии в пространстве трех полей.

Метод главных компонент при комплексном анализе L -полей по сети наблюдений из $(n \times N)$ точек до сих пор остается нереализованным, т.е. иначе, отсутствует учет корреляционных связей между полями (признаками) по площади наблюдений.

Ниже приводится алгоритм многопризнаковой фильтрации полей (многопризнаковый анализ) по сети профилей, приведенный в [7].

Реализация этого алгоритма сводится к двухэтапной процедуре метода главных компонент. На первом этапе осуществляется вычисление первой главной компоненты по комплексу полей (признаков) для каждого профиля сети в отдельности, т.е. как при традиционном варианте метода.

На втором этапе проводится расчет первой главной компоненты для уже полученных первых главных компонент по каждому профилю, что позволяет учесть корреляционные связи между полями (признаками) по всей площади наблюдений. При этом значения собственного вектора для такой первой главной компоненты характеризуют «вес» каждого профиля в многопризнаковом пространстве, а сами значения первой главной компоненты определяют «вес» каждой точки (пикета) в многопризнаковом пространстве. Перемножение вектора-столбца значений первой главной компоненты на вектор-строку значений ее собственного вектора позволяет восстановить значения первой главной компоненты по всей площади наблюдений.

Последовательность операций при реализации предлагаемого алгоритма для данных исходных матриц L -полей по площади из N профилей и n точек по каждому профилю

$$\begin{pmatrix} x_{11k} & x_{21k} & \dots & x_{L1k} \\ x_{12k} & x_{22k} & \dots & x_{L2k} \\ \dots & \dots & \dots & \dots \\ x_{1nk} & x_{2nk} & \dots & x_{Lnk} \end{pmatrix}, k = 1, \dots, N \quad (5.1)$$

$$l = 1, l = 2, \dots, l = L$$

включает:

1. Расчет коэффициентов корреляции для L -признаков (или L -полей) по каждому из N профилей съемки:

$$\begin{pmatrix} 1 & r_{12} & \dots & r_{1L} \\ \dots & \dots & \dots & \dots \\ r_{L1} & r_{L2} & \dots & 1 \end{pmatrix}, k = 1, \dots, N. \quad (5.2)$$

2. Построение N -корреляционных матриц для каждого из N -профилей размерностью $L \times L$

$$\begin{pmatrix} 1 & r_{12} & \dots & r_{1L} \\ r_{21} & 1 & \dots & r_{2L} \\ \dots & \dots & \dots & \dots \\ r_{L1} & r_{L2} & \dots & 1 \end{pmatrix} (k), k = 1, \dots, N \quad (5.3)$$

3. Вычисление для каждой k -й матрицы ($k = 1, \dots, N$) (5.3) максимального собственного значения $\lambda_{\max}(k)$, характеризующего энергетическое отношение сигнал/помеха в пространстве L -признаков (или L -полей) для каждого профиля.

4. Нахождение значений собственных векторов, соответствующих $\lambda_{\max}(k)$, для каждой из N -матриц (5.3):

$$(a_{11}, a_{12}, \dots, a_{1L})(k), k = 1, \dots, N, \quad (5.4)$$

которые характеризуют «вес» каждого признака (поля) в наблюдаемые значения признаков (полей) при их нормировке $\sum_{l=1}^L a_{1l}^2 = 1$.

5. Расчет значений первой главной компоненты для каждого профиля:

$$Y_{1k} = \sum_{l=1}^L a_{1l} X_{lk}, k = 1, \dots, N,$$

соответствующих $\lambda_{\max}$:

$$\begin{pmatrix} x_{11} & x_{21} & \dots & x_{L1} \\ x_{12} & x_{22} & \dots & x_{L2} \\ \dots & \dots & \dots & \dots \\ x_{1n} & x_{2n} & \dots & x_{Ln} \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{12} \\ \dots \\ a_{1L} \end{pmatrix} = \begin{pmatrix} Y_{11} \\ Y_{21} \\ \dots \\ Y_{n1} \end{pmatrix} (k), k = 1, \dots, N. \quad (5.5)$$

Значения $Y_{il(k)}$ характеризуют весовые коэффициенты для каждого i -го пикета, т.е. получаем N -векторов $(Y_{i11}, Y_{i12}, \dots, Y_{i1N})$, каждый размерностью n .

Причем первый индекс указывает на изменения по точкам (пикетам). Второй индекс подчеркивает первую главную компоненту, а третий индекс указывает на номер профиля. Следовательно, в результате вычислений по формуле (5) реализуется построение матрицы для N первых главных компонент размерностью $n \times N$, т.е.

$$\begin{pmatrix} Y_{111} & Y_{112} & \dots & Y_{11N} \\ Y_{211} & Y_{212} & \dots & Y_{21N} \\ \dots & \dots & \dots & \dots \\ Y_{n11} & Y_{n12} & \dots & Y_{n1N} \end{pmatrix}. \quad (5.6)$$

6. Расчет коэффициентов корреляции между всевозможными парами первых главных компонент R_{ij} и построение по этим коэффициентам корреляционной матрицы R размерностью $N \times N$.

$$[R] = \begin{bmatrix} 1 & R_{12} & \dots & R_{1N} \\ R_{21} & 1 & \dots & R_{2N} \\ \dots & \dots & \dots & \dots \\ R_{N1} & R_{N2} & \dots & 1 \end{bmatrix}, R_{ij} = R_{ji}. \quad (5.7)$$

7. Вычисление для матрицы (7) максимального собственного значения $\lambda_{\max}$, характеризующего энергетическое отношение сигнал/помеха уже в многопризнаковом пространстве по всей площади наблюдения.

8. Нахождение значений собственного вектора, соответствующего $\lambda_{\max}$ для матрицы (7) ($b_{11}, b_{12}, \dots, b_{1N}$) из системы уравнений (5.8):

$$\begin{bmatrix} 1 - \lambda_{\max} & R_{12} & \dots & R_{1N} \\ R_{21} & 1 - \lambda_{\max} & \dots & R_{2N} \\ \dots & \dots & \dots & \dots \\ R_{N1} & R_{N2} & \dots & 1 - \lambda_{\max} \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{12} \\ \dots \\ b_{1N} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \dots \\ 0 \end{bmatrix}. \quad (5.8)$$

При этом используется нормировка $\sum_{k=1}^N b_{1k}^2 = 1$, которая обеспечивает определенность решения системы линейных уравнений (8) при равенстве всех уравнений в правой их части нулю.

Значения вектора ($b_{11}, b_{12}, \dots, b_{1N}$) характеризуют «весовые» коэффициенты каждого профиля в многопризнаковом пространстве.

9. Нахождение значений первой главной компоненты для совокупности первых главных компонент, рассчитанных по каждому из N -профилей съемок путем умножения вектора-столбца ($b_{11}, b_{12}, \dots, b_{1N}$) на матрицу (5.6):

$$\begin{bmatrix} Y_{111} & Y_{112} & \dots & Y_{11N} \\ Y_{211} & Y_{212} & \dots & Y_{21N} \\ \dots & \dots & \dots & \dots \\ Y_{n11} & Y_{n12} & \dots & Y_{n1N} \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{12} \\ \dots \\ b_{1N} \end{bmatrix} = \begin{bmatrix} z_{11} \\ z_{21} \\ \dots \\ z_{n1} \end{bmatrix}, \quad (5.9)$$

где значения $z_{11}, z_{21}, \dots, z_{n1}$ характеризуют «весовые» коэффициенты для каждого пикета в многопризнаковом пространстве.

10. Расчет комплексного параметра для L -признаков (полей) по сети $N \times n$ точек путем перемножения вектора-столбца ($z_{11}, z_{21}, \dots, z_{n1}$) на вектор-строку ($b_{11}, b_{12}, \dots, b_{1N}$), что обеспечивает восстановление значений комплексного параметра по всей сети наблюдений, т.е.:

$$\begin{pmatrix} z_{11} \\ z_{21} \\ \dots \\ z_{n1} \end{pmatrix} (b_{11} \quad b_{21} \quad \dots \quad b_{1N}) = \begin{pmatrix} z_{11}b_{11} & z_{11}b_{12} & \dots & z_{n1}b_{1N} \\ z_{21}b_{11} & z_{21}b_{12} & \dots & z_{n1}b_{1N} \\ \dots & \dots & \dots & \dots \\ z_{n1}b_{11} & z_{n1}b_{12} & \dots & z_{n1}b_{1N} \end{pmatrix}. \quad (5.10)$$

При этом для каждого столбца матрицы (5.10) следует добавить среднее значение каждой первой компоненты, поскольку при расчете коэффициентов корреляции в матрице (5.7) средние значения главных компонент вычитались.

Адаптивный вариант многопризнаковой фильтрации реализуется на основе применения описанных выше процедур (5.1) – (5.10) в скользящем окне из $(N \times n)$ точек путем его перемещения на один дискрет по X , а затем на один дискрет по Y . Рассмотренный выше алгоритм многопризнаковой фильтрации (5.1) – (5.10) может быть использован при анализе комплекса сейсмических атрибутов по временному или глубинному разрезу. Размер скользящего окна определяется по изолинии ДАКФ, равной 0,1 – 0,2 максимального значения ДАКФ.

При комплексном анализе сейсмических атрибутов по временному разрезу выбор размера скользящего окна ориентируется на предполагаемые размеры нефтегазовой залежи в выделенном интервале целевого горизонта. Так, например, для детальной сейсмической съемки масштаба 1:25 000 (1:50 000) при расстоянии между сейсмоприемниками в 25 м и при интервале целевого горизонта в 0,1 с с целью обнаружения залежи углеводородов размерами до 1 км по простиранию размер окна составит 40 точек по оси x ($n = 40$) и 50 точек по оси t ($N = 50$) т.е. размеры матриц (1) будут 40×50 для каждого L -атрибута (признака).

Выделение искомой неоднородности, характеризующейся максимальным энергетическим отношением сигнал/помеха в пространстве L -полей, на поверхности земли (или залежи углеводородов в плоскости временного размера по L -сейсмическим атрибутам) осуществляется по значению изолинии первой главной компоненты для матрицы (10), превышающему величину $2\sqrt{\lambda_{\max}}$.

В качестве примера приведем результат многопризнаковой фильтрации по шести атрибутам волнового поля, рассчитанным в целевом интервале временного разреза размером 0,2 сек (100 точек по времени) и 4 км (80 точек по оси X). Два из этих атрибутов изображены на рис. 5.2 (а, б), а результат многопризнаковой фильтрации – на рис. 5.2 (в). По максимальным значениям четко выделяется местоположение нефтегазовой залежи.

В качестве другого примера использования алгоритма многознаковой фильтрации в работе [7] приведено выделение субгоризонтальных границ земной коры по региональному профилю 3DB по комплексу сейсмических атрибутов.

Заметим, что в пределах больших территорий корреляционные свойства полей существенно изменяются. Это приводит к необходимости использования адаптивного варианта алгоритма многопризнаковой фильтрации. Скорости современных ЭВМ обеспечивают возможность реализации адаптивных алгоритмов только по площади съемки, но в объемном (3D) варианте.

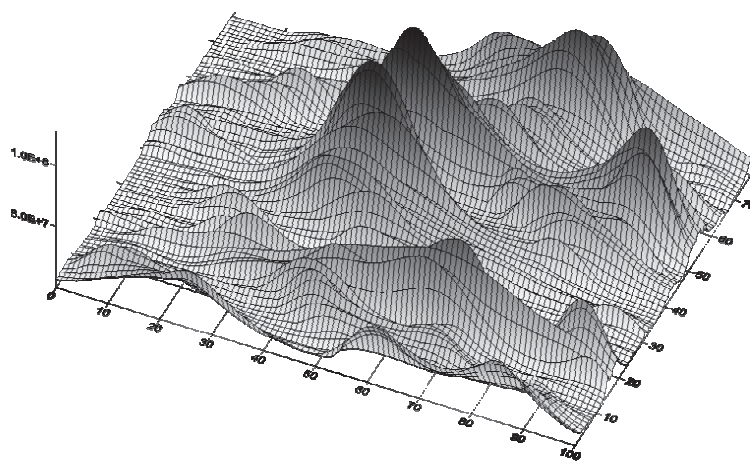


Рис. 5.2, а. Атрибут энергия волнового поля

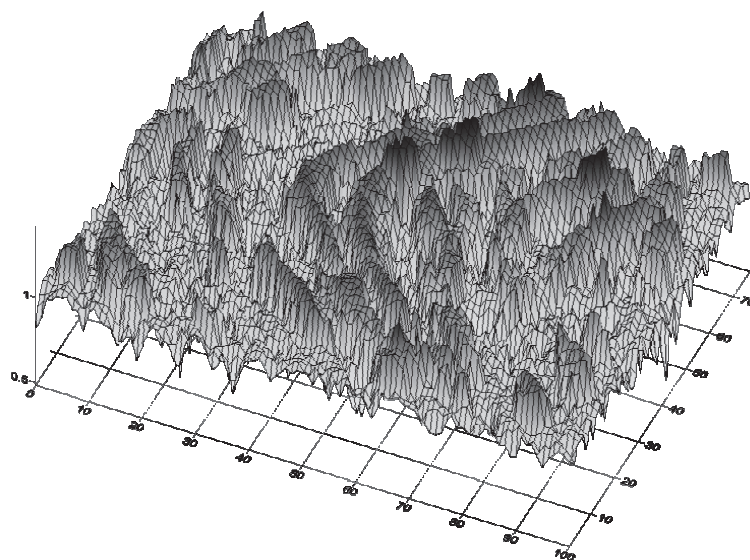


Рис. 5.2, б. Атрибут энтропия волнового поля

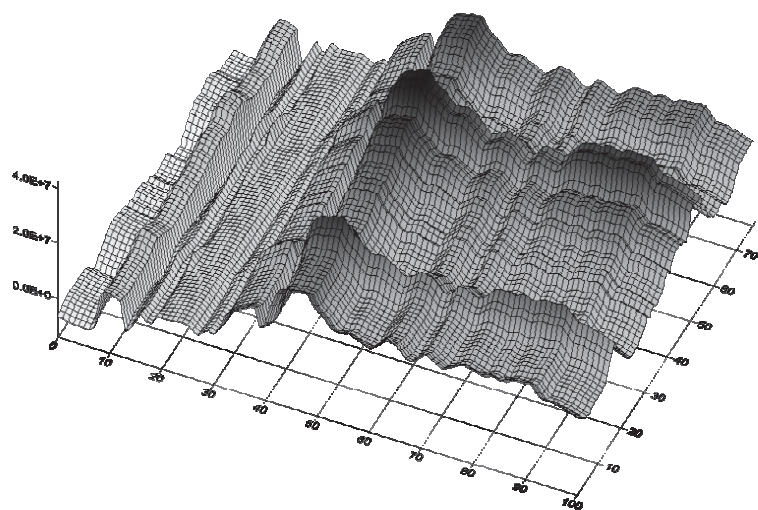


Рис. 5.2, в. Результат многопризнаковой фильтрации

ГЛАВА VI. МАТЕМАТИЧЕСКАЯ ПОСТАНОВКА ЗАДАЧ КОЛИЧЕСТВЕННОЙ КОМПЛЕКСНОЙ ИНТЕРПРЕТАЦИИ

Математическая постановка задач количественной комплексной интерпретации по определению геометрических и физических параметров $\vec{\theta}_l$, $l = 1, \dots, L$ – число полей, реализуется как при детерминированном подходе, так и при вероятностно-статистическом.

При детерминированном подходе наблюдаемые данные считаются точно измеренными, без помех, и решение задачи дается либо в виде числа, либо в виде функции.

При вероятностно-статистическом подходе наблюдаемые данные считаются измеренными неточно, неполно, искажены помехами. При этом решение задачи дается в виде распределения вероятностей над возможным значением результата по оценке того или иного параметра.

Основной задачей количественной комплексной интерпретации является построение согласованных *физико-геологических моделей* (ФГМ). Согласованная ФГМ представляет распределение набора физических параметров (свойств), осредненных на единой геометрии слоев и блоков в плоскости геологического разреза (2D-модель) или в объеме геологической среды (3D-модель), построенных в пределах погрешностей наблюдений.

Комплексная интерпретация при детерминированном подходе.

При детерминированном подходе количественная комплексная интерпретация сводится к минимизации функционалов для каждого поля, представляющих квадрат разности наблюдаемых (измеренных) полей и полей, рассчитанных для соответствующих моделей разреза (среды), т.е.

$$\sum_{i=1}^n \left[(f_{li}^{набл} - f_{li}(\vec{\theta}_l)) \right]^2 \Rightarrow \min, \quad (6.1)$$

$f_{li}^{набл.}$ – наблюдаемое значение l -го поля в i -й точке; $f_{li}(\vec{\theta}_l)$ – значение модельного поля в i -й точке для вектора параметров $\vec{\theta}_l$, включающего геометрические и физические параметры.

Достижение минимума такого функционала даже для двух полей представляет непростую задачу, поскольку требуется найти не локальный, а глобальный минимум.

В связи с этим практическая реализация количественной комплексной интерпретации для двух полей, например, сейсморазведки и гравиразведки, сводится сначала к определению геометрии слоев (субгоризонтальных

границ раздела) по данным сейсморазведки. Затем для выделенных слоев осуществляется разбиение их на блоки с заданными значениями плотностей. Плотности оцениваются по скоростям, либо путем линейных регрессионных зависимостей, либо по формуле Гарднера $\sigma_{пл} = 0,31V^{1/4}$, где V в м/с, а плотность σ в г/см³.

Для полученного таким образом плотностного разреза решается прямая задача гравиразведки. При совпадении модельной кривой Δg с наблюдаемым полем в пределах погрешности наблюдений (например, для масштаба 1:50 000 – это 0,5 мГал) можно считать, что задача построения согласованной по плотности и скорости модели решена. При отсутствии подобного согласования переходят к переинтерпретации данных сейсморазведки, на основании результатов которой вновь задается плотностная модель и для нее решается прямая задача гравиразведки.

Поскольку решение любой обратной задачи геофизики сводится к решению системы линейных уравнений, а сама задача является, как правило, некорректно поставленной из-за наличия помех, в частности при решении задач гравиразведки и магниторазведки, осуществляется регуляризация. При наличии некоррелированных помех параметр регуляризации представляет просто дисперсию помех, которая добавляется к диагональным членам матрицы, полученной при решении системы линейных уравнений. Когда решение обратной задачи по данным комплекса двух методов не может быть достигнуто с необходимой погрешностью, используются данные других методов.

В последнее время наиболее эффективным средством решения задачи количественной комплексной интерпретации считается использование генетических алгоритмов. При наличии современных высокоскоростных ЭВМ с их помощью даже для десятка полей (при геофизических исследованиях скважин) в течение нескольких часов удается достичь глобального минимума при определении параметров геологического разреза.

Заметим, что Г.Я. Голиздра в работе [2] при минимизации L -функционалов в (6.1) предложил введение весовых множителей h_i , что в целом не приводит к желаемому результату.

Им же введено понятие о трех основных моделях среды:

Комплексная модель типа S не предполагает зависимости между физическими и геометрическими параметрами для разных геофизических методов. Комплексная модель получается путем простого объединения геометрических и физических параметров, полученных для источников аномалий отдельно взятых методов.

Модель типа U предполагает полное совпадение геометрических и физических параметров для разных методов. В этом случае можно ограничиться интерпретацией лишь одного метода.

Наиболее распространенной является модель типа M , занимающей промежуточное положение между моделями S и U . При этом требуется учет корреляционных связей параметров разных методов.

Комплексная интерпретация при вероятностно-статистическом подходе определяется с позиции теории статистических решений. Впервые ее постановка была сформулирована Ф.М. Гольцманом [3].

По сравнению с моделью (6.1) аномальная составляющая определена вектором параметров $\vec{\theta}_l$ каждого l -го поля с учетом наличия помех, т.е. $\vec{f}_{li} = \vec{a}_l(\vec{\theta}_l) + \vec{n}_l, l = 1, \dots, L$.

Комплексная интерпретация с позиции теории статистических решений. Методы теории статистических решений для количественной интерпретации были развиты Ф.М. Гольцманом и Т.Б. Калининой. Модель количественной интерпретации представляется следующим образом. Наблюдаемое l -е поле $\vec{f}_l$ является суммой полезной аномалии $\vec{a}_l$ и помехи $\vec{n}_l$: $\vec{f}_l = \vec{a}_l(\vec{\theta}_l) + \vec{n}_l$; $l = 1, \dots, L$. Аномальное поле $\vec{a}_l(\vec{\theta}_l)$ считается заданным из решения прямой задачи и определяется вектором параметров $\vec{\theta}_l(\theta_{l1}, \dots, \theta_{ls})$, где s – число параметров, определяемых по полю $\vec{f}_l$.

Решение задачи интерпретации при этом принимается на основе оценки вектора $\vec{\theta}$ методом максимального правдоподобия. Для случая некоррелированных и нормально распределенных помех функция правдоподобия равна:

$$P(\vec{f}_1, \dots, \vec{f}_L / \vec{\theta}) = \prod_{l=1}^L (2\pi\sigma_l^2)^{-\frac{n_l}{2}} \exp \left\{ -\frac{1}{2} \sum_{l=1}^L \frac{1}{\sigma_l^2} \sum_{i=1}^{n_l} [f_{li} - a_{li}(\vec{\theta})]^2 \right\}, \quad (6.2)$$

где n_l – число значений l -го поля; σ_l^2 – дисперсия помехи l -го поля.

Для практических расчетов достаточно ограничиться функцией

$$\varphi(\vec{\theta}) = -\sum_{l=1}^L \frac{1}{\sigma_l^2} \sum_{i=1}^{n_l} [f_{li} - a_{li}(\vec{\theta})]^2. \quad (6.3)$$

Максимум $\varphi(\vec{\theta})$ дает те же оценки $\vec{\theta}$, что и максимум функции правдоподобия. Из формулы (6.3) следует, что модельные аномальные эффекты a_{li} для разных полей одновременно зависят от одних и тех же параметров $\vec{\theta}_l$. Это означает, что при определении вектора $\vec{\theta}$ экстремум по l

не совпадает с экстремумами по отдельным составляющим в пространстве $\vec{\theta}$ и тем самым количественная интерпретация данных комплекса геофизических методов принципиально отличается от количественной интерпретации раздельного определения параметров по каждому полю f_i .

В качестве примера рассмотрим построение функции $\varphi(\vec{\theta})$ для магнитного f_{1i} и гравитационного f_{2i} аномальных полей от шара, полученных в одних и тех же точках $x_i = i\Delta x$; $i = 1, \dots, n$ прямолинейного профиля наблюдений при наличии помех.

Модели экспериментальных полей имеют вид

$$f_{1i} = \mu \frac{2h_1^2 - x_i^2}{(h_1^2 + x_i^2)^{3/2}} + n_{1i}; \quad f_{2i} = \nu \frac{h_2}{(h_2^2 + x_i^2)^{3/2}} + n_{2i}, \quad (6.4)$$

где μ – магнитный момент шара; h_1 и h_2 – глубины залегания его центра; ν – произведение гравитационной постоянной на избыточную плотность шара; n_{1i} и n_{2i} – помехи, осложняющие аномальные поля, принимаемые нормально распределенными, стационарными и некоррелированными по x , с известными дисперсиями σ_1^2 и σ_2^2 .

Примем также, что $h_1 = h_2 = h$. Для упрощения будем считать величину ν заранее известной. Тогда $\mu = \theta_1$; $h_1 = h = \theta_2$; $h_2 = h = \theta_2$.

Искомymi параметрами, таким образом, будут $\theta_1 = \mu$ и $\theta_2 = h$. Функция $\varphi(\theta_1, \theta_2)$ равна

$$\varphi(\theta_1, \theta_2) = \varphi(\mu, h) = \frac{1}{\sigma_1^2} \sum_{i=1}^n \left[f_{1i} - \frac{\mu(2h^2 - x_i^2)}{(h^2 + x_i^2)^{3/2}} \right]^2 - \frac{1}{\sigma_2^2} \sum_{i=1}^n \left[f_{2i} - \frac{\nu h}{(h^2 + x_i^2)^{3/2}} \right]^2. \quad (6.5)$$

Максимумы $\varphi(\theta_1, \theta_2)$ при изменении μ и h дают искомые оценки неизвестных параметров.

Решение задачи оценки даже для случая лишь двух параметров по комплексу двух полей представляет непростую математическую проблему. В то же время в случае весьма простой модели (наклонный уступ, пласт) число параметров достигает пяти-шести. Специальные исследования, проведенные Т.Б. Калининой, показывают, что *при одновременном определении всех параметров требуется весьма высокая точность исходных данных и корректность модели*. Чтобы получить эффективные оценки нескольких параметров одновременно, нужно иметь много исходной информации. При ее недостатке погрешность определения отдельных параметров достигает десятков и даже сотен процентов. *Геологический разрез средней сложности, включающий несколько структурных границ, тектонических нарушений,*

блоков, тел и т.п. *описывается физико-геологической моделью, которая характеризуется не менее, чем несколькими десятками геометрических и физических параметров.* По мере усложнения разреза число параметров увеличивается, поэтому необходимо использование многих независимых источников информации, по-разному освещающих элементы геологического строения, т.е. *необходимость в комплексировании различных методов и комплексной интерпретации геолого-геофизических данных совершенно очевидна.*

При недостатке информации число искомых параметров следует уменьшать путем упрощения физико-геологической модели. Это ведет к более грубому, но более надежному решению. Кроме того, почти всегда имеются параметры, медленно изменяющиеся по площади исследования (например, скорость распространения упругих волн в некоторой толще, плотность изверженных пород определенной формации). Опыт предшествующих работ обычно избавляет от необходимости определения этих параметров. А вновь поступающую информацию целесообразно направить на оценку параметров модели, отличающихся большей изменчивостью.

ГЛАВА VII. КРИТЕРИИ АДЕКВАТНОСТИ ФИЗИКО-ГЕОЛОГИЧЕСКИХ МОДЕЛЕЙ

Критерии адекватности построенных ФГМ реальной среде определяют качество моделей.

При детерминированном подходе в качестве критерия качества служит величина невязки между наблюдаемыми значениями поля и значениями, полученными путем решения прямой задачи от модели. Величина невязки определяется суммой квадратов разности наблюдаемых и рассчитанных полей. Сама величина невязки непосредственно связана с масштабом исследований.

В последних рекомендациях по гравиразведке, изданных Мингео СССР в 1984 г. для масштаба 1:200 000 допустимая величина невязки составляет 2 мГал. Учитывая возросшую точность, в частности при аэрогравиметрических наблюдениях величина невязки составляет 0,5 мГал. Соответственно при более детальных исследованиях величина невязки существенно снижается. Для аэрогравиметрических наблюдений масштаба 1:50 000 она составляет 0,2 мГал, а по старым рекомендациям Мингео СССР – 0,5 мГал. При современных аэромагнитных наблюдениях величина невязки практически любого масштаба составляет 1-2 нТл.

При комплексировании геофизических методов достигается повышение однозначности решения обратных задач, а следовательно, и качества (адекватности) построенных согласованных ФГМ реальной среде.

Кроме того, необходимость количественной комплексной интерпретации диктуется теми соображениями, что блоковые и слоистые модели среды в разных полях отображаются по-разному.

Количественная комплексная интерпретация реализуется лишь для достаточно интенсивных аномальных полей, т.е. при больших отношениях сигнал/помеха (при малых отношениях сигнал/помеха решаются в основном задачи обнаружения сигналов по установлению факта их наличия). При этом, в отличие от методов распознавания образов и классификации по комплексу признаков, количественная интерпретация невозможна без аппарата математической физики, теории полей. В то же время методы регрессионного анализа, теории статистических решений позволяют находить взаимные корреляционные связи между изучаемыми элементами геологического строения (физическими и геометрическими параметрами) и на основе таких связей осуществлять прогноз наиболее вероятных параметров искомых объектов.

Оценка адекватности комплексных ФГМ реальным объектам.

Необходимость комплексной интерпретации диктуется теми соображениями, что решение, т.е. модель, грубо согласующееся с результатами нескольких геофизических методов, является предпочтительнее решения, которое хорошо удовлетворяет результатам интерпретации лишь одного какого-то метода и в то же время не удовлетворяет результатам других методов. Практически всегда для данных наблюдений одного метода можно подобрать модель, параметры которой при решении прямой задачи обеспечат совпадение расчетов с результатами эксперимента с высокой точностью. Однако решение прямой задачи для параметров этой же модели для другого метода может резко отличаться от его экспериментальных данных.

Кроме того, разные геологи придерживаются разных гипотез или концепций о геологическом строении конкретного региона (например, слоистая или блоково-слоистая модели), поэтому возникает задача оценки адекватности комплексной ФГМ той или иной геологической концепции. Учитывая изложенное, в качестве критерия оценки адекватности комплексной ФГМ реальным объектам может быть использована величина обобщенного расстояния μ , равная

$$\mu = \sum_l^L \sum_i^n h_l [f_{li}^M(\vec{\Theta}_K) - f_{li}^H]^2, \quad (7.1)$$

где f_{li}^H – наблюдаемое значение l -го поля в i -й точке; $f_{li}^M(\vec{\Theta}_K)$ – теоретически рассчитанное (модельное) значение l -го поля в i -й точке для вектора параметров объекта (среды) $\vec{\Theta}_K$, под которыми понимаются как петрофизические, так и геометрические параметры; h_l – весовые множители для l -го поля.

Весовые множители определяются координатами собственного вектора корреляционной матрицы, построенной по коэффициентам корреляции отклонений $f_{li}^M - f_{li}^H$ данных для любой пары модельного и наблюдаемого полей.

Координаты этого собственного вектора рассчитываются для максимального собственного значения $\lambda_{\max}$ корреляционной матрицы. Именно процедура нахождения весовых коэффициентов только для $\lambda_{\max}$ обеспечивает подбор модели, хотя и грубо, но согласующейся с наблюдениями по всем методам.

С помощью обобщенного расстояния μ оценивается адекватность различных физико-геологических моделей реальной ситуации и разным геологическим построениям. За адекватную модель принимается модель, для которой величина обобщенного расстояния будет наименьшей. В качестве примера рассмотрим табл. 7.1, в которой приведены значения обобщенного

расстояния, полученные для наблюдаемых и теоретических годографов, а также аномалий силы тяжести по трем вариантам геологического разреза. Разрез характеризует строение крупной синклинали, выполненной эффузивно-осадочными породами. Если использовать только сейсмические данные, то следует принять вариант 2 разреза с величиной обобщенного расстояния 1,2. Однако этот вариант по совокупности сейсмических и гравиметрических данных наименее приемлем (обобщенное расстояние 5,43). Небольшое ухудшение сходимости сейсмических данных позволяет улучшить общую сходимость и выбрать вариант 3 с расстоянием 1,78.

Таблица 7.1

Оценка близости вариантов интерпретации наблюдаемых данных по величине обобщенного расстояния (по И.Г. Клушину)

Геофизические данные	Варианты разреза структуры		
	1	2	3
	Обобщенное расстояние		
Только сейсмические	1,66	1,20	1,29
Сейсмические и гравиметрические	1,88	5,43	1,78

При комплексной интерпретации данных геофизических методов скважин (ГИС) на основе регрессионных нейронных сетей критерием качества является величина множественного коэффициента корреляции [8], определяемая по информативной совокупности методов. Информативная совокупность методов ГИС реализуется на эталонных объектах, в частности, для пластов, содержащих углеводороды. Для модели линейной регрессии, включающей разные методы ГИС, рассчитывается величина множественного коэффициента регрессии, и при достижении стабильности этого коэффициента, устанавливается информативная совокупность методов, отражающая качество построенной модели – регрессионная зависимость нефтегазосодержания от набора наиболее информативных методов.

Оценка адекватности с позиции регрессионного анализа.

Множественная регрессия используется при исследовании связи между несколькими (тремя и более) величинами. Линейная множественная регрессия y на x_1 и x_2 имеет вид:

$$y_i = a_0 + a_1 x_{1i} + a_2 x_{2i}.$$

Система нормальных уравнений для нахождения коэффициентов a_0, a_1, a_2 , в матричной форме записывается следующим образом [8]:

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \end{bmatrix} \begin{bmatrix} 1 & x_{11} & x_{21} \\ 1 & x_{12} & x_{22} \\ \dots & \dots & \dots \\ 1 & x_{1n} & x_{2n} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix}.$$

После перемножения матриц получаем:

$$\begin{bmatrix} n & \sum x_1 & \sum x_2 \\ \sum x_1 & \sum x_1^2 & \sum x_1 x_2 \\ \sum x_2 & \sum x_1 x_2 & \sum x_2^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum y \\ \sum x_1 y \\ \sum x_2 y \end{bmatrix};$$

$$a_0 n + a_1 \sum x_1 + a_2 \sum x_2 = \sum y;$$

$$a_0 \sum x_1 + a_1 \sum x_1^2 + a_2 \sum x_1 x_2 = \sum x_1 y;$$

$$a_0 \sum x_2 + a_1 \sum x_1 x_2 + a_2 \sum x_2^2 = \sum x_2 y.$$

Теснота связи оценивается по величине множественного коэффициента корреляции R , которая определяется из выражения:

$$R_{k+1} = \sqrt{\sum_{i=1}^k b_i r_{iy}},$$

где r_{iy} – коэффициент корреляции величины x_i с y , а значения коэффициентов b_i находятся из решения системы линейных уравнений вида $R\vec{b} = \vec{r}$. В этой системе R – корреляционная матрица, составленная из коэффициентов парной корреляции r_{ij} между x_i и x_j . Для линейной регрессии $y_i = a_0 + a_1 x_{1i} + a_2 x_{2i}$ матрица $R = \begin{bmatrix} 1 & r_{12} \\ r_{21} & 1 \end{bmatrix}$. Тогда уравнение $R\vec{b} = \vec{r}$ примет вид

$$R = \begin{bmatrix} 1 & r_{12} \\ r_{21} & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} r_{1y} \\ r_{2y} \end{bmatrix},$$

из которого получаем:

$$b_1 + b_2 r_{12} = r_{1y},$$

$$b_1 r_{12} + b_2 = r_{2y}.$$

Из последней системы находим коэффициенты:

$$b_1 = \frac{r_{1y} - r_{12} r_{2y}}{1 - r_{12}^2}; \quad b_2 = \frac{r_{2y} - r_{12} r_{1y}}{1 - r_{12}^2}.$$

Следовательно, множественный коэффициент корреляции:

$$R_3 = \sqrt{b_1 r_{1y} + b_2 r_{2y}} = \sqrt{\frac{r_{12}^2 - 2r_{12}r_{1y}r_{2y} + r_{2y}^2}{1 - r_{12}^2}}.$$

Величина множественного коэффициента регрессии по-существу представляет оценку качества (адекватности) построенных моделей.

Множественная линейная регрессия успешно используется для определения глубины залегания складчатого фундамента H_i по данным гравirazведки (Δg_i) и магниторазведки ΔT_i в виде: $H_i = a_0 + a_1 \Delta g_i + a_2 \Delta T_i$.

Множественная регрессия широко применяется для построения петрофизических моделей продуктивных залежей углеводородов по данным геофизических исследований скважин. Так, для месторождений Южного Мангышлака использование данных о сопротивлении ρ_k по значениям каротажа КС, интенсивности нейтронного гамма каротажа $I_{НГК}$ и скорости v_p по акустическому каротажу приводит к регрессии вида:

$\sigma_{плотнос.} = 2,28 + 0,28 \lg \rho_k + 0,36 I_{НГК} - 0,017 v_p$ при коэффициенте множественной корреляции $R=0,75$.

Множественная регрессия служит основой построения нейронных сетей регрессионного типа. Входом сети являются данные различных методов каротажа (атрибутов) или данные атрибутов сейсмического (временного) разреза. Промежуточный слой сети представляет совокупность информативных атрибутов, отобранных путем перебора сочетаний различных исходных атрибутов по критерию достижения множественного коэффициента корреляции его заданного значения. Выходом нейронной сети является комплексный параметр, в качестве которого при анализе, например, данных каротажа могут быть такие физические параметры, как плотность и т.д. в виде множественной регрессии, построенной по информативной совокупности атрибутов.

В общем случае следует иметь в виду, что качество количественной интерпретации существенно зависит от корреляционной связи между определяемыми параметрами разреза (среды). Мерой сходства параметров являются коэффициенты корреляции. Если коэффициент корреляции между параметрами близок к нулю, то корреляционная связь мала и обеспечиваются наиболее легкие условия определения параметров.

Если коэффициент корреляции близок к единице, т.е. корреляционная связь велика, то совместное определение параметров путем учета их различного влияния на наблюдаемое поле представляет трудно решаемую задачу. Такое распределение касается, прежде всего, определения значений параметров по одному полю, например, размеров объектов и их физических свойств.

Например, для гравитационного поля совместное определение глубины, размеров и физических свойств объектов могут достигать сотни процентов. Именно в таких случаях подчеркивается решающая роль комплексной интерпретации данных разных методов.

Совершенно очевидно, что при оценке глубины залегания субгоризонтальных границ определяющую роль для повышения качества моделей играют данные сейсморазведки и методы электромагнитного зондирования, а при оценке блокового строения моделей – данные гравиразведки, магнито-разведки и электропрофилирования.

ГЛАВА VIII. КОМПЛЕКСНАЯ ИНТЕРПРЕТАЦИЯ ГЕОФИЗИЧЕСКИХ ПОЛЕЙ ПРИ ПОСТРОЕНИИ ГЛУБИННЫХ МОДЕЛЕЙ ЗЕМНОЙ КОРЫ

С целью изучения глубинного строения земной коры проводятся комплексные геофизические исследования методами КМПВ-ГСЗ, МТЗ, гравиразведки и аэромагниторазведки масштаба 1:200 000 по региональным и опорным (геотраверсам) профилям. Опорные профили включают хотя бы одну сверхглубокую скважину.

К настоящему времени геофизические исследования по региональным и опорным профилям составляют сотни тысяч погонных километров. Общая схема расположения этих профилей приведена на рис. 8.1.

Электромагнитные зондирования МТЗ обычно проводятся по профилям с интервалом 20-50 км.

Гравитационные и аэромагнитные съемки масштаба 1:200 000 имеются по всей территории РФ, а в обе стороны от региональных профилей на расстоянии 50 км осуществляются более детальные съемки масштаба 1:50 000. Таким образом, комплекс геофизических методов КМПВ-ГСЗ, МТЗ, гравиразведка и магниторазведка обеспечивает возможность построения согласованных по основным физическим параметрам физико-геологических моделей на всю мощность земной коры и верхов мантии. Комплексный анализ и комплексная интерпретация геофизических полей в особенности гравимагнитных наблюдений проводятся несколько поразному в разных организациях, принимающих участие в построении моделей земной коры. Такими организациями являются ВСЕГЕИ, ВНИИгеосистем, ВНИИгеофизика, СНИИГИМС и др. В то же время для сопоставления моделей земной коры, полученных в разных организациях, следует придерживаться достаточно стандартного графа обработки и интерпретации геофизических полей.

Исследования ГСЗ, в частности проведенные в 70-е годы XX века по инициативе чл.-корр. АН СССР В.В. Федынского с использованием ядерных взрывов общей протяженностью региональных профилей 20 тыс. погонных километров, позволяют обеспечить построение моделей земной коры на глубине границы поверхностей Мохоровичича (~40 км), К-1 (~20 км) и К-2 (~10 км) и кристаллического фундамента. Эти границы и их скоростные характеристики обеспечивают четырехслойную геометрию земной коры. Результаты МТЗ, обычно до глубин 60-80 км, дополняют скоростные границы значениями проводимости. Однако границы ГСЗ и МТЗ могут не совпадать. При этом наиболее информативными являются участки профилей, на которых



Рис. 8.1. Схема расположения государственной сети опорных геофизических профилей России

отмечается их совпадение, а расхождение границ требует истолкования их геологической природы.

В целом результаты КМПВ (изучение фундамента), ГСЗ и МТЗ обеспечивают геометризацию земной коры с выделением крупных разломов в земной коре.

Применение метода ОГТ требует больших разносов, что на практике не всегда достижимо, и глубинность исследований ограничивается до 20-30 км.

По результатам МОГТ строятся сейсмоэнергетические разрезы, обеспечивающие выделение зон пониженной и повышенной рефлексивности отраженных волн. Пониженные зоны рефлексивности на глубинах 10-15 км представляют один из критериев регионального прогноза нефтегазоносности. По результатам МТЗ подобным критериям являются зоны пониженного сопротивления на тех же глубинах.

По результатам интерпретации данных сейсморазведки и электроразведки в конечном итоге строятся скоростные модели и модели проводимости земной коры, используемые в дальнейшем для создания согласованных по всем методам моделей.

Более детальную характеристику геометрии разреза земной коры обеспечивают результаты трансформаций гравитационного и магнитного полей в нижнее полупространство. С этой целью на практике используются методы Б.А. Андреева, его модифицированный вариант для 3D-моделей, А.В. Петрова и алгоритм А.И. Кобрунова и И.И. Приезжева, реализованный в спектральной области для 2D и 3D-моделей [6].

По результатам комплексного анализа атрибутов гравитационного и магнитного полей осуществляется структурно-тектоническое районирование территории исследований. Поэтому стандартный граф обработки и интерпретации потенциальных полей включает, с одной стороны, процедуры интерпретационной обработки, направленной на создание структурно-тектонической основы территории исследований с поверхности, с другой стороны, процедуры непосредственной количественной интерпретации этих полей по созданию плотностных и магнитных моделей земной коры.

Процедуры интерпретационной обработки (термин «интерпретационная обработка» определяет неразрывную связь между обработкой и интерпретацией) потенциальных полей включают, с одной стороны, разложение поля на составляющие (региональную и локальные) путем двумерной адаптивной энергетической фильтрации, с другой стороны расчеты различных статистических, градиентных, спектральных атрибутов полей в скользящих окнах. Разложение полей позволяет оконтурить наиболее выраженные по энергии области, отражающие аномалиеобразующие источники. Их

привязка по глубине ориентируется на размеры выделенных аномальных зон. Чем глубже источник, тем большую по размерам зону он характеризует. Статистические, градиентные атрибуты определяют положение зон нарушений. Возможность их картирования по глубине осуществляется по данным трансформаций поля в нижнее полупространство.

Совпадение аномальных зон, выделенных одновременно по гравитационному и магнитному полям указывает на их единый источник.

Совпадение зон тектонических нарушений, выделенных по гравитационным полям подтверждает их однозначность, а участки пересечения нарушений различного простирания, как правило, приурочены к возможным рудопроявлениям и образованиям нефтегазовых ловушек.

Отметим целесообразность использования самонастраивающейся фильтрации по трассированию зон нарушений [8]. Непрерывный характер аномальных зон нарушений свидетельствует о более молодых по геологическому возрасту структурах, а прерывистый характер о более поздних фазах заложения.

В конечном итоге результаты интерпретационной обработки позволяют обеспечить построение структурно-тектонической схемы исследуемой территории.

Процедуры количественной интерпретации потенциальных полей направлены на построение плотностных и магнитных моделей земной коры в 2D варианте по разрезам и 3D вариантах в объеме изучаемой среды.

Основные процедуры интерпретации стандартного графа включают:

- 3D (2D) трансформации поля в нижнее полупространство, обеспечивающие построение объемных или плоскостных (по разрезу) распределения эффективных (избыточных) значений плотности и намагниченности.

При транспонировании магнитного поля предварительно производится приведение к магнитному полюсу.

Трансформация поля в нижнее полупространство приводит к выделению блоковой структуры исследуемого глубинного разреза:

- структурный анализ результатов трансформаций по созданию блоковых моделей земной коры. При этом для сужения неоднозначности моделей следует ориентироваться на результаты сейсморазведки, в частности, на сейсмоэнергетические разрезы и глубинной электроразведки;

- задание априорной плотностной (и магнитной) модели. При отсутствии скоростной модели ориентируются на задание градиентной модели с учетом изменения плотности от 2,6 г/см³ по фундаменту до 3,30 г/см³ для поверхности Мохоровичича. Обычно используется пяти-шестислойная модель. При наличии данных ГСЗ скорости переводятся в плотности.

При задании априорной магнитной модели исходят из возможных по составу слагающих пород земной коры, например гранитный, диоритовый и базальтовый. В общем случае ориентируются на увеличение основности горных пород с глубиной.

– построение плотностной и магнитной моделей земной коры на основе данных структурного анализа и заданных априорных моделей, путем использования методов подбора при решении прямых задач. Подбор моделей завершается при достижении заданной величины невязки между наблюдаемыми и модельными (расчетными) значениями полей;

На рис. 8.2 и 8.3 соответственно приведены исходные разрезы и результаты методной интерпретации данных глубинной сейсморазведки (рис. 8.2) и гравимагниторазведки и МТЗ (рис. 8.3). На рис. 8.4 представлена согласованная комплексная модель земной коры по профилю I-СБ, в которой каждый выделенный блок охарактеризован соответственно значениями плотности (рис. 8.4, а), скорости (рис. 8.4, б) и сопротивления (рис. 8.4, в). В дальнейшем по этим данным реализован прогноз вещественного состава этих блоков путем сопоставления физических свойств с породами верхней, средней и нижней частей земной коры. Прогноз вещественного состава приведен на рис. 8.5.

– уточнение моделей. Для уточнения моделей, связанного, прежде всего, с прослеживанием субгоризонтальных границ разреза, неоднозначно проявляемых в результатах структурного анализа потенциальных полей можно рекомендовать ряд специализированных процедур. Одной из процедур является расчет вертикального градиента по результатам трансформаций полей. Наличие локальных максимумов и минимумов вертикального градиента указывает на положение субгоризонтальной границы.

На рис. 8.6 по профилю 3 ДВ последовательно представлены: временной глубинный разрез (а), разрез распределения избыточной плотности (б), плотностная модель (в), формализованная структурная модель (г) и сейсмоэнергетический разрез с наложенной структурной моделью (д).

На рис. 8.7 представлен результат многопризнаковой фильтрации данных трансформаций гравитационного поля по профилю I-СБ, с выделением субгоризонтальных границ.

В качестве специализированной процедуры интерпретации по выделению субгоризонтальных границ предлагается применение алгоритма многопризнаковой фильтрации [7]. С этой целью из куба результатов 3D трансформаций поля выбирается блок, включающий несколько (10-12) профилей, расположенных через 2 км (для масштаба 1:200 000), протяженностью 100 км, а по глубине в 40 км при интервале 1 км. Количество точек в таком блоке по пространственным осям 50х10х40.

В качестве признака принимается отдельный профиль в 50 точек с его результатами трансформации по глубине в 40 точек, т.е. матрица 50х40.

Алгоритм многопризнаковой фильтрации сводится к двухэтапной процедуре метода главных компонент. На первом этапе рассчитывается первая главная компонента по результатам трансформации каждого профиля. Количество значений первой главной компоненты соответствует числу глубинных уровней профиля – 40. Значение такой компоненты определяют весовые множители (веса) для каждого глубинного уровня. На втором этапе для полученных значений первых главных компонент, для 10 профилей их будет 10, вновь рассчитывается первая главная компонента, значения которой характеризуют уже «веса» каждого пикета, т.е. точки вдоль профилей, всего точек – 50. Перемножение значений первой главной компоненты для 40 глубинных уровней на значения первой главной компоненты для 50 пикетов, позволяет охарактеризовать результаты в виде матрицы из 50 значений вдоль оси X и 40 значений по оси Z (глубине). По сравнению с исходными результатами трансформации полученная матрица обеспечивает выделение наиболее энергоемкой составляющей с учетом применения корреляционных связей как по X, так и по Z. Такая составляющая подчеркивает субгоризонтальную границу в земной коре.

Оценка качества моделей обеспечивается:

1) совпадением наблюдаемого и расчетного от модели полей при заданной невязке. Так, для съемок масштаба 1:200 000 для гравитационного поля эта невязка составляет 2 мГал, для магнитного поля 1-2 нТл. Учитывая возросшую точность наблюдений гравитационного поля при использовании современных аэрогравиметров для того же масштаба до 0,5 мГал, соответственно уменьшается величина невязки;

2) совпадением плотностных и магнитных границ раздела земной коры в местах пересечения региональных и опорных профилей:

3) совпадением зон пересечения тектонических нарушений на поверхности земли с выходами зон глубинных разломов. Этот критерий может служить признаком при обнаружении месторождений как углеводородного, так и рудного сырья;

4) непротиворечивостью построенных моделей современным геологическим представлениям и гипотезам о строении земной коры.

Создание согласованной физико-геологической модели земной коры реализуется путем построения методных моделей. На их основе образуется единая геометризованная среда в 2D или 3D вариантах.

Для такой усредненной единой геометрии определяются физические свойства: плотность, намагниченность, скорость и сопротивление для выделенных блоков земной коры. Сопоставление набора физических свойств

блоков со справочными данными позволяет перейти к вещественному составу земной коры. Оценка вещественного состава земной коры для регионального профиля I-СБ по территории Восточной Сибири, пересекающего обширную Юрубчено-Тохомскую нефтегазоносную зону, проведена Н.Н. Пимановой. Эти результаты приведены на рис. 8.5, включают определение плотности, скорости и проводимости, и на их основе реализуется переход к вещественному составу пород земной коры. Магнитная восприимчивость не учтена по причине малой глубинности магниторазведки, составляющей около 15 км.

Построение согласованных моделей и оценки вещественного состава пород земной коры осуществлены в среде геоинформационной системы ГИС ИНТЕГРО и ее модуля ГИС ИНТЕГРО ГЕОФИЗИКА [2].

В заключение подчеркнем целесообразность построения согласованных ФГМ земной коры для решения задач регионального прогноза месторождений полезных ископаемых.

ГЛАВА IX. ТЕХНОЛОГИЯ «ПРЯМОГО» ПРОГНОЗА УГЛЕВОДОРОДОВ НА ОСНОВЕ КОМПЛЕКСНОГО АНАЛИЗА АТТРИБУТОВ ГЕОФИЗИЧЕСКИХ ПОЛЕЙ

Впервые термин «прямые поиски» углеводородов появился в СССР в начале 1960-х годов и был связан с установленным И.Г. Медовским фактом отсутствия регулярных отражений ниже нефтяных залежей на 16 месторождениях Азербайджана. Опытные-методические работы по проблеме прямых поисков нефти и газа были проведены в 1961-1962 гг. на двух нефтяных месторождениях (Колодезное и Величаевское) и на хорошо разбуренной пустой структуре Максимовская на территории Ставропольского края. Эти работы показали наличие четких отражений под нефтяными месторождениями и их отсутствие под пустой структурой. Таким образом, критерий снижения рефлексивности отраженных волн под продуктивными структурами оказался неоднозначным. Поэтому дальнейшие исследования по проблеме «прямых поисков» были направлены на расширение признаков (атттрибутов) прямого прогноза. До сих пор термин «прямой прогноз» используется в ряде как отечественных, так и зарубежных публикаций. При этом основное внимание уделяется поглощению сейсмических волн. Долгое время АВО-эффект (эффект, связанный с увеличением амплитуды сигнала при удалении от залежи) также считался «прямым» признаком залежи углеводородов (УВ). Однако этот эффект достаточно уверенно фиксируется лишь для крупных по толщине газовых коллекторов и практически не проявляется для маломощных коллекторов нефти.

Привлечение широкого набора кинематических и динамических атттрибутов волнового поля совместно с атттрибутами потенциальных и электромагнитных полей обеспечивает существенное сужение неоднозначности при выявлении ловушек УВ, эффективность бурения при открытии которых составляет не более 30% при ограниченном числе атттрибутов волнового поля.

Заметим, что разрешающая способность сейсморазведки МОГТ составляет половину длины волны, т.е. пласты мощностью 8-10 м не могут быть выявлены сейсморазведкой. Однако согласно гипотезе Головкинского-Вальтера нефтегазовая залежь оказывает воздействие на вмещающую среду, и ее влияние может превышать для маломощных залежей разрешающую способность.

К настоящему времени сужение неоднозначности при установлении нефтегазоносности разреза целиком базируется на комплексном анализе атттрибутов волнового поля и атттрибутов других геофизических полей. Расчет статистических атттрибутов волнового поля для наиболее распространенной съемки масштаба 1:50 000 осуществляется в скользящих окнах размером 7-9

трасс при 13-15 дискретов по времени. Это обеспечивает состоятельность и эффективность оценок атрибутов.

К наиболее информативным динамическим атрибутам относятся:

- увеличение (реже уменьшение) амплитуды отражений от коллектора;
- увеличение затухания упругих колебаний, определяемые по отношению амплитуд волн, отраженных от подошвы и от кровли коллектора, либо по величине коэффициента поглощения p в сейсмическом импульсе $S(f) = e^{-px} \cos \beta x$ (β – видимый период);
- понижение частотного состава отражений;
- наличие положительных аномалий AVO-эффекта в пределах целевого горизонта;
- наличие низкочастотного резонанса, обусловленного эмиссией геодинамического шума. При его оценке в пределах целевого горизонта производится вычитание всех регулярных волн путем двумерной адаптивной энергетической фиксации, а по остаточному полю геодинамического шума, создаваемого самой залежью, реализуется вейвлет-анализ. По результатам вейвлет-анализа рассчитываются амплитуды и частоты коэффициентов вейвлета, в качестве которого обычно используется Морле-вейвлет. Совпадение повышенных значений (аномалий) амплитуд вейвлета с пониженными значениями частоты вейвлета указывает на наличие низкочастотного резонанса геодинамического шума (рис. 9.1, см. Приложение).

В качестве информативных кинематических атрибутов выступают:

- уменьшение интервальной скорости до 20-30%, определяемое по амплитудам отраженных волн от пласта-коллектора;
- снижение рефлексивности (отражательной способности) волн под залежью;
- появление на временном разрезе горизонтальных отражающих площадок, особенно проявляемых на фоне наклонных границ структурных форм вмещающей среды. Такие площадки приурочены к водо-нефтяным и водо-газовым контактам (ВНК и ВГК). Наличие ВНК и ВГК может быть установлено путем расчетов асимметрии и эксцесса распределений приращений Δt между соседними каналами (трассами). Эксцесс сказывается на островершинности распределения Δt . Само распределение Δt оценивается во временных окнах в 0,1 с;
- смена полярности волн, часто наблюдаемая при переходе от водного коллектора к газонефтяному;
- резкое изменение градиентных атрибутов отражений.

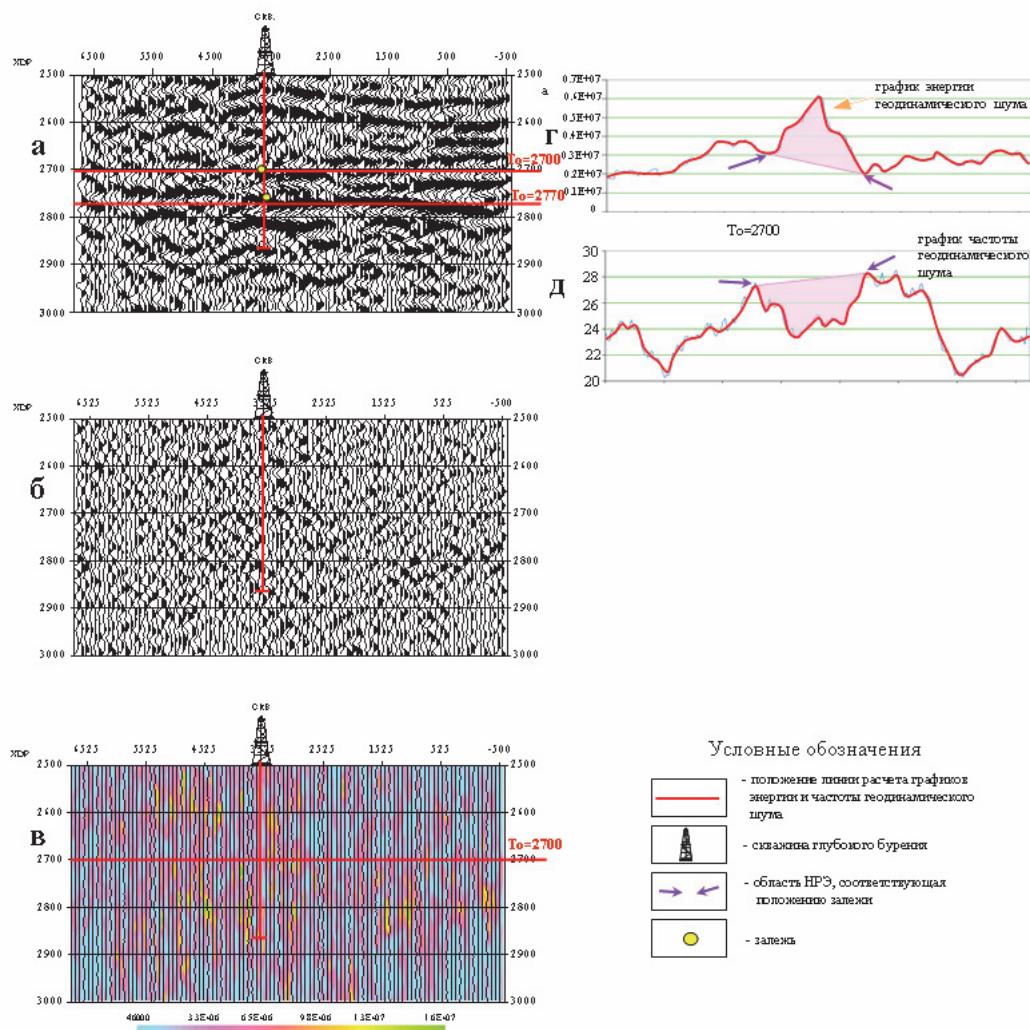


Рис. 9.1. Выделение УВ насыщенной части разреза на основе применения технологии НРЭ геодинамического шума (Краснодарский край)
а – временной разрез МОГТ;
б – временной разрез геодинамического шума;
в – результат вейвлет-преобразования геодинамического шума;
г, д – графики энергии и частоты соответственно.

Вычисление атрибутов анизотропии и энтропии в скользящих окнах временного разреза часто позволяет оконтурить расположение залежей УВ. Использование только динамических атрибутов (амплитуды, энергии, частоты, затухания) обеспечивает локализацию залежей УВ на основе метода главных компонент и метода *K*-средних [8].

В гравитационном поле, ввиду дефекта плотности газа (плотность $0,65 \text{ г/см}^3$) и нефти ($0,9-1 \text{ г/см}^3$) по сравнению с плотностью пластовых вод в $1,1 \text{ г/см}^3$ коллектора и с вмещающими породами с плотностью $2,3 \text{ г/см}^3$ наблюдается либо отрицательная аномалия Δg , либо выполаживание кривой Δg над залежью. Этот эффект достигается для локальных составляющих поля Δg при детальном съемках масштаба 1:50 000 и крупнее, для предварительно исключенной региональной составляющей. Для многопластовых залежей отрицательный эффект поля Δg может составлять 1-2 мГал. Существенная надежность аномальных эффектов поля Δg обеспечивается за счет их корреляции по нескольким профилям съемки.

В то же время следует заметить, что подобный эффект может быть обусловлен повышенной трещиноватостью пород, перекрывающих залежи УВ, которой сопровождается антиклинальные складки.

Достаточно эффективными для поисков залежей УВ являются электрические и электромагнитные зондирования: ВЭЗ, ВЭЗ-ВП, ЗСПБ. Их применение связано с тем, что нефтегазонасыщенные коллекторы характеризуются повышенными по сравнению с вмещающими породами, удельными сопротивлениями. Это объясняется, с одной стороны, более высоким сопротивлением самих нефтегазовых залежей за счет непроводящих ток нефти и газа в пористых породах, с другой стороны, играет роль уплотнение пород за счет высокого пластового давления, а также карбонитизации пород. По целому ряду нефтегазовых месторождений в контуре нефтегазонасыщенности и над залежью параметры поляризуемости становятся выше вокруг залежи. Это связано с наличием вкрапленности пирита и других продуктов окисления, образующихся за счет миграции и окисления углеводородов залежи. Следовательно, полезными становятся наблюдения методом вызванных потенциалов на постоянном токе (ВЭЗ-ВП) и частотно-меняющемся поле (УЭЗ-ВП).

Эффективность методов электроразведки достигается при выявлении достаточно мощных залежей ($\sim 100 \text{ м}$), залегающих на глубинах 2-2,5 км. Обычно считается, что эффект от залежи по величине сопротивления проявляется, если мощность коллектора составляет 1/10 от мощности вышележащих отложений.

В.А. Сидоровым и др. [4] высказывалась гипотеза о влиянии залежи на электрическое сопротивление вышезалегающих пород, вплоть до дневной поверхности. Физической основой такого эффекта является насыщение газом вышезалегающих пород, в которых газ находится в рассеянном виде. При этом аномальные эффекты проводимости обусловлены главным образом приконтактными зонами. Последние являются в тектоническом плане наиболее ослабленными, а над самой залежью экраны в виде глинистых покровов препятствуют проникновению газовых фракций.

Аномалии проводимости в приконтактных к залежи зонах по интенсивности могут превосходить аномалии от залежи УВ. Так, если считать, что 8-10 метровая залежь содержит 100% газа, а в законтурных отложениях в «столбе» вышезалегающих пород до 3000 м содержание газа составляет 1%, то произведение мощности таких пород на содержание газа в 1% ($3000 \cdot 0,01 = 30$) эквивалентно влиянию залежи 30-метровой толщины.

Комплексный анализ за счет интеграции аномальных эффектов различных полей и их атрибутов в конечном итоге обеспечивает повышение однозначного выделения залежей УВ и как бы приводит к их «прямому» прогнозу. Однако многообразие самих залежей и условий их залегания в различных геологических условиях приводит к необходимости расширения диапазона применения различных геофизических методов и разработке новых информативных атрибутов волновых, потенциальных и электромагнитных полей.

В целом технологию «прямого» прогноза УВ можно свести к двух-этапной процедуре: 1) установление диагностически значимых (информационных) атрибутов геофизических полей и выбора их информативной совокупности; 2) применение методов распознавания образов, рассмотренных в главах III, IV, V.

Появление новых типов ловушек УВ, ловушек в породах кристаллического фундамента, газогидратов, сланцевой нефти, силицитов вынуждает искать все новые диагностически значимые атрибуты геофизических полей.

ГЛАВА X. ТЕХНОЛОГИИ ПОСТРОЕНИЯ СОГЛАСОВАННЫХ ФГМ ОСАДОЧНОГО ЧЕХЛА И ФУНДАМЕНТА ПРИ ИНТЕРПРЕТАЦИИ ДАННЫХ НАЗЕМНЫХ И СКВАЖИННЫХ МЕТОДОВ

Построение согласованных ФГМ, т.е. моделей, представляющих распределение в нижнем полупространстве изучаемой геосреды набора физических свойств, осредненных на единой геометрии слоев и блоков, получило в настоящее время широкое применение при оценке строения осадочного чехла и кристаллического фундамента. При этом комплексная количественная интерпретация реализуется, главным образом, с использованием данных двух методов: сейсморазведка – гравиразведка, сейсморазведка – электро-разведка, сейсморазведка – геофизические методы исследования скважин, а при оценке строения фундамента эффективным комплексом также является гравиразведка – магниторазведка.

Основным методом в технологии построения согласованных ФГМ служит сейсморазведка, с помощью которой определяются горизонтальные и субгоризонтальные границы осадочного чехла и фундамента. В качестве основного метода при изучении глубинного строения осадочного чехла до 2-3 км могут служить методы электромагнитного зондирования при достаточно большой мощности слоев.

В целом, если точность определения глубины залегания субгоризонтальных границ по данным сейсморазведки составляет 1-5%, то по данным электрозондирования 10-30%, по данным гравиразведки более 30%. В то же время положение субвертикальных границ с большей точностью оценивается по данным электропрофилирования, гравиразведки и отчасти магниторазведки.

Технология построения сейсмоплотностных моделей включает:

1) определение глубин залегания субгоризонтальных границ и скоростей слоев по данным сейсморазведки МОГТ;

2) оценку блокового строения по данным гравиразведки и разбиение каждого слоя на блоки с учетом детальности проведенных исследований по пересчету поля Δg в нижнее полупространство;

3) оценку плотностей выделенных блоков с использованием либо формулы Гарднера $\sigma_{пл} = 0,31V^{1/4}$, где плотность $\sigma_{пл}$ в г/см³, а скорость в м/с, либо путем линейных регрессионных зависимостей, имеющих для данного региона. В формуле Гарднера $\sigma_{пл} = aV^b$ коэффициенты a и b несколько различны для терригенных и карбонатных пород осадочного чехла.

Например, для северо-западной части Прикаспия для терригенных пород $a=0,309$; $b=0,225$; для карбонатных пород $c=0,325$;

4) для полученного плотностного разреза путем изменения границ блоков и их плотностей решается прямая задача гравиразведки и оценивается невязка между наблюдаемыми и рассчитанными значениями Δg , которая не должна превышать 0,4 мГал для масштаба 1:50 000.

В том случае, когда не удастся путем применения скоростей и размеров блоков добиться требуемой невязки, возвращаются к переоценке положения субгоризонтальных границ и скоростей по данным сейсморазведки, с учетом которых вновь осуществляется разбиение слоев на блоки и определение их плотностей. Подобные итерации позволяют за приемлемое время добиться создания согласованной сейсмо-плотностной модели. Время расчетов существенно сокращается при использовании имеющихся данных акустического и плотностного каротажа, а также с учетом данных пересчета гравитационного поля в нижнее полупространство, как это было описано в главе VIII.

Определение значений скорости и плотности обеспечивает возможность расчета акустического импеданса, т.е. произведения скорости на плотность. Разрезы акустического импеданса позволяют осуществить более надежную корреляцию границ отражающих горизонтов. При этом требуется привлечение данных акустического и плотностного каротажа, о чем пойдет речь в разделе интерпретации данных сейсморазведки и ГИС.

Технология построения согласованных сейсмoeлектрических временных разрезов (СЭВР).

Эта технология базируется на установленной нелинейной связи между временем t_0 , полученном по годографам сейсмического временного разреза и времени $t_{ЗСБ}$, полученным по годографам электромагнитного зондирования методом становления поля в ближней зоне (ЗСБ) $t_0 = 2a\sqrt{t_{ЗСБ}}$.

Для выяснения физического смысла коэффициента «а» рассмотрим многослойную среду мощностью H в качестве сейсмoeлектрической модели, характеризующей:

– по данным сейсморазведки – средней скоростью упругих колебаний $V_{ср.}$, равной

$$V_{ср.} = \frac{h_1 + \dots + h_n}{t_1 + \dots + t_n} = \frac{h_1 + \dots + h_n}{\frac{h_1}{V_1} + \dots + \frac{h_n}{V_n}} = \frac{\sum h_i}{\sum \frac{h_i}{V_i}} = \frac{H}{t_0},$$

где h_i , t_i и V_i – мощность, время пробега и скорость продольных волн в каждом из n -однородных слоев, H и t_0 – мощность и время пробега волны через всю среду по нормали к границам раздела;

– по данным электромагнитного зондирования – суммарной продольной проводимостью S :

$$S = \frac{h_1 + \dots + h_n}{\rho_1 + \dots + \rho_n} = \frac{h_1 + \dots + h_n}{\frac{h_1}{S_1} + \dots + \frac{h_n}{S_n}} = \frac{\sum h_i}{\sum \frac{h_i}{S_i}} = \frac{H}{\rho_l},$$

где h_i , ρ_i , S_i – соответственно мощность, удельное сопротивление и проводимость каждого из n -однородных слоев, H и ρ_l – мощность и среднее продольное сопротивление среды.

Таким образом, по данным сейсморазведки $H = \frac{t_0 V_{cp}}{2}$, а по данным ЗСБ: $H_k = k \sqrt{t_{ЗСБ} / \mu \sigma}$, где k – коэффициент характеристики направленности зондирования; σ – удельная продольная проводимость среды мощностью H ; μ – ее магнитная проницаемость ($\mu = 4\pi \cdot 10^{-7} \text{ Гн / м}$).

При совпадении опорного электрического и отражающего сейсмического горизонтов получаем

$$\frac{t_0 V_{cp}}{2} = k \sqrt{t_{ЗСБ} / \mu \sigma},$$

$$\text{т.е. } t_0 = \frac{2k}{V_{cp} \sqrt{\mu \sigma}} \sqrt{t_{ЗСБ}} \text{ или } a = \frac{k}{V_{cp} \sqrt{\mu \sigma}}.$$

Величина a зависит от геологического строения исследуемого района и колеблется обычно в пределах от 0,5 до 0,6. Относительная ошибка определения a для большинства нефтегазоносных провинций не превышает 5% и зависимость можно считать функциональной.

Совпадение сейсмических и геоэлектрических границ по методике СЭВР позволяет картировать высокоомные аномальные объекты, представляющие интерес для прогноза нефтегазоносности, выделять в подсоловых отложениях зоны тектонических нарушений, свидетельствующие о блоковом строении района, скорректировать структурную основу модели для дальнейшего построения плотностной модели среды. Последнее обстоятельство уже приводит к возможности комплексной интерпретации данных сейсмо-, электро- и гравиразведки (методика Н.П. Смилевец).

Комплексная интерпретация данных сейсмо- электро- и гравиразведки в единой информационной среде проведена по одному из региональных

профилей, пересекающих Казанско-Кажимский авлакоген (ККА) в широтном направлении. Использование данной методики позволило:

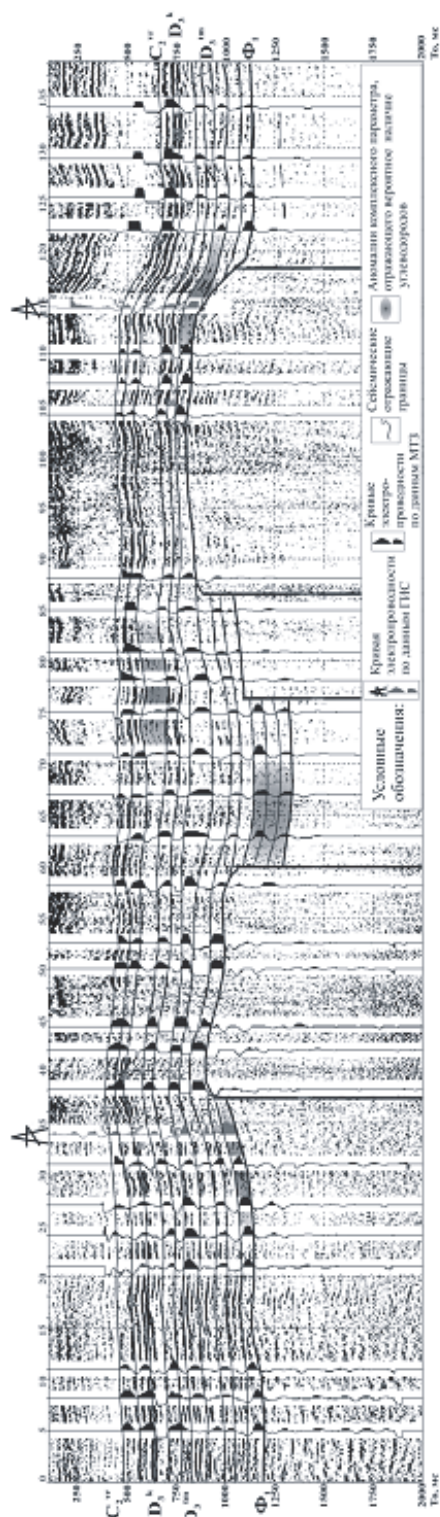
- уточнить строение перспективных комплексов палеозойского интервала разреза;
- выполнить прогноз нефтегазоносности отложений среднего и нижнего карбона, а также карбонатного и терригенного комплексов девона;
- выделить зоны тектонических нарушений;
- изучить рельеф поверхности и внутреннюю структуру кристаллического фундамента, выявить его блоковую природу;
- построить согласованную физико-геологическую региональную модель ККА, нашедшую отображение в сейсмическом, электрическом и гравитационном полях (см. рис. 10.1).

Комплексная интерпретация данных сейсморазведки и ГИС.

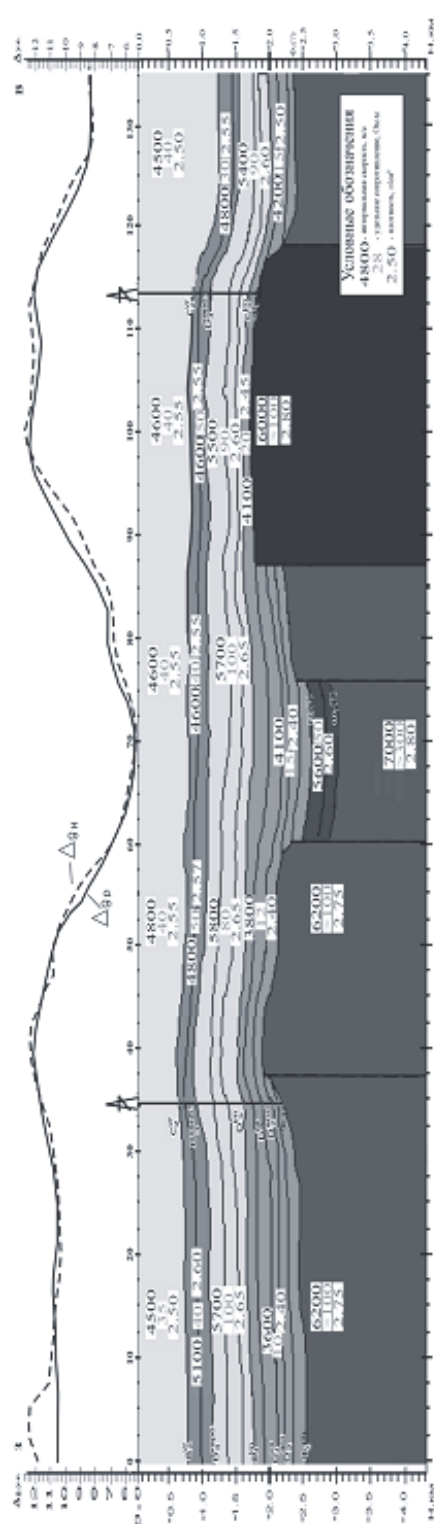
Данные геофизических исследований скважин по своей природе уже являются комплексными. По сочетанию результатов различных методов ГИС осуществляется прогноз литологического состава пород, а по результатам атрибутивного анализа данных сейсморазведки обеспечивается прогноз литологического состава в межскважинном и заскважинном пространстве при наличии регрессионных зависимостей между атрибутами сейсмической записи и данными ГИС.

С целью прогноза нефтегазоносности пластов наиболее популярны в настоящее время регрессионные и вероятностные нейронные сети. При наличии нескольких нефтегазоносных пластов производится отбор диагностически значимых сочетаний данных методов ГИС. Путем перебора диагностически значимых сочетаний выбирается оптимальный комплекс ГИС для оценки нефтегазоносности пластов. Критерием отбора служит величина множественного коэффициента корреляции. В том случае, когда величина множественного коэффициента регрессии для регрессионной зависимости отобранных сочетаний методов ГИС перестает значимо увеличиваться, процесс отбора диагностически значимых сочетаний прекращается.

Вероятностные нейронные сети базируются на использовании формулы Байеса. При этом критерием отбора диагностически значимых сочетаний методов ГИС является заданная величина апостериорной вероятности. По существу вероятностные нейронные сети представляют статистические методы распознавания образов при наличии эталонных объектов, рассмотренных в главе III. В качестве эталонных объектов выступают нефтегазоносные и водоносные пласты при распознавании образов на два класса.



А. Сейсмoeлектромагнетичный временной разрез



Б. Согласованная параметрическая модель

Рис. 10.1. Пример построения региональной согласованной модели разреза на основе комплексной интерпретации данных сейсмо-, электро- и гравитационной информации в единой информационной среде

При оценке литологического состава пород по скважине формула Байеса используется для распознавания образов (объектов) на несколько классов, для которых необходимо иметь достаточный статистический материал для обучения, т.е. порядка десяти. Большую роль при оценке прогноза ловушек углеводородов играет построение разрезов акустического импеданса, т.е. произведение скорости на плотность и соответственно инверсионные преобразования временного разреза.

Инверсионные преобразования типа ПАК

Основой акустической инверсии типа псевдоакустического каротажа (ПАК) является уравнение-свертка $u(t) = r(t) * w(t)$, где $u(t)$ – сигнальная часть сейсмотрассы, $r(t)$ – трасса коэффициентов отражения, $w(t)$ – сейсмический импульс.

Обычно трасса коэффициентов отражения определяется путем дифференцирования логарифма единичного импульса по времени t :

$$z(t) = \frac{1}{2} \frac{\Delta}{\Delta t} [\ln I(t)].$$

Суть акустической инверсии сводится к определению зависимости акустического импеданса (произведение скорости на плотность) от времени t по записи отраженных волн при нормальном падении пластов – $u_z(t)$. В качестве этой зависимости используется трасса временного разреза.

Выражение:

$$\hat{I}(t + \Delta t) = \hat{I}(t) \frac{1 + \hat{r}(t)}{1 - \hat{r}(t)}$$

отражает суть акустической инверсии, а именно, зная начальное значение импеданса $\hat{I}$ в какой-либо точке среды можно перейти от коэффициентов отражения, характеризующих внутренние границы среды, к непосредственной оценке импеданса. В то же время следует иметь в виду, что акустический импеданс определяется в предположении, что суммотрасса ОГТ является результатом свертки коэффициентов отражения с одним и тем же импульсом $w(t)$. Для свертки обычно используется модельный сигнал, который нельзя извлечь из сейсмограммы, представляющей сложное интерференционное поле.

При этом наибольшие погрешности определения импеданса обусловлены наличием кратных волн, которые предварительно следует отфильтровать. Кроме того, акустическая инверсия реализуется в предположении о том, что длина волны существенно меньше размера неоднородностей среды.

В то же время инверсионные преобразования, как правило, позволяют обеспечить детализацию акустических свойств особенно для сложно-построенных ловушек УВ.

Высокий коэффициент корреляции 0,8-0,9 акустической жесткости с пористостью, отмечаемый по данным каротажа, в частности, для карбонатных коллекторов, определяет особое значение инверсии сейсмической записи в акустическую жесткость (акустические импедансы). Ниже рассматриваются акустическая инверсия суммарных сейсмотрасс при нормальном падении пласта.

Основные этапы реализации акустической (сейсмической) инверсии включают:

- контроль качества и подготовка исходных данных;
- стратиграфическая привязка сейсмических отражений и оценка формы сейсмического импульса (импульсов);
- построение априорной геоакустической модели, содержащей отсутствующую в сейсмическом диапазоне информацию о низкочастотной составляющей скорости и плотности, которая распространяется по изучаемому объему с учетом поведения отражающих горизонтов;
- подбор параметров инверсии и проведение инверсии с получением разрезов (кубов при 3D-сейсморазведке) упругих параметров и контроль качества инверсии;
- интерпретация результатов инверсии в терминах целевых свойств ловушек УВ.

Наиболее распространенными видами инверсии является детерминистическая и стохастическая.

При детерминистической акустической инверсии суммотрассы ОГТ преобразуются в разрезы (кубы) акустического импеданса (возможно скорости V_p). Основным ограничением на модель среды является требование «вертикальной редкости» последовательности коэффициентов отражения и диапазон возможных изменений акустических свойств.

Вертикальная разрешенность для куба Р-импеданса составляет около 1/4 эффективной длины сейсмического импульса. При наличии в скважинах надежной связи акустического импеданса с пористостью, литологией и нефтегазонасыщенностью значения импеданса распространяются в межскважинном пространстве для этих свойств ловушек УВ. Кроме того, Р-импеданс может быть использован для уточнения стратиграфической привязки отражающих горизонтов.

Для варианта детерминистической синхронной (AVO/AVA) инверсии несколько угловых кубов синхронно преобразуются в кубы упругих пара-

метров: скорости V_P , V_S и объемной плотности. Далее полученные кубы используют для расчета V_P/V_S и соответственно таких упругих модулей, как коэффициент Пуассона, модуль сдвига, модуль (постоянная) Ламэ, которые обеспечивают более надежный прогноз УВ-насыщения пористости, литологии по сравнению с данными лишь акустического импеданса.

Стохастическая инверсия связана с расчетом множества моделей среды, для которых вычисляются такие характеристики как среднее, дисперсия, корреляционные функции, по которым строится вероятностная модель ловушки УВ. Разнообразные реализации модели согласуются с сейсмическими и скважинными данными. Полученные наборы реализаций кубов пористости, насыщенности и литологии используются для оценки распределения вероятности целевых параметров коллектора.

Детерминистическую акустическую инверсию используют для прогноза свойств коллектора при наличии тесной связи между коллекторскими свойствами и акустическим импедансом. Данная инверсия не требует большого количества скважин, ограничивается стандартными методами ГИС, даже не требует измерения величины V_S и данных по объемной плотности пород.

От сейсмических данных требуется наличие суммарного куба с сохранением амплитуд суммотрасс. Однако надежный прогноз коллекторских свойств и нефтегазонасыщенности реализуется при скважинных измерениях скоростей V_S . При отсутствии данных о V_S ограничиваются отношением $V_S/V_P = 0,5$, при определении плотности используется формула Гарднера $\sigma_{пл} = 0,31V^{1/4}$.

Применение геостатической инверсии требует достаточного объема априорной геолого-геофизической информации, включая исследования керна и методов ГИС. При акустической инверсии особое значение представляет выбор формы моделирующего импульса. Для однотипных сейсмогеологических условий обычно наблюдается сохранение формы импульса. Для нуль-фазового импульса при этом можно ограничиться расчетом автокорреляционной функции по динамически выраженным горизонтам с высоким отношением «сигнал/помеха». Длина оценки при извлечении сигнала (импульса) составляет не менее 3-5 его длин. Расчет автокорреляционной функции производится адаптивно, ориентируясь на сохранении стационарного характера сейсмозаписи, которая составляет 0,7-0,9 с.

В том случае, когда оценка фазы сигнала существенно отлична от нуля (более $\pm 30-40^\circ$) и выдержана от скважины к скважине, тогда для приведения сейсмических данных к нуль-фазовому виду либо проводится корректировка

фазы сейсмических трасс, либо используется импульс, отличный от нуль-фазового. Интервал расчета импульса содержит несколько амплитудно выдержанных отражений, захватывает целевой интервал и составляет не менее 300 мс. При этом важен высокий коэффициент корреляции синтетических и сейсмических трасс – не менее 0,7, что обеспечивает хорошее качество инверсии.

Оценка упруго-деформационных характеристик

Упруго-деформационные свойства среды определяются на основе упругих параметров, к которым относятся:

- коэффициент всестороннего сжатия, иначе *модуль упругости* K , равный отношению нормального напряжения σ , возникающего при всестороннем сжатии (объемного давления) к относительному изменению объема $\theta = \Delta V / w$ – дилатации среды:

$$K = \frac{\sigma}{\theta}.$$

Обратная величина $\beta = 1 / K$ характеризует *сжимаемость* среды.

- *модуль продольной упругости Юнга*, равный отношению нормального напряжения σ к продольной деформации ε :

$$E = \sigma / \varepsilon.$$

Модуль Юнга может различаться в сотни и тысячи раз.

- *модуль сдвига* (модуль поперечной упругости) G – равный отношению касательного напряжения τ к углу сдвига γ :

$$G = \tau / \gamma;$$

- коэффициент Пуассона:

$$\mu = \varepsilon_1 / \varepsilon,$$

где ε_1 – относительная поперечная деформация. Коэффициент Пуассона изменяется от 0 до 0,5.

Имея данные о скоростях V_P и V_S и плотности ρ , полученные по результатам широкополосного акустического каротажа (АКШ), многоволнового ВСП и плотностного каротажа, нетрудно определить:

- модуль продольной упругости Юнга $E = \rho \frac{V_S^2 (3V_P^2 - 4V_S^2)}{V_P^2 - V_S^2}$;
- модуль сдвига $G = \rho V_S^2$ (или $G = T \frac{1}{2(1 + \mu)}$);

– коэффициент Пуассона $\mu = \frac{V_P^2 - 2V_S^2}{2(V_P^2 - V_S^2)}$.

Определение приведенных выше упруго-деформационных модулей обеспечивает в конечном итоге оценку пористости, фильтрационно-емкостных свойств коллекторов и в комплексе с данными ГИС приводит к надежному прогнозу УВ.

Комплексная интерпретация данных гравиразведки и магниторазведки. Наиболее эффективна для оценки рельефа кристаллического фундамента при его сравнительно неглубоком залегании до 3-4 км и прогноза его вещественного состава.

Имеющиеся технологии комплексной интерпретации данных гравиразведки и магниторазведки реализуют построение согласованных по плотности и магнитной восприимчивости моделей фундамента в 2D (по разрезам) и 3D (в объеме) пространствах.

В технологии СИГМА-3D (Ю.И. Блох) обеспечивается построение модели субгоризонтального слоя с латеральной изменчивостью намагниченности и плотности. Верхняя кромка моделируемого слоя задается, например, по данным сейсморазведки или электроразведки или на основе результатов пересчета потенциальных полей в нижнее полупространство. Нижняя кромка слоя принимается горизонтальной, а ее глубина оценивается по спектру интерпретируемого поля. Кроме того, нижняя кромка, т.е. ее граница, может быть задана из тех соображений, что и верхняя кромка. Для решения обратных задач гравиразведки и магниторазведки сформированный субгоризонтальный слой аппроксимируется совокупностью квадратных, либо треугольных в плане вертикальных однородных по плотности и намагниченности призм, расположенных в слое. Такие призмы располагают не только в самом слое, но и его обрамлениях, выходящих за пределы съемки с целью учета краевых эффектов.

Размер призм в плане выбирается исходя из средней глубины залегания верхней кромки, т.е. границы слоя и примерно принимается ей равным. Далее по наблюдаемым полям путем спектрального анализа определяется разность между плотностью или намагниченностью каждой из призм, по отношению к одной из них, принимаемой в качестве базовой. При моделировании магнитного поля требуется задание направления вектора намагниченности пород, которое принимается совпадающим с направлением главного геомагнитного поля в изучаемом регионе.

Число точек наблюдений гравитационного и магнитного полей должно быть значительно больше числа призм, что необходимо для осуществления контроля расчетов по имеющимся избыточным наблюдениям.

Для гравиразведки решение обратной задачи реализуется на основе системы линейных уравнений, а для магниторазведки решение обратной задачи нелинейно, что отличает технологию СИГМА-3D от других технологий [11].

Результаты моделирования представляют собой эффективные значения плотности и намагниченности, равные разностям между истинными значениями физических свойств каждой из элементарных призм и соответствующими значениям в «базовой» призме. Такой прием полностью устраняет влияние регионального фона.

Помимо распределения физических параметров (плотности и намагниченности) для изучаемого слоя (фундамента) получает остаточное поле, как разность между наблюдаемым полем и полем от модели непосредственно в точках наблюдений.

Остаточное поле обусловлено влиянием осадочного чехла при моделировании кристаллического фундамента. Это остаточное поле содержит также помехи техногенного происхождения и связанные с погрешностями аппроксимации моделируемого слоя.

В то же время остаточное поле позволяет эффективно решать задачи по изучению осадочного чехла.

Технология СИГМА-3D явилась основой построения карт фундамента по территории Русской (Восточно-Европейской) платформы.

ГЛАВА XI. ГЕОИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ КОМПЛЕКСНОГО АНАЛИЗА И ИНТЕРПРЕТАЦИИ ГЕОФИЗИЧЕСКИХ ДАННЫХ

В 70-80 гг. XX века было реализовано достаточно много различного типа автоматизированных систем по комплексному анализу геофизических полей.

Все такие системы включали построение базы данных, главным образом, реляционной, систему управления базой и пакет прикладных программ.

В качестве прикладного математического аппарата нашли применение методы математической логики системы КОРА-3 (МИНХиГП, ныне РГУ нефти и газа), система ТЕСТ (СО АН СССР), регрессионный анализ система РЕТР, АСКИНТ (МИНХиГП), потенциальные функции система ПОТЕНЦИАЛ (ВИМС), методы проверки статистических гипотез – система РЕГИОН (МНИИПУ), КОСКАД (МГРИ-РГГРУ) и др., реализующие распознавание образов с обучением на эталонных объектах.

Среди методов распознавания образов без обучения на эталонных объектах, так называемые классификационные алгоритмы, нашли применение метод главных компонент и метод К-средних, впервые реализованные в Казахстанской опытно-методической экспедиции Мингео КазССР П.Н. Горбуновым и А.Н. Кленчиным в середине 70-х годов XX века. В дальнейшем эти методы были включены в различные геоинформационные системы и технологии.

В 90-е годы XX века в связи с появлением географических информационных систем (Arc GIS, Arc Views, AutoCAD), которые обеспечивали точную координатную привязку любой точки площади исследований, топологическое покрытие, т.е. наложение точечных, линейных, площадных объектов и рельефа местности на единую картографическую основу, а также включение векторных данных, привело к разработке геоинформационных систем и технологий, в которых использовалось прикладное программное обеспечение АСОД по комплексному анализу геополей, отсутствующие в географических системах. Географические системы, а затем и геоинформационные системы позволяли строить карты и разрезы с их легендами. Те системы, в которых не использовались все достоинства географических систем, получили название технологий.

Одновременно с появлением географических систем резко, в десятки и сотни раз, возросла скорость персональных компьютеров, что обеспечило возможность реализации таких современных алгоритмов комплексного анализа как нейронные регрессионные и вероятностные сети и генетиче-

ские алгоритмы. Одной из первых отечественных геоинформационных систем в середине 90-х годов была создана система ПАРК (прогноз – анализ – распознавание – картирование) под руководством М.А. Белоборова. Практически одновременно во ВНИИгеосистем была создана система ГИС ИНТЕГРО под руководством Е.Н. Черемисиной и М.Я. Финкельштейна. Развитие ГИС ПАРК в настоящее время реализуется в Московском Филиале ВСЕГЕИ. На ее основе издаются базы картографической и объектно-привязанных данных, использование методов распознавания образов на эталонных объектах, главным образом, при решении задач рудной геофизики. Отличительной особенностью ГИС ПАРК является использование критериев качества прогноза на основе функции риска $R = c_1 L + c_2 T$, где c_1 и c_2 – цены поисков по площади исследований и конкретного объекта, L и T , – параметры, характеризующие локальность и надежность прогноза. При этом надежность характеризуется долей пропущенных объектов, не вошедших при распознавании в качестве эталонных объектов. Локальность прогноза определяется риском пустых затрат по изучению явно безрудных площадей.

ГИС ИНТЕГРО является специализированной объектно ориентированной, направленной на решение широкого круга задач недр- и природопользования среди которых следует выделить:

- автоматизированное построение геологических карт и разрезов;
- автоматизированное построение легенд для листов от миллионного масштаба до двухсоттысячного;
- структурно-тектоническое районирование по атрибутам геофизических полей и их комплексного анализа;
- компьютерный прогноз твердых полезных ископаемых и подсчет их запасов;
- экологическое районирование;
- построение глубинных моделей земной коры по данным комплекса методов ГСЗ–КМПВ, МОВ–ОГТ, МТЗ, гравиразведки и аэромагниторазведки масштаба 1:200 000. С этой целью создана В.И. Галуевым ГИС ИНТЕГРО ГЕОФИЗИКА по построению глубинных моделей земной коры по сети региональных и опорных профилей, реализуемых на территории Русской платформы, Западной и Восточной Сибири [2].

Структурно-функциональная схема ГИС ИНТЕГРО приведена на рис. 11.1 [10], в которой выделена база геоданных в виде ГИС-конструктора, СУБД и приведены основные блоки решения прикладных задач.

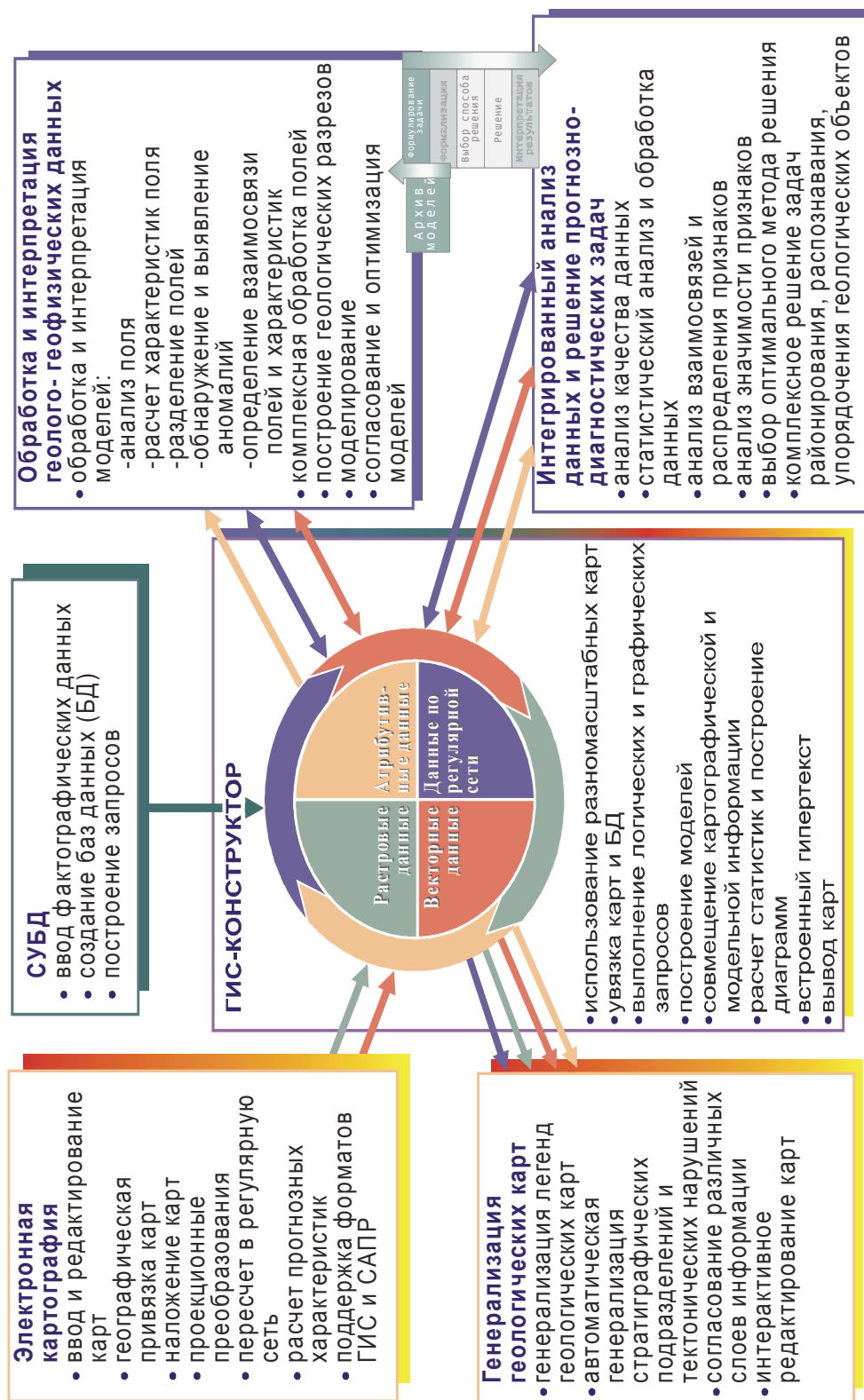


Рис. 11.1. Структурно-функциональная схема ГИС INTEGRO

Достаточно универсальной и простой при реализации является компьютерная технология КОСКАД-3D (А.В. Петров, А.А. Никитин), нашедшая широкое применение во многих учебных ВУЗах, научно-исследовательских и производственных организациях. КОСКАД-3D включает шесть основных разделов:

«Сервис» – обеспечивает выполнение стандартных функций системы управления базой данных. С их помощью осуществляется ввод/вывод содержательной информации, объединение и фрагментация сетей, восполнение отсутствующих в отдельных точках наблюдения значений признака, интерполяция сетей, различные преобразования с данными и т.д.

«Графика» – включает графический интерфейс, позволяющий оперативно просматривать одномерную, двумерную и трехмерную информацию из базы данных на экране дисплея в виде растровых карт, отдельных графиков, карт графиков и т.д.

«Статистика» – программы данного раздела предназначены для расчета статистических, спектральных и корреляционных характеристик геополей. Анализ этих характеристик позволяет получить дополнительную, полезную информацию об исследуемом поле и правильно выбрать граф его дальнейшей обработки.

«Фильтрация» – в модулях данного раздела реализованы наиболее распространенные в разведочной геофизике линейные оптимальные фильтры, позволяющие решать задачи разложения поля на составляющие, исключения тренда, оценки формы слабых аномалий. Особый интерес представляют уникальные адаптивные фильтры, позволяющие корректно обрабатывать нестационарные по спектрально-корреляционным характеристикам геофизические поля.

«Обнаружение» – с помощью программ данного раздела решается задача обнаружения слабых аномалий, соизмеримых по амплитуде с уровнем помех, линейной и изометричной формы, по одному или нескольким признакам.

«Комплекс» – использование программ этого раздела позволяет решать задачи разбиения анализируемой площади на однородные области (классы) с равными средними значениями признаков, распознавания комплексных аномалий по эталонной аномалии. Кроме этого возможно проведение компонентного анализа многопризнаковых данных.

Для программ этой группы в качестве входной информации могут быть использованы значения различных геолого-геофизических признаков и их производных, полученных с помощью программ из других разделов комплекса. В настоящее время существенно расширена область моделиро-

вания, включены модифицированные алгоритмы методов главных компонент и К-средних.

Система Mult Alt (Ф.М. Гольцман, Д.Ф. Калинин) целиком использует все преимущества статистического распознавания образов с обучением для решения прогноза и поисков рудных объектов.

Достаточно широкий комплекс алгоритмов комплексного анализа и комплексной интерпретации содержат система ПАНГЕЯ (ЗАО ПАНГЕЯ, В.В. Колесов и др.). ПАНГЕЯ завоевала рынок в Китае.

Методы количественной комплексной интерпретации геофизических полей реализованы главным образом в двух методных геоинформационных технологиях: сейсморазведка–гравиразведка; сейсморазведка–электроразведка, сейморазведка – геофизические исследования скважин, гравиразведка–магниторазведка.

Если первые три технологии решают в основном задачи прогноза нефтяных коллекторов и картирования осадочного чехла, то последняя предназначена для картирования кристаллического фундамента.

Компьютерная технология интерпретации данных сейсморазведки и гравиразведки

Система GCIS (геоинформационная комплексная информационная система), созданная А.И. Кобруновым и А.П. Петровским, основана на совместном решении обратных задач сейсморазведки и гравиразведки. При этом изучаемый объект задается в виде модели слоистой среды с априорными значениями скоростей и плотностей.

Построение согласованной сейсмоплотностной (согласование по скорости и по плотности) модели среды осуществляется путем минимизации обобщенного функционала $\Phi(\xi, f)$ на основе интерактивного подбора. При этом предполагается, что минимум такого функционала $\Phi(\xi, f)$ реализуется при условии соответствия с одной стороны, скоростной модели волновому полю отражений ОГТ, а с другой стороны, плотностной модели – гравитационному полю наблюдений. При каждом фиксированном векторе параметров среды индекс ξ означает, что «сейсмические» границы ξ зафиксированы и вошли в конструкцию функционала. Аналогично при каждом фиксированном f функционал $\Phi(\xi, f)$ представляет собой критерий оптимальности на множестве эквивалентности для плотностных границ. Итерационный процесс состоит в последовательном нахождении решений, при котором результат, полученный при интерпретации данных гравиразведки, передается для конструирования функционала по данным сейсморазведки и последующего нахождения решения, далее

найденное решение для сейсморазведки передается для конструирования функционала по данным гравиразведки и так далее. Описанный итерационный процесс сходится. Его результатом является пара слоистых моделей – скоростная и плотностная, максимально близких между собой насколько это допустимо согласованностью волнового и гравитационного полей. Важным элементом приведенной технологии является параметризация модели среды, определяющая допустимые виды моделей в данной конкретной геологической ситуации, выражаемые через варьируемые в процессе решения обратной задачи параметры среды. Параметром служит одна единственная функция – функция Лагранжа, через которую и критерий оптимальности выражаются все границы среды. В рамках параметризации используется аппроксимация среды наборами призм, что обеспечивает расчет прямой задачи. Технология GCIS развивается путем математического моделирования динамики формирования изучаемого объекта, результаты которого являются последовательность статистических моделей, а параметризация заключается в оценке параметров по управлению динамикой процесса формирования углеводородов.

Компьютерная технология комплексной интерпретации данных сейсморазведки и электроразведки по методу зондирования становлением поля в ближней зоне (ЗСБ) создана в НВНИИГГ и ЗАО «Геонефтегаз» под руководством Н.П. Смилевец. Эта технология реализует построение согласованных сейсмoeлектрических временных разрезов (СЭВР) при решении прогнозно-поисковых задач нефтегазовой геологии. В ее основе заложено использование практически функциональной зависимости между t_0 временных сейсмических разрезов и временем становления поля $t_{ЗСБ}$, а именно $t_0 = a\sqrt{t_{ЗСБ}}$. Величина коэффициента связи « a » меняется от 0,5 до 0,8 для различных типов пород. Коэффициент « a » является функцией средней скорости и удельной электропроводности горизонтально-слоистой модели геологического разреза. Переход к глубинам контактных поверхностей разреза осуществляется на том основании, что в сейсморазведке аналогом глубины является время отраженных волн t_0 , а в электромагнитной разведке аналогом кажущейся глубины исследований – соответственно величина $\sqrt{t_{ЗСБ}}$. Для высокоскоростных и высокоомных отложений (это модели соляных куполов, сложенных породами сульфатно-карбонатного состава) отмечается возрастание коэффициента « a » с глубиной исследований, при этом « a » равен 0,6. Для разрезов, представленных терригенными (проводящими) отложениями, величина коэффициента « a » составляет 0,8.

Необходимым условием для достоверного построения согласованных сейсмоэлектрических временных разрезов является совпадение электромагнитных зондирований с сейсмическими пикетами по линии общего профиля. Относительные погрешности определения коэффициента «а» составляют 5-10% в интервале глубин 2-5 км, максимальные значения погрешностей «а» отмечаются в пределах склона соляных куполов и в мульдах, которые достигают до 20-30%.

По различным типам зависимостей $a = f(\sqrt{t_{3CB}})$ осуществляется прогноз литологии и нефтегазоносности коллекторов при поисково-разведочных работах, а по данным зависимости $t_0 = a\sqrt{t_{3CB}}$ – более точная привязка по глубине залегания горизонтов (контактных поверхностей).

Развитием технологии СЭВР является технология COMINTER, включающая программно-алгоритмическое обеспечение по преобразованию данных электромагнитных и потенциальных (гравитационных и магнитных) полей в масштаб временного сейсмического разреза. Это позволяет в едином координатном пространстве (x, t) обеспечить построение согласованного по физическим параметрам (скорости, электропроводности, плотности) временного разреза с целью более надежного прогноза литологии и нефтегазоносности коллекторов. Пример построения такой согласованной физико-геологической модели приведен на рис. 8.6. В связи с технологией COMINTER, предназначенной в основном для поисков углеводородов на глубинах 2-3,5 км. В связи с этим отметим технологию **ГИС ИНТЕГРО ГЕОФИЗИКА**, рассмотренную в разделе 6.5 и предназначенную для построения согласованных по физическим и геометрическим параметрам моделей земной коры до глубин 40 км и более по региональным опорным профилям.

Компьютерные технологии интерпретации данных сейсморазведки и геофизических исследований скважин

В России используются такие технологии, как Geo Frame, TIGRESS, LANDMARK и другие. Все указанные технологии для прогноза вещественного состава включают данные сейсмических трасс.

Программные пакеты сейсмической интерпретации SHARISMA и IESX полностью интегрированы с другими пакетами системы Geo Frame. Поэтому, независимо от того, какая из интерпретационных процедур используется, геолог-геофизик имеет возможность плавно переходить из одной области в другую, и от одной функции к обработке другой.

База данных системы Geo Frame обеспечивает отдельный доступ и содержит большую библиотеку прикладных программ, которая включает данные, связанные с процессами разведки и разработки месторождений. Система реализует широкий набор средств управления данными, что позволяет пользователю осуществить легкий доступ и работу с БД, а также поддерживает все стандартные форматы, используемые в каротаже.

Система Geo Frame содержит программные модули, необходимые для геологической и петрофизической интерпретации: Well Edit – редактирование скважинных данных; Well Composite – обобщение всех данных по скважине; Petro View Plus – определение петрофизических параметров; Start Log – построение стратиграфической колонки; Rock Class – классификация литологического состава.

С помощью Well Edit контролируется качество данных, реализуется просмотр и редакция каротажных диаграмм по глубине. С помощью Rock Class формируется стратиграфическая колонка на основе статистической классификации данных каротажа, расчета минерального состава или данных исследования керна.

Геологическая интерпретация начинается с помощью взаимосвязанных пакетов Well Plus и Start Log, которые обеспечивают обработку данных – корреляцию каротажных диаграмм, составление геологических разрезов и картирование горизонтов.

Картирование в системе Geo Frame реализуется в модуле CPS-3, содержащем алгоритмы для моделирования поверхностей в нефтегазовой геологии.

В настоящее время на рынке интерпретационных продуктов сейсморазведки и ГИС господствующее положение занимает система Petrel фирмы Schlumberger.

Отечественной разработкой при проведении обработки и комплексной интерпретации всех видов геолого-геофизических данных по скважинам и сейсморазведки на единой информационной основе является система ГЕММА (И.М. Чуринова и др. ОАО «Центральная геофизическая экспедиция»). Система ГЕММА решает задачи площадной обработки и комплексного анализа данных геофизических исследований скважин (подсистема ГИС), анализов результатов исследования керна (подсистема КЕРН), интерпретации данных сейсморазведки (подсистема СЕЙС), вертикального сейсмического профилирования (подсистема ВСП) и гидродинамических исследований скважин (подсистема ГДИ), построения геологической модели (подсистема МОДЕЛЬ). База данных системы содержит каротажную, геологическую, собственно промысловую, керновую и сейсмическую информатическую, собственно промысловую, керновую и сейсмическую информацию. Ее СУБД

SYBASE обеспечивает оперирование большими объемами разнородной информации, работу в сети многих клиентов и функционирование на компьютерах разных платформ. Сервер SYBASE, БД и управляющие ядром системы расположены на RISC-компьютерах.

Для комплексной интерпретации результатов анализа керна и данных каротажа используются средства математической статистики (гистограммные преобразования), корреляционно-регрессионного анализа и классификационные алгоритмы с построением разрезов скважин. Комплексная интерпретация данных геофизических исследований скважин и сейсморазведки реализуется в подсистеме ИНПРЕСС. Построение геологических моделей геосреды и месторождений и их объемная визуализация осуществляется с помощью программных средств динамической визуализации DV.

Интегрирование скважинной и сейсморазведочной информации с целью оценки литофациальной зональности, тектонических нарушений и прогноза коллекторских свойств во внескважинном пространстве реализуется в компьютерной технологии ИНТЕГРАН (ВНИИгеосистем С.А. Каплан).

Входная информация в этой технологии представлена временными разрезами ОГТ (до и после миграции), ВСП, стратиграфической приуроченностью опорных отражений, данными геофизических исследований скважин.

Комплексирование скважинной и сейсморазведочной информации носит итерационный характер в зависимости от получаемых результатов на каждом этапе комплексной интерпретации. Технология ИНТЕГРАН включает семь основных этапов.

Технология ИНТЕГРАН успешно была адаптирована к задачам исследований на опорных геофизических профилях (геотраверсах), а также при обработке волновых картин акустического каротажа. Она получила широкую апробацию на ряде крупных месторождений нефти и газа, включая Астраханское газоконденсатное (Прикаспийская провинция), Оренбургское газоконденсатное, Теректинское (Волго-Уральская провинция), Ново-Уренгойское, Сугмутское, Кынское Бованенковское (Западно-Сибирская провинция), что обеспечило ее внедрение в геологические организации России: Ямалгеофизика, Оренбург-геология, Восток-геология и др.

Следует подчеркнуть, что сравнение технологий ГЕММА, ПАРМКОЛЛЕКТОР, ИНТЕГРАН (в части решения нефтяных задач) и других подобных систем, например, технологий СВАН и НЕДРА (СВАН – спектрально-временной анализ для картирования типов геологического разреза – авторы Е.А. Давыдова, Е.А. Копилевич, И.А. Мушин, НЕДРА – технология интерпретации наземно-скважинных измерений Краснодарской опытно-методической экспедиции «Союзморгео» – автор Д.П. Земцова) по комплексной интерпрета-

ции данных сейсморазведки и геофизических исследований скважин имеют много общего в постановке задачи, различаясь использованием тех или иных сейсмических атрибутов, а также критериями по увязке сейсмических и каротажных данных и прогнозу коллекторских свойств.

Следует отметить в ряде технологий использование новых математических приемов: вейвлет-анализа, нейронных сетей и генетических алгоритмов (Petrel ГИС ИНТЕГРО, КОСКАД-3Д, ИНТЕГРАН, DV-Seis Geo и др.).

Компьютерные технологии геологического моделирования залежей углеводородов семейства DV

В настоящее время в разведочной геофизике активно разрабатываются технологии трехмерного цифрового моделирования месторождений полезных ископаемых, и в первую очередь, залежей углеводородов (УВ). Цифровые геологические модели – это численные представления разномасштабных геологических объектов на геометрических сетках в трехмерном и четырехмерном пространствах. Основные достижения по созданию таких моделей принадлежат разработчикам технологии семейства DV (Dynamic visualization) в ОАО «Центральная геофизическая экспедиция» под руководством А.С. Кашика, С.И. Билибина и др. [10].

В отличие от главных мировых производителей трехмерных моделей Schlumberger, Halliburton (Landmark), Roxar, Paradigm Geophysical в области нефтегазовой индустрии в технологии семейства DV, наряду с координатами X, Y, Z введен четвертый параметр T – время геологических эпох.

Технологии всех указанных производителей при построении цифровых моделей базируются на комплексной интерпретации данных детальных сейсмических наблюдений и данных каротажа. Возможности компаний Schlumberger и Hallibarton были отражены выше.

Технологии семейства DV представлены рядом программных модулей:

- DV1-Discovery – детальный анализ и интерпретация сейсмических данных совместно с данными геофизических исследований скважин по созданию детальной модели геосреды.
- Геология DV-Geo – построение трехмерных геомodelей для подсчета запасов и подготовки данных для гидродинамического моделирования.
- Сейсмогеология DV-SeisGeo – создание цифровой объемной структурно-параметрической модели на основе интерпретации 2D/3D сейсмических, скважинных, геологических, петрофизических и промысловых данных; палеотектонический и палеоседиментационный анализ, подсчет запа-

сов, автоматическая корреляция сейсмических и скважинных данных, трассирование нарушений.

Характеристика систем геологического моделирования семейства DV

Программные комплексы DV: DV1-Discovery, DV-Geo и DV-SeisGeo являются многофункциональными и высокотехнологическими при организации работы с большими и разнородными массивами геолого-геофизических данных.

Уникальность математического обеспечения технологии DV для моделирования залежей УВ и геологических процессов состоит в использовании четырехмерной метрики XYZT. Комплексы DV «динамического видения», визуализации отличаются оригинальностью инженерного решения, при котором доступ к данным куба исходных данных или их атрибутов при движении слайса, т.е. трехмерного куба сечения (XYZ; XYT; XZT; YZT), осуществляется в режиме реального времени. Результатом является сплошное сканирование исследуемой среды, что в десятки и сотни раз увеличивает объем информации, который предоставляется пользователю. Динамическая визуализация играет ключевую роль при поиске структурных ловушек УВ, корреляции отражающих горизонтов и трассировании тектонических нарушений.

Основные типы данных, используемые в технологии семейства DV при построении объемных геологических моделей, приведены на рис. 11.2. В качестве примера визуализации данных в DV на рис. 11.3 приведены результаты по анализу сейсмических данных, а на рис. 11.4 – результаты комплексного анализа сейсмических атрибутов и петрофизических параметров. Схема основных технологий DV1-Discovery изображена на рис. 11.5.

В заключение подчеркнем, что автоматизированные и геоинформационные системы решают хорошо формализованные задачи, когда решение прямых и обратных задач доводится до числа.

Решение слабоформализованных задач в середине 80-х годов XX века привело к созданию так называемых экспертных систем, а решение неформализованных задач уже в XXI веке к информационно-аналитическим системам (ИАС). Экспертные системы по своей сути близки к методам распознавания образов с использованием эталонных объектов и их использование в настоящее время наблюдается лишь в нефтегазовой отрасли.

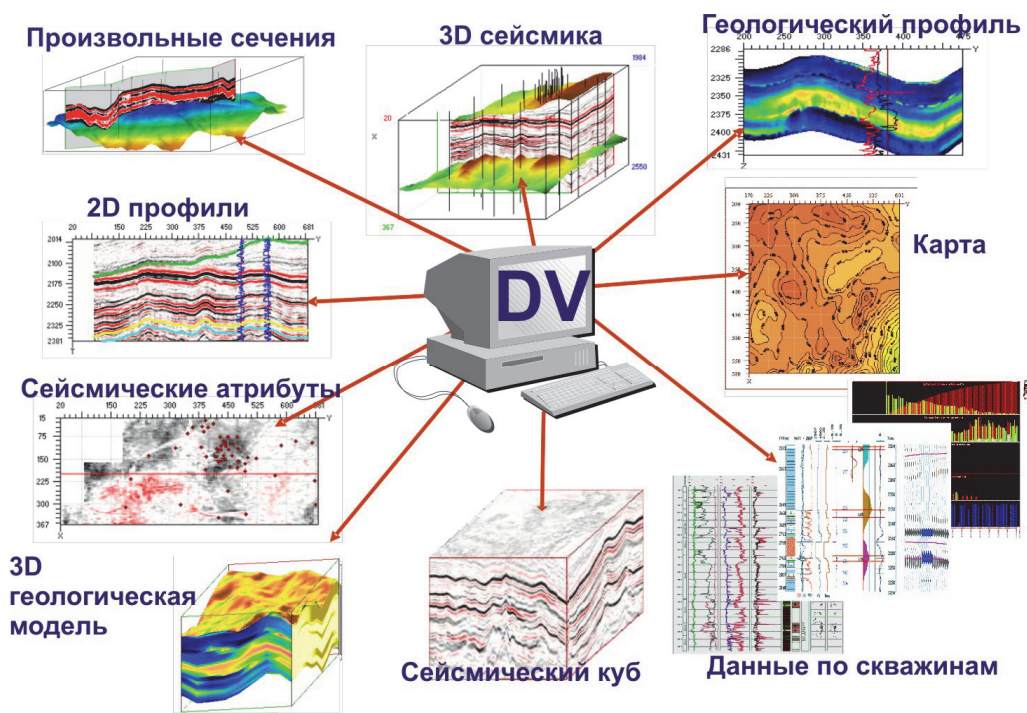


Рис. 11.2. Типы данных, используемых для построения трехмерных геологических моделей в проектах DV

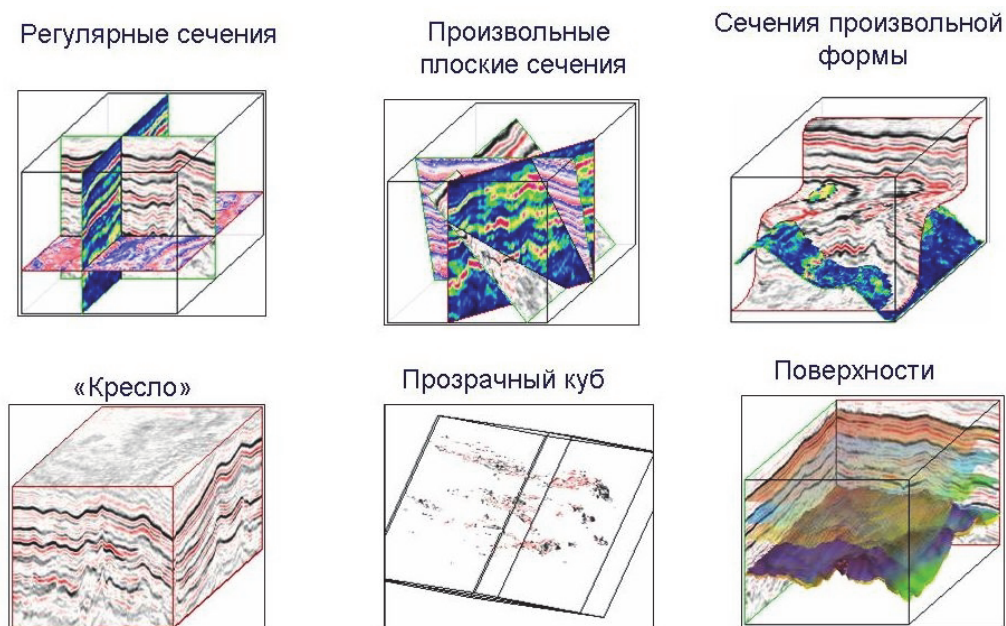


Рис. 11.3. Динамический визуальный анализ сейсмических данных

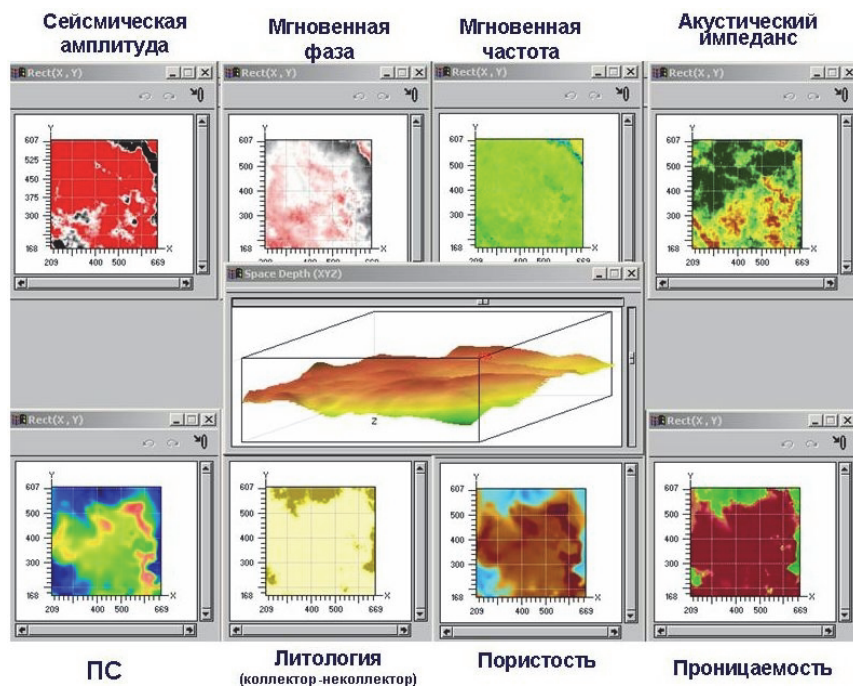


Рис. 11.4. Комплексный анализ сейсмических атрибутов и петрофизических параметров на поверхности

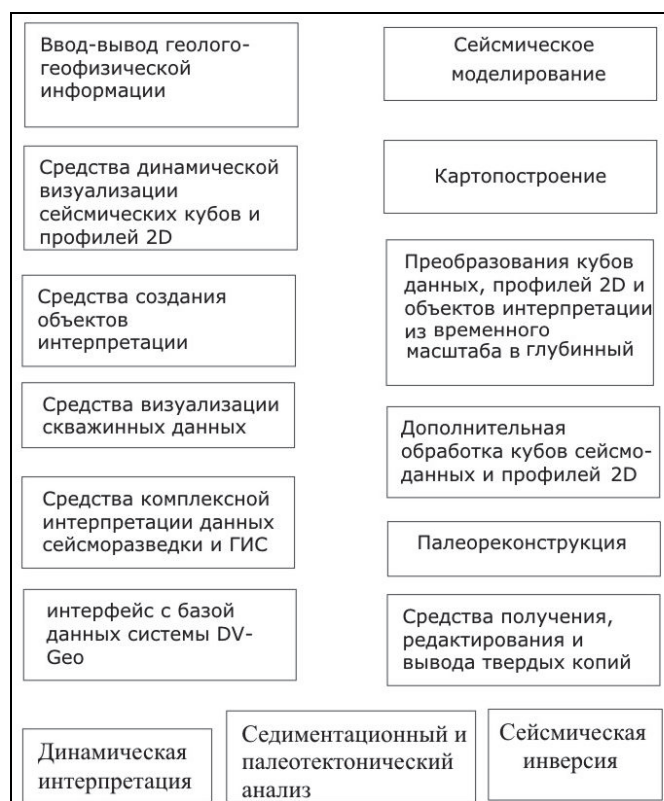


Рис. 11.5. Схема основных блоков пакета DV1-Discovery

Информационно-аналитические системы применяются в задачах управления крупными регионами, министерствами. С этой целью создаются так называемые Ситуационные центры.

В целом развитие программно-аналитического обеспечения при геофизических исследованиях можно представить в виде схемы на рис. 11.6.

В экспертных системах решение задач прогноза в настоящее время базируется главным образом на байсовском подходе.

В информационно-аналитических системах решение принимается на основе количественных критериев системного анализа.

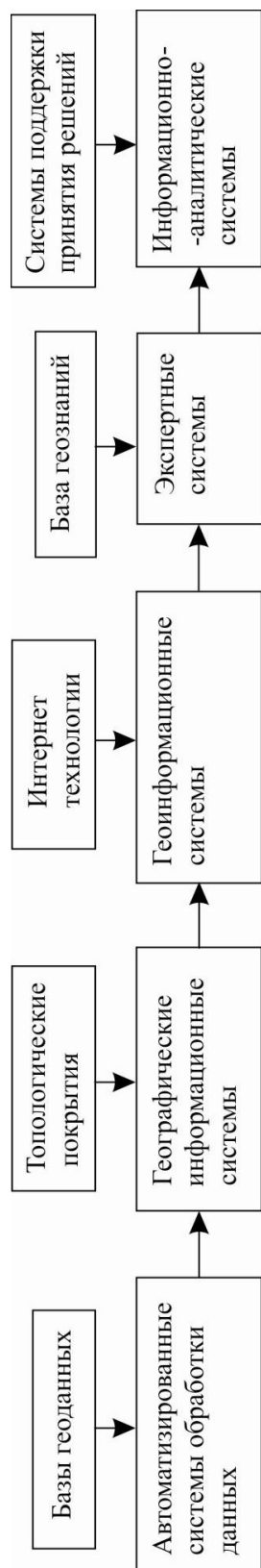


Рис. 11.6. К развитию геоинформационных систем

Контрольные вопросы по комплексному анализу и комплексной интерпретации геофизических данных

1. Охарактеризуйте основные процедуры комплексного анализа при наличии эталонных объектов.
2. Как оцениваются вероятность прогноза искомых объектов при байесовском подходе.
3. Каким образом оценивается информативность комплекса признаков (полей) в задачах поисков и геокартирования.
4. Назовите основные приемы комплексного анализа геополей при отсутствии эталонных объектов.
5. Каким образом оценивается ошибка при распознавании образов при наличии эталонных объектов и при их отсутствии.
6. В чем состоит суть метода главных компонент.
7. Каким образом можно учесть корреляционные связи между различными геопольями (или их атрибутами) при площадных наблюдениях.
8. Приведите математическую постановку задачи количественной комплексной интерпретации геофизических полей при детерминированном подходе.
9. В чем смысл математической постановки задачи количественной комплексной интерпретации с позиции вероятностно-статистического подхода.
10. Каким образом оценивается адекватность построенных комплексных ФГМ реальной среде.
11. Приведите основные процедуры стандартного графа обработки и интерпретации гравитационного и магнитного полей при построении глубинных моделей земной коры.
12. В чем состоит физический смысл признаков «прямого» прогноза углеводородов по волновому и потенциальным полям.
13. На чем основано построение согласованных сейсмоэлектрических разрезов (СЭВР).
14. Что собой представляют DV-технологии.

ЛИТЕРАТУРА

Основная

1. Соколова Т.Б., Булычев А.А., Лыгина И.В. и др. Интерпретация геофизических материалов. Тверь, ГЕРС, 2011, 208 с.
2. Голиздра Г.Я. Комплексная интерпретация геофизических полей при изучении глубинного строения земной коры. М. Недра. 1988, 212 с.
3. Гольцман Ф.М. Статистические модели интерпретации. М. Недра, 1971, 327 с.
4. Никитин А.А., Хмелевской В.К. Комплексирование геофизических методов. Учебник для ВУЗов. М. ГНЦ РФ ВНИИгеосистем, 2012, 344 с.

Дополнительная

5. Галуев В.И., Каплан С.А., Никитин А.А. Технология создания физико-геологических моделей земной коры по опорным профилям на базе геоинформационных систем. ГНЦ РФ ВНИИгеосистем, 2009, 228 с.
6. Кобрунов А.И. Математические основы теории интерпретации геофизических данных. М. ЦенрЛитНефтегаз, 2008, 288 с.
7. Никитин А.А. Энергетическая фильтрация геофизических полей. НТЖ ГЕОФИЗИКА, № 3, 2014, с. 2-13.
8. Никитин А.А., Петров А.В. Теоретические основы обработки геофизической информации. Учебное пособие. ГНЦ РФ ВНИИгеосистем, 2013, 117 с.
9. Приезжев И.И. Уточнение геологической модели по данным гравитационного поля на основе критериальных методов решения обратных задач геофизики. НТЖ ГЕОФИЗИКА, 2010, № 11, с. 65-68.
10. Черемисина Е.Н., Никитин А.А. Геоинформационные системы и технологии. ГНЦ РФ ВНИИгеосистем, 2011, 375 с. Учебник для ВУЗов.
11. www.Sigma3d.com.

ПРИЛОЖЕНИЕ

Вейвлет-анализ геофизических полей

Вейвлет-анализ представляет разложение исходных данных по системе заданных по форме сигналов и реализуется путем линейной свертки входных значений поля с весовыми функциями, в качестве которых выступают заданные по форме сигналы. Система заданных по форме сигналов или вейвлетов (*«вейвлет»* – всплеск или небольшая волна, этот термин возник из применения вейвлет-анализа в самом начале его развития при обработке данных сейсморазведки) представляет совокупность таких функций (сигналов), каждая из которых является сдвинутой (по времени) и масштабируемой, т.е. сжатой или растянутой копией одной и той же функции, так называемого порождающего вейвлета. Для того, чтобы функция (сигнал) $\psi(t)$ называлась вейвлетом должны выполняться два условия: среднее значение $\psi(t)$ равно нулю и $\psi(t)$ должна быстро убывать при $t \rightarrow \pm\infty$, т.е. говорят, что всплеск (вейвлет) компактен в пространстве и локализован по частоте.

Вейвлет-анализ является обозначением большого класса разложений, поскольку существующие виды порождающих вейвлетов достаточно сильно отличаются друг от друга своими определениями, свойствами и приложениями. В определенном смысле вейвлет-анализ подобен Фурье-анализу с его разложением исходных значений по системе синусов и косинусов. Однако вейвлет-анализ имеет два существенных отличия от Фурье-анализа. Во-первых, Фурье-анализ не различает сигналы из двух синусоид с разными частотами, один из которых является суммой синусоид, а второй представляет следующие друг за другом последовательно синусоиды. В обоих случаях спектры таких сигналов представлены двумя пиками на фиксированных частотах этих синусоид. Во-вторых, Фурье-анализ слабо приспособлен для обработки нестационарных сигналов, в том числе локализованных на некотором временном интервале, поскольку в спектре теряется информация о временных характеристиках сигнала, т.е. вейвлет-анализ реализует спектральный анализ одновременно и по частоте и по времени. Порождающие или базисные функции вейвлет-анализа в отличие от Фурье-анализа обладают частотно-временной локализацией. Базисной функцией в Фурье-анализе является $e^{j\omega t}$. Масштабирование базисной функции по времени осуществляется в вейвлет-анализе умножением (или делением) ее абсциссы на число, т.е. $\psi(t) = \psi(at)$, что приводит к сжатию или растяжению базисной функции, т.е. заданного по форме сигнала, то же самое касается масштабирования спектра по частоте.

В практике обработки геофизических данных получил применение непрерывный вейвлет-анализ. **Непрерывным вейвлет-анализом** (преобразованием) исходной функции $f(t)$ называют функцию двух переменных $F(\omega, b)$. При этом вводится базис (порождающий вейвлет), отвечающий условиям равенства его среднего значения нулю и быстрого затухания по времени в виде:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right), \quad (1)$$

где множитель $\frac{1}{\sqrt{|a|}}$ требуется для сохранения масштабов.

Если a и b – произвольные вещественные значения, тогда пара преобразований непрерывного вейвлет-анализа принимает вид:

$$w_f(a, b) = \langle f(t), \psi_{a,b}(t) \rangle = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} f(t) \psi\left(\frac{t-b}{a}\right) dt, \quad (2)$$

$$f(t) = C_{\psi}^{-1} \int_{-\infty}^{\infty} \frac{da}{a^2} \int_{-\infty}^{\infty} w_f(a, b) \psi\left(\frac{t-b}{a}\right) db,$$

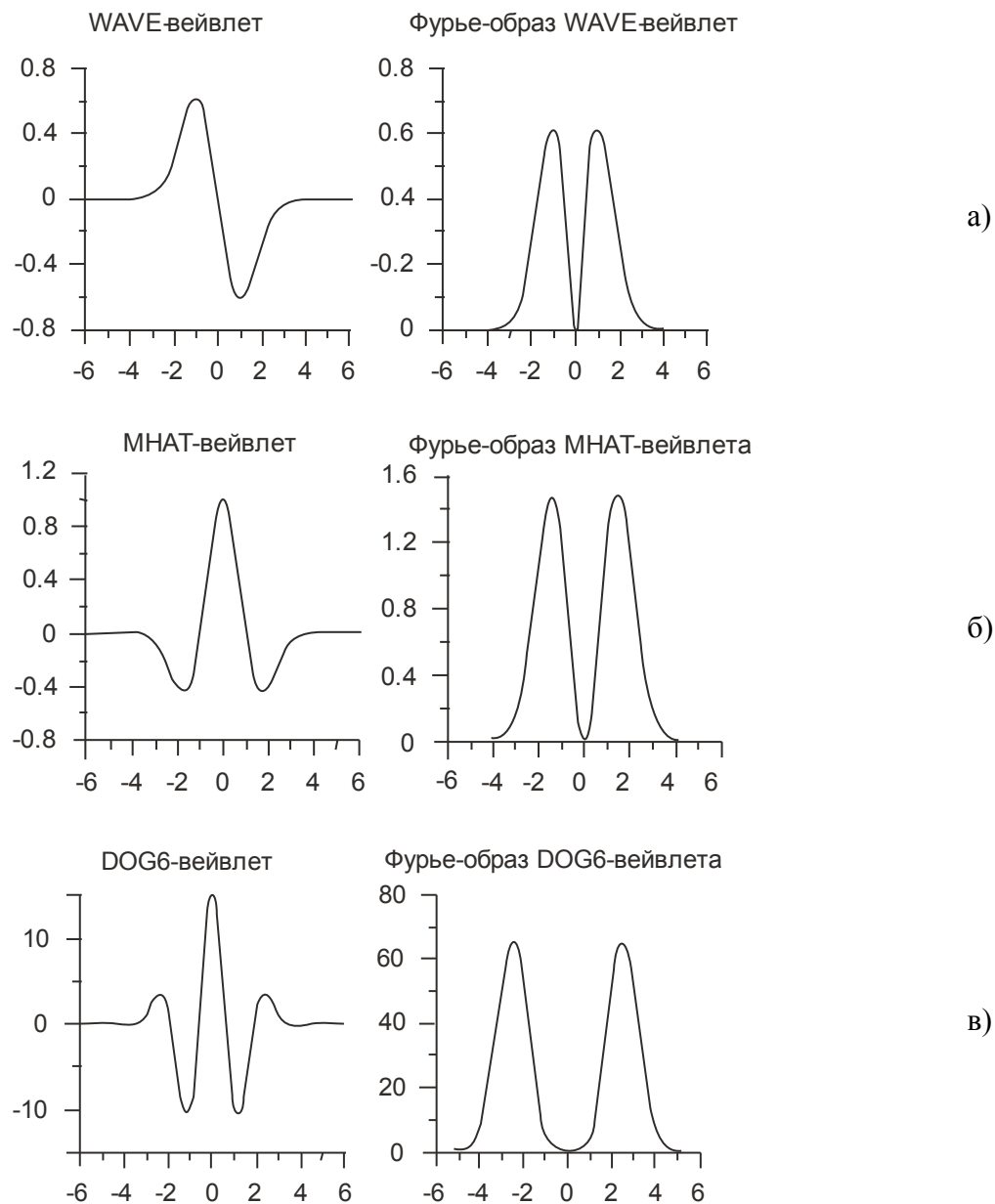
где $C_{\psi} = \int_{-\infty}^{\infty} \frac{|\psi(\omega)|^2}{|\omega|} d\omega$ – нормализующий множитель, $\langle f(t), \psi_{a,b}(t) \rangle$ – скалярное произведение исходного сигнала $f(t)$ с материнским вейвлетом $\psi(t)$, представленное их линейной сверткой.

Сравнивая формулы (2) с парой преобразований Фурье, легко видеть, что в преобразованиях Фурье роль функции $\psi_{a,b}(t)$ играет функция $e^{j\omega t}$, а C_{ψ} аналогичен коэффициенту 2π , роль частоты играет масштабный множитель $\frac{1}{a}$. Но базисная функция $\psi\left(\frac{t-b}{a}\right)$ зависит еще от параметра ее сдвига по времени b . Следовательно, для каждой пары значений a и b функция $w_f(a, b)$ определяет амплитуду соответствующего вейвлета. В Фурье-анализе каждой частоте соответствует всего одна гармоническая составляющая, а для вейвлет-анализа каждой частоте соответствует множество сдвинутых друг относительно друга функций.

В геофизике широко используются порождающие вейвлеты на основе производных функции Гаусса: $\psi_m(t) = (-1)^m \frac{d^m}{dt^m} \exp(-t^2/2)$, как хорошо локализованной и по частоте и по времени.

На рис. 1 приведены WAVE-вейвлет при $m = 1$ (а), МНАТ-вейвлет при $m = 2$ (б), DOG-вейвлет, как разность двух гауссиан (в), Morlet-вейвлет: $\psi(t) = \exp(j\omega_0 t)e^{-t^2/2}$ (г) и Paul-вейвлет при $m = 4$.

Справа от этих вейвлетов представлены их Фурье-образы. Морле-вейвлет наиболее распространен при обработке данных сейсморазведки. МНАТ-вейвлет – при обработке данных электромагнитных зондирований.



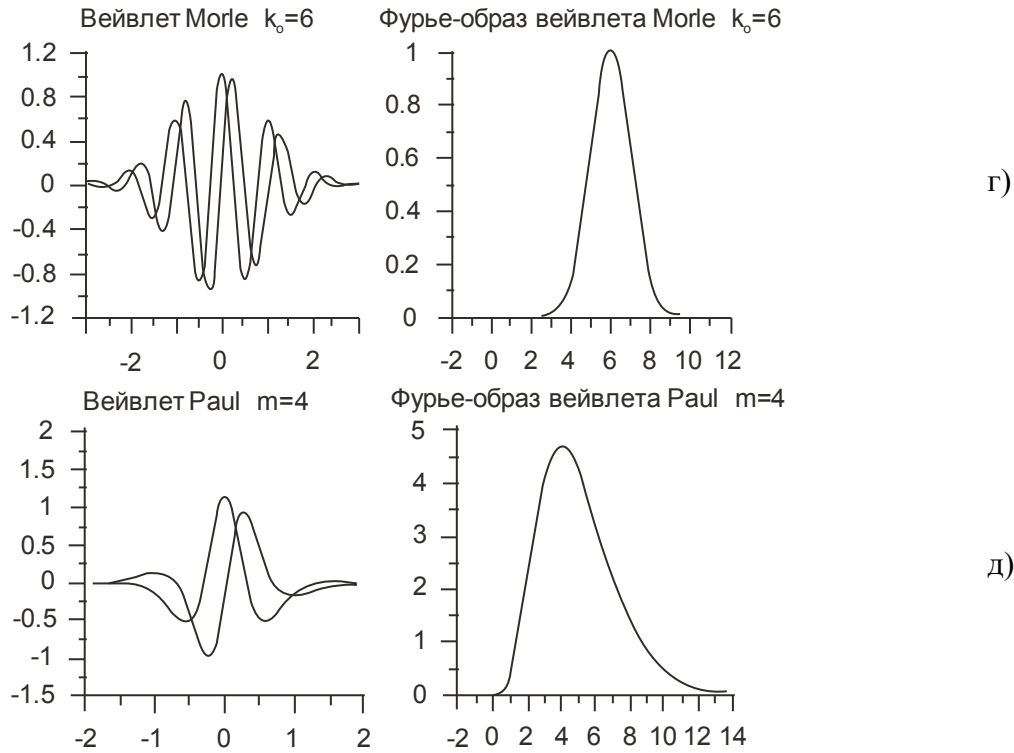


Рис. 1. Вейвлеты в зависимости от времени и их образы Фурье:

а – WAVE, б – MHAT, в – DOG с $m = 6$, г – Morlet с $k_0 = 6$,

д – Paul с $m = 4$ соответственно

При обработке данных сейсморазведки обычно используется Морле-вейвлет.

При анализе аномальных гравитационных и магнитных полей логично использовать вейвлеты, построенные, как частные производные ядра Пуассона:

$$K^{nx,mz}(x,z) = \frac{\partial^{n+m}}{\partial^n x \partial^m z} \left(\frac{z}{x^2 + z^2} \right).$$

Полученные таким образом вейвлеты будут удовлетворять необходимым условиям равенства их среднего значения нулю и компактности. Так вейвлеты, соответствующие вычислению первых частных производных поля (вейвлеты 1-го порядка) на высоте h , будут иметь вид:

$$\psi_0^z(\xi) = \frac{1 - \xi^2}{(\xi^2 + 1)^2}, \quad \psi_{x,h}^z(\xi) = \frac{1}{\sqrt{h}} h^2 \frac{h^2 - (\xi - x)^2}{[(\xi - x)^2 + h^2]^2}, \quad (3)$$

$$\psi_0^x(\xi) = \frac{2\xi}{(\xi^2 + 1)^2}, \quad \psi_{x,h}^x(\xi) = \frac{1}{\sqrt{h}} h^2 \frac{2(\xi - x)h}{[(\xi - x)^2 + h^2]^2}. \quad (4)$$

В формулах (1 – 2) параметр a заменен на h (в формулах 3 – 4) в качестве параметра глубины, а t на ξ .

Вейвлеты $\psi_0^z(\xi)$ и $\psi_{x,h}^z(\xi)$ соответствуют вычислению 1-ой вертикальной производной, а вейвлеты $\psi_0^x(\xi)$ и $\psi_{x,h}^x(\xi)$ – 1-ой горизонтальной производной исходной функции. Очевидно, что при заданном параметре h функция $W(x,h)$, рассчитанная с помощью вейвлетов (3, 4) связана с первой частной производной поля $g(\xi)$ на высоте соотношением:

$$g_1(\xi, h) = \frac{1}{\pi} \frac{\sqrt{h}}{h^2} W(\xi, h). \quad (5)$$

Здесь под функцией g_1 понимается функция g_x или g_z , соответствующая первой частной производной функции g по параметру x или z .

Аналогично (3, 4) могут быть представлены вейвлеты Пуассона более высоких порядков. Отметим, что для анализа двумерных потенциальных полей удобно использовать комплексный вейвлет Пуассона.

На основании формулы (2) по полученному вейвлет-спектру $W(x,h)$ возможна реконструкция исходного сигнала $g(\xi)$. Это в свою очередь позволяет использовать вейвлет-преобразования для решения таких “классических” задач, как фильтрация исходных сигналов, редуцирование двумерных магнитных аномалий к полюсу, продолжение поля и вычисление высших производных полей в верхнем и нижнем полупространствах.

В то же время для задач интерпретации аномальных гравитационных и магнитных полей представляет интерес анализ поведения самого вейвлет-спектра $W(x,h)$. Это связано с тем, что характер его поведения отражает положение особых точек потенциальных полей и распределение избыточных масс и намагниченностей. Так, например, можно показать, что по вейвлет-спектру $W(x,h)$ аномального гравитационного поля g , рассчитанному вейвлетом Пуассона $\psi_{x,h}^z(\xi)$, можно определить эквивалентное распределение плотности, создающее исходное поле:

$$\delta(x, h) = \frac{1}{\pi^2 G} \frac{\sqrt{h}}{h^2} W(x, h).$$

Применение различных типов вейвлетов позволяет подчеркнуть различные особенности анализируемого поля и их связь с источниками.

Для анализа площадных аномальных полей аналогично одномерному можно определить двумерное вейвлет-преобразование на основе ядра Пуассона. В то же время такое преобразование будет иметь некоторые особенности по сравнению с одномерным. В частности восстановление исходного сигнала будет возможно только для осесимметричного вейвлета.

Учебное издание

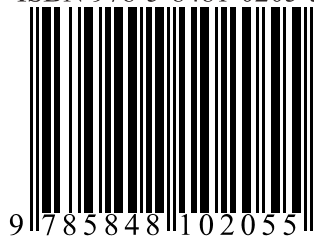
**Булычев Андрей Александрович,
Никитин Алексей Алексеевич**

**Комплексный анализ и комплексная интерпретация
геофизических полей**

Учебное пособие

Издано в авторской редакции

ISBN 978-5-8481-0205-5



Подписано в печать 12.11.2015.
Формат 60х90/8. Усл. печ. л. 10,5.
Тираж 100 экз. Заказ № 34.

Отпечатано в Московском филиале ФГУНПП «Росгеолфонд»
«ВНИИГеосистем»
117105, Москва, Варшавское шоссе, 8
Тел. (495) 952-21-57. E-mail: artur@geosys.ru