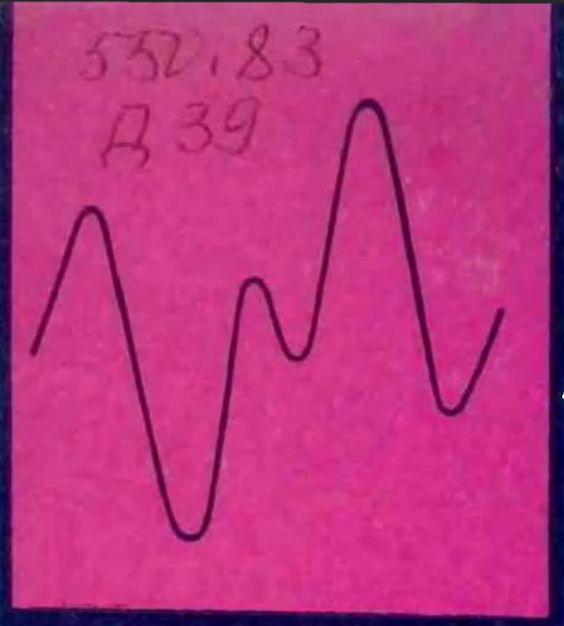
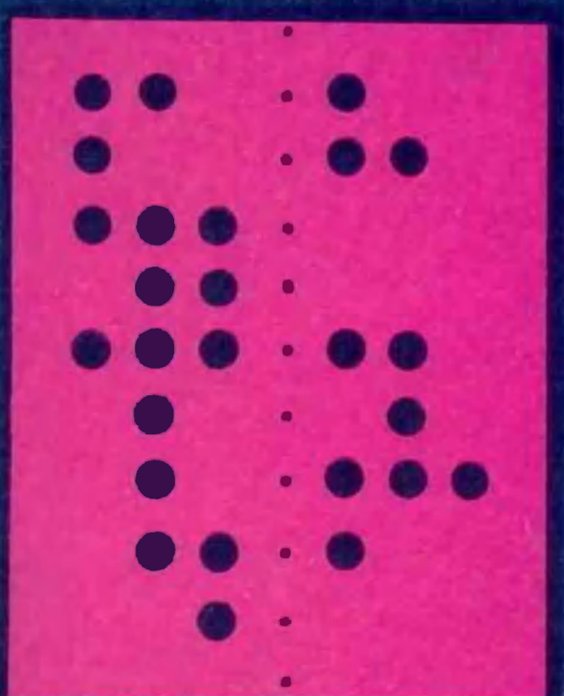
532.183
Д 39



В. Н. ДЕЧ
Л. Д. КНОРИНГ

НЕТРАДИЦИОННЫЕ МЕТОДЫ КОМПЛЕКСНОЙ ОБРАБОТКИ И ИНТЕРПРЕТАЦИИ ГЕОЛОГО-ГЕОФИЗИЧЕСКИХ НАБЛЮДЕНИЙ В РАЗРЕЗАХ СКВАЖИН



УДК 550.837 : 622.241 : 519.242.32

Деч В. Н., Кноринг Л. Д. Нетрадиционные методы комплексной обработки и интерпретации геолого-геофизических наблюдений в разрезах скважин. Л., «Недра», 1978. 192 с.

Рассмотрены способы комплексной обработки и интерпретации материалов геолого-геофизических исследований глубоких скважин с целью изучения геологического строения разрезов, оценки коллекторских свойств пород, выделения нефте- и газоносных пластов. Особое внимание уделено выявлению и количественной оценке природных факторов, определивших изменчивость состава и физических свойств отложений.

В основу интерпретации положены методы математической теории оптимального эксперимента (МТОЭ). Описаны математический аппарат этих методов и соответствующие модели интерпретации. Показано, какие методы наиболее приемлемы; определены условия рационального использования конкретных методов; оценены возможности и эффективность методов МТОЭ, указаны пути их реализации на ЭВМ.

Приведены примеры интерпретации материалов, характеризующих разрезы отложений разного возраста и географического положения в условиях, при которых традиционные методы не всегда оказывались приемлемыми и результативными.

Книга предназначена для промысловых геологов и геофизиков.

Табл. 32, ил. 23, список лит. 96 назв.

Д $\frac{20804-353}{043(01)-78}$ 126—78

© Издательство «Недра», 1978

ОГЛАВЛЕНИЕ

Предисловие 4

Часть первая

Изучение осадочных пород

Глава I. Оценка коллекторских и литолого-петрографических характеристик 13

Постановка и содержание задачи . . . . . —
Методы решения и примеры их использования 14
Обсуждение результатов 48

Глава II. Выявление факторов, определяющих изменчивость состава и физических свойств осадочных пород 49

Постановка и содержание задачи . . . . . —
Основы компонентного анализа . . . . . —
Изучение разрезов карбонатных толщ 55
Обсуждение результатов 87

Глава III. Диагностика газоносных и водоносных пластов . . . 91

Постановка задачи и методы ее решения . . . . . —
Последовательный статистический анализ 97
Дискриминантный анализ 104
Обсуждение результатов 107

Часть вторая

Изучение разрезов осадочных толщ на основе анализа структуры геолого-геофизических регистрограмм

Глава IV. Изучение взаимосвязи между последовательными значениями регистрограммы 110

Предварительные замечания и определения . . . . . —
Геологические аспекты задачи 112
Методы решения и примеры их использования 113
Обсуждение результатов 132

Глава V. Выявление и анализ периодичностей в регистрограммах 132

Постановка и содержание задачи . . . . . —
Основы метода выделения скрытых периодичностей . . . 134
Изучение разрезов карбонатных толщ 141
Обсуждение результатов 160

Глава VI. Расчленение и корреляция разрезов скважин на основе анализа конфигурации регистрограмм 161

Постановка и содержание задачи . . . . . —
Метод непараметрического описания структуры регистрограмм (спектральный анализ) . . . . . 163.
Расчленение осадочной толщи и составление корреляционной схемы 170
Некоторые интегральные параметры при описании и изучении конфигурации регистрограмм 176
Обсуждение результатов 185

Заключение 186

Список литературы 188

ПРЕДИСЛОВИЕ

При решении многих задач, связанных с изучением разрезов осадочных толщ, широко используются данные геофизических исследований скважин. Существующие методы регистрации физических свойств горных пород в скважинах доставляют ценную, детальную и в то же время обильную и многообразную информацию, характеризующую различные черты разрезов осадочных толщ.

В подавляющем большинстве случаев при обработке и интерпретации геолого-геофизических данных используются либо теоретические методы математической физики, либо эмпирические методы в их традиционной постановке. В основе этих методов, как известно, лежат детерминистические принципы, т. е. предполагается, что существует строго однозначное соответствие между изменением геологических характеристик горных пород и отражающих их геофизических полей (показателей). Поэтому основная часть исследований в промысловой геофизике была направлена как на поиск зависимостей между геофизическими полями (показателями) и геологическими характеристиками (свойствами) горных пород, так и на устранение различных неслучайных эффектов, осложняющих оценку этих свойств. В подобных исследованиях реальную горную породу, как правило, заменяют более простой (идеализированной) средой со значительно меньшим, чем в реальной ситуации, числом переменных, способных вызвать изменчивость геофизических показателей. Только в этом случае теоретические методы математической физики дают ощутимые результаты при решении прямых и обратных задач промысловой геофизики [1, 33, 38].

Для иллюстрации можно сослаться на методы интерпретации геофизических данных, характеризующих коллекторы с межзерновым типом пористости. Так, при исследовании воздействия межзерновой пористости на электрические показатели структура пористой среды принимается известной, вещественный состав скелета породы полагается электрически непроводящим, а заполнитель пор считается неизменным. Переменной характеристикой в этом случае является лишь собственно значение межзерновой пористости. В такой ситуации на основе методов математической физики сравнительно легко получены аналитические выражения,

связывающие межзерновую пористость пород с удельным электрическим сопротивлением, собственными потенциалами и т. п.

Подобный подход используется и при анализе данных, характеризующих глинистые терригенные, а также карбонатные коллекторы. Последние, как известно, обладают пористостью смешанного типа (межзерновой, трещинной, каверновой); в этих условиях раздельная оценка пористости каждого типа крайне осложнена. Здесь теоретические методы математической физики оказываются ценными лишь в том отношении, что они позволяют аналитически, без дорогостоящих экспериментов, вскрыть функциональные связи между геологическими параметрами коллекторов и геофизическими данными [27, 50].

Однако стремление заменить реальные горные породы адекватными идеализированными агрегатами не всегда приводит к успеху, ибо с усложнением структуры породы и ее состава приходится учитывать все возрастающее число переменных, которые воздействуют на тот или иной исследуемый геофизический показатель. Это заставляет отказываться от использования методов математической физики и искать другие, в частности эмпирические.

С помощью эмпирических методов (в их традиционной постановке) выявляются и оцениваются сила и характер в основном двумерных зависимостей (очень редко трехмерных) между геологическими свойствами и разного рода геофизическими показателями. В этом направлении проведено большое число исследований, так как каждый интерпретатор старается представить результат в виде эмпирического графика, связывающего те или иные исследуемые им свойства горных пород с изменением геофизических характеристик. Однако результаты эмпирических методов, особенно нацеленных на восстановление одних характеристик (как правило, геологических) по другим (как правило, геофизическим), хорошо согласуются с действительностью только тогда, когда число переменных, способных вызвать изменчивость геофизических показателей, невелико, а искомые зависимости носят функциональный либо близкий к нему характер. Построением эмпирических графиков анализ обычно исчерпывается. Выразить обнаруженные зависимости аналитически исследователи, если и стремятся, то не всегда, ибо это связано с трудоемкими математическими операциями.

С появлением ЭВМ при интерпретации геофизических данных стали широко использовать комплексирование методов математической физики с экспериментальными, что, как правило, сводится к выработке определенного интерпретационного алгоритма с последующей его реализацией на ЭВМ. В этом случае ЭВМ как бы подменяет опытного интерпретатора и освобождает его от трудоемких «ручных» операций [39, 52, 77]. Безусловно, что при этом можно учесть влияние большого числа свойств и особенностей горных пород, ввести ряд необходимых поправок и т. п.

Оценивая изложенные традиционные методы, прежде всего необходимо подчеркнуть хорошо известный факт, что геофизические показатели неспособны отражать изменение отдельных геологических характеристик и особенностей горных пород (например, изменение пористости, вариацию вещественного состава пород или физико-химических свойств флюида, заполняющего поровое пространство). Эти показатели отражают изменение совокупности характеристик, которые к тому же сложным образом связаны между собой. Более того, в области, ограниченной даже малым радиусом, геологические, а следовательно, и геофизические показатели подвержены различным флуктуациям — случайной рассеянности (изменчивости). В одних случаях можно установить, что эти флуктуации имеют геологическую природу, в других они бывают настолько сложными, что кажутся случайными.

Большое число геологических характеристик и их флуктуации, получая отображение в геофизических данных, могут настолько осложнить интерпретацию, что при использовании традиционных методов результаты истолкования становятся неоднозначными. Так, методы математической физики в принципе не предназначены учитывать воздействие случайных вариаций, и, следовательно, полученные с их помощью аналитические геолого-геофизические зависимости справедливы для идеализированных, но не для реальных сред. То же самое можно сказать об эмпирических способах интерпретации и о методах комплексирования. Поскольку последние основаны на строго детерминистических принципах, не реагирующих на флуктуационную природу анализируемых данных, то использование комплексных алгоритмов при интерпретации исходных геофизических показателей не всегда целесообразно.

Следует отметить, что привлечение традиционных методов интерпретации, сформировавшихся на основе строгих физических и математических принципов, часто оправданно. Так, при истолковании геофизических данных, характеризующих «классические» песчано-глинистые разрезы (пласты песчаников в этих разрезах сравнительно мощные, неглинистые или слабоглинистые), общеизвестные методы дают достаточно хорошие результаты. Однако кризис традиционных приемов интерпретации уже наметился. Например, эти приемы малоэффективны при изучении разрезов карбонатных коллекторов, а также разрезов, сложенных плотными неоднородными (по глинистости) песчаниками со сложной структурой порового пространства.

Все изложенное заставляет при интерпретации геофизических измерений искать другие методы обработки данных, которые лучше соответствовали бы специфике реальных природных объектов. Работы последних лет показали, что целесообразно использовать методы, которые позволяют детально исследовать изменчивость или взаимосвязь геолого-геофизических показателей, причем под изменчивостью показателей понимается их способность

к систематическому (закономерному) и случайному (стохастическому) отклонениям. Если систематические отклонения, как правило, отражают эволюционное (поступательное) развитие и изменение свойств и состава геологической среды во времени или пространстве (по разрезу или площади), то случайные колебания связаны с «диффузностью» объекта, т. е. со случайным искажением поступательного развития. Поэтому в самом общем виде исходная изменчивость регистрируемых показателей может быть представлена следующим образом:

Изменчивость общая = изменчивость закономерная + + изменчивость случайная.

Задача интерпретации сводится к разложению общей изменчивости на соответствующие составные элементы. При этом очень часто одну часть исходной информации определяют как полезную, а другую — как мешающую. Однако было бы не совсем верно лишь случайную изменчивость связывать с мешающей информацией. Такое разделение в известной мере условно, и все зависит от той конкретной задачи, которая стоит перед исследователем. Правильнее под полезной понимать ту часть исходной информации, которая содержит сведения только об интересующих исследователя свойствах и составе пород. Расщепление интегральных сведений целесообразно и потому, что составные элементы могут быть проинтерпретированы намного полнее и эффективнее. Более того, устранив случайные или неинтерпретируемые («мешающие») эффекты, к оставшейся части информации можно применить методы, основанные на детерминистических принципах, и при этом получить более достоверные результаты. Отсюда вытекает немаловажное следствие: указанный подход не только не исключает использование общепринятых методов интерпретации, но и может существенно их дополнить.

При решении задачи, связанной с расщеплением исходной информации, необходимо ответить на следующие вопросы. Каким образом можно определить, что собственно следует понимать под закономерной изменчивостью, которая составляет часть общей изменчивости геолого-геофизических данных? Как найти отличительные черты и критерии, позволяющие одну часть информации отнести к закономерной, а другую к случайной, если оба ее вида объединены и содержатся в исходных результатах наблюдений?

Неоднозначность геофизической информации является лишь одной стороной рассматриваемой проблемы. Другая заключается в том, что интерпретатор, как правило, имеет дело с анализом и обработкой большого объема разнохарактерных наблюдений: геологических (определения по керну, шлифам и пр.) и геофизических (измерения непосредственно в скважине или на образцах). Поэтому совершенно необходимы эффективные способы обработки больших массивов информации. Целесообразно сначала предста-

вить эти данные в компактном, легко обозримом виде, или, иначе говоря, провести свертку исходной информации. Результаты свертки позволяют оценить силу и характер взаимозависимости между разными параметрами объекта, определить, какую часть исходной информации следует считать полезной, а какую мешающей. Это помогает отыскать соответствующие методы расщепления информации, которые существенно облегчают распознавание и восстановление по косвенным (геофизическим) данным искомых геологических характеристик.

Изложенным требованиям свертки и расщепления информации лучше всего отвечают модели методов математической теории оптимального эксперимента (МТОЭ), покоящихся на идеях теории вероятностей, теории информации, математической статистики и кибернетики.

В геологии модели интуитивно применяются во многих случаях. Уже сама словесная зарисовка наблюдаемого геологического явления, оцененного геолого-геофизическими данными, может быть воспроизведена в математических терминах и, следовательно, представляет собой математическую модель. Действительно, словесное описание природного явления, как правило, всегда схематизировано, доведено до крайне простых представлений, хотя само явление, безусловно, может быть несравненно более сложным. Так, если утверждается, что с увеличением глубины имеется тенденция уменьшения пористости горных пород, то исследователь в зависимости от полноты данных, теоретического освещения вопроса, опыта и интуиции в своем воображении создает тот или иной идеализированный закон связи пористости с глубиной. При этом он может представить, что пористость с глубиной уменьшается по линейному, показательному или более сложному закону. В принципе такое представление и есть модель (закон) изменения пористости, реализованная уже не только мысленно, но и математически, ибо каждому закону соответствует вполне определенное математическое выражение.

Приняв такую модель в качестве возможной рабочей схемы, исследователь, бесспорно, отдает себе отчет, что на самом деле изменение пористости с глубиной намного сложнее, но упрощенными, схематизированными понятиями ему оперировать гораздо легче. В общем случае такой подход позволяет быстрее уловить как общие тенденции, так и частные отклонения в фактических кривых, которые отражают характер поведения пористости по разрезу осадочной толщи. Следовательно, появляется возможность сравнительно легко обобщить обширный материал, уловить в нем специфические детали, раскрывающие особенности строения исследуемого геологического объекта.

Можно поступить и иначе, а именно: в начале исследования на основе различных априорных сведений попытаться в самых общих чертах представить, как может функционировать природный механизм (динамическая система), чтобы изменение пористости

было примерно таким, какое наблюдается в реальной ситуации. Вероятно, для не очень сложных случаев это функционирование может быть описано математически в рамках тех или иных стохастических дифференциальных уравнений. Очевидно, что чем больше при этом имеется априорных сведений, тем более полную модель можно создать и, следовательно, тем эффективнее можно объяснить изменчивость исходных данных, характеризующих исследуемый геологический объект.

Образное определение математической модели дано физиком-теоретиком Я. И. Френкелем: «Математическая модель напоминает своеобразную зарисовку явления, воспроизведенную художником-карикатуристом. Так, художник должен воспроизвести оригинал не во всех деталях, подобно фотографическому аппарату, но упростить и схематизировать его таким образом, чтобы выявить и подчеркнуть наиболее характерные черты. Фотографической точности можно и следует требовать лишь от теоретического описания простейших систем. Хорошая теория сложных систем должна представлять лишь хорошую «карикатуру» на эти системы, утрирующую те свойства их, которые являются наиболее типичными, и умышленно игнорирующую все остальные, несущественные, свойства. Хорошая карикатура на какого-либо человека не может существенно улучшиться от более аккуратного и точного изображения нехарактерных деталей его лица или фигуры».

Следовательно, основное достоинство математической модели состоит в том, что она позволяет в форме определенных математических выражений описать характерные, феноменологические, черты изменчивости эмпирических данных. Кроме того, если проявляется определенная согласованность в поведении эмпирических данных разного рода, характеризующих один и тот же геологический объект, то соответствующие математические модели могут описать эту согласованность, а точнее, взаимозависимость между регистрируемыми геолого-геофизическими параметрами. Очевидно, что при этом непременно будет существовать расхождение между математической моделью и эмпирическими данными, которое можно учесть заранее введенными мерами и критериями.

Надо отметить, что при построении модели отдельные черты изменчивости или взаимозависимости геолого-геофизических данных, характеризующих исследуемый объект, могут оказаться неоправданно подчеркнутыми, другие затушеванными, слишком упрощенными или даже искаженными. Как правило, это происходит из-за недостатка априорных сведений. Необходимость априорных сведений для осуществления грамотной обработки и эффективной интерпретации результатов экспериментальных наблюдений подтверждают самые строгие теоретические доказательства. Так, на основе теоретико-информационного подхода Л. А. Халфин [85] показал, что интерпретация геофизических сигналов (это в полной мере относится и к разного рода диаграммам и кривым, регистри-

руемым по скважине) может считаться комплексной геолого-геофизической, если кроме статистических свойств шумовой компоненты (случайной изменчивости) известны априорные геологические данные о закономерной составляющей. Такой же результат получен и при статистической оценке сигнала, смешанного с шумом [2, 37, 43, 54]. Образно говоря, прежде чем что-то выделять, надо иметь четкое представление, что выделять.

Очевидно, что необходимые априорные сведения можно почерпнуть лишь из суммы накопленных геологических знаний. Однако в геологии эти знания бывают недостаточно конкретными, что сильно затрудняет создание модели, согласующейся с исследуемым явлением. Более того, недостаток априорных сведений может пагубно отразиться на результатах интерпретации (что часто наблюдается в практике), ибо принятая модель может оказаться неадекватной экспериментальным данным. Правда, для таких ситуаций в МТОЭ разработаны специальные приемы [60, 84, 89]. Так, в подобных случаях можно предложить и апробировать несколько моделей, отличных друг от друга, и выбрать ту, которая по заранее заданным критериям лучше всего согласуется с эмпирическими наблюдениями. Очень часто бывает и так, что все созданные модели имеют право на существование и не конкурируют друг с другом, ибо каждая модель подчеркнуто и контрастно описывает специфические черты явления, важные для одного рода исследований и второстепенные для другого.

Методы МТОЭ под тем или иным названием получили широкое распространение. Например, в разведочной геофизике они являются первостепенными и дают интересные практические результаты [2, 24, 73]. Эти методы внедряются также при обработке и анализе собственно геологических данных с целью решения ряда теоретических вопросов [18—20, 47, 48, 57]. Что касается использования методов МТОЭ при интерпретации скважинных геолого-геофизических данных, то здесь в какой-то мере применяются лишь регрессионный и корреляционный анализы в их традиционной постановке [58, 68]. Критические замечания теоретического характера по этому поводу приведены в работах А. Е. Кулинковича [51, 52]. Имеются примеры использования методов распознавания образов [6, 29, 30, 58]. Однако такие методы, как факторный анализ, регрессионный анализ по методу главных компонент, теория случайых функций, ортогональное разложение переменных и ряд других, в промыслово-геофизической практике применяются исключительно редко [35, 36, 45, 46].

Данная книга в определенной мере восполняет этот пробел. В ней рассмотрена интерпретация геофизических и геологических наблюдений в скважинах на основе их обработки методами МТОЭ. Установлено, какие из методов наиболее приемлемы; определена область, где те или иные конкретные методы наиболее рациональны; показаны и оценены возможности и эффективность этих методов; проиллюстрирована их ценность в практических при-

ложениях; указаны пути численной реализации методов МТОЭ на основе широкого использования ЭВМ.

Можно утверждать, и в книге это обстоятельство подчеркивается особо, что интерпретация результатов геофизических исследований скважин осуществляется для решения задач двух принципиально разных классов. Один класс — это получение геологической информации об интересующем исследователя объекте, т. е. о свойствах и составе осадочных пород, слагающих изучаемые разрезы. В этом случае по геофизическим данным надо определить литологический состав пород, оценить емкость коллектора и установить его тип, выяснить вид флюида, насыщающего поровое пространство, расчленить и сопоставить разрезы и т. д. Этот круг вопросов в геофизике поставлен уже давно и решается на протяжении многих лет, но, как правило, с иных, чем принятые в данной книге, позиций.

Другой класс задач в принципе не является для геофизических исследований традиционным. Его задача — раскрыть на основе геофизической информации природу явлений, обусловливающих облик объекта, например, объяснить строение разреза, характер чередования слоев, упорядоченное изменение свойств пород и т. д. Но, вероятно, этим не следует ограничиваться, ибо геофизическая информация в ряде случаев оказывается настолько детальной и емкой, что разрешает поставить и другой круг задач, например, оценить характерные черты динамики геологического процесса, ответственного за распределение и формирование тех или иных свойств горных пород; вскрыть факторы (причины), определившие изменчивость исследуемых свойств по разрезу. Именно этим вопросам в книге и уделено особое внимание, так как при традиционных методах обработки и анализа исходных данных они. как правило, выпадают из поля зрения.

Обычно с помощью общепринятых методов решают чисто прагматические задачи. Например, стараются как можно точнее оценить пористость или установить, какими породами представлены отложения в данной точке разреза. Но при этом не задаются вопросом, в результате каких процессов реализовалось наблюдаемое изменение пористости по разрезу или определенное чередование слоев разного литологического состава, как связаны свойства и состав последовательных слоев и вследствие каких причин эта связь возникла. Между тем известно, что пористость осадочных пород и особенно характер ее распределения по разрезу могут рассматриваться не только как собственно коллекторские параметры осадочных толщ, а в гораздо более общем плане. Так, распределение пористости межзернового типа по вертикали несет информацию о характерных особенностях седиментационных и постседиментационных процессов [19, 35, 45]; картина трещиноватости объекта отражает интенсивность тектонических напряжений [44, 76]; каверновая пористость карбонатных пород позволяет судить об интенсивности вторичных процессов [23, 76].

Естественно предположить, что диаграммы результатов геофизических наблюдений содержат в себе геологическую информацию и генетического характера, но в сильно завуалированной форме. Выявив эту информацию, можно обоснованнее и увереннее интерполировать и экстраполировать свойства осадочных толщ, детальнее анализировать исследуемый геологический объект в целом для его эффективной реконструкции.

В соответствии с решаемыми задачами книга разделена на две части. В первой рассматриваются методы изучения осадочных пород по данным геофизических исследований скважин. Здесь ставятся следующие цели: оценить коллекторские и литолого-петрографические характеристики горных пород, выявить факторы, определяющие их изменчивость, отличить водоносные пласты от продуктивных. Вторая часть посвящена изучению строения разрезов осадочных толщ по геофизическим данным. Основное внимание здесь сосредоточено на анализе структуры геолого-геофизических регистрограмм. Такой анализ позволяет в целом ряде случаев объяснить наблюдаемую конфигурацию регистрограммы действием определенного природного механизма, т. е. провести генетическую интерпретацию.

Использование специфических методов, естественно, потребовало изложить часть материала на особом языке, который в какой-то степени непривычен для геологов и геофизиков, обрабатывающих скважинные данные. Этот язык, порожденный теорией вероятностей, кибернетикой и многомерной математической статистикой, сначала применялся в теории радиосвязи, теории автоматического регулирования, статистической физике. Лишь недавно он вошел в разведочную геофизику (особенно в сейсморазведку). В область же промысловой геологии и геофизики этот язык только начинает проникать. Неподготовленному читателю рекомендуем ознакомиться с рядом руководств по этому вопросу [47, 48, 57, 91], выпущенных специально для геологов.

Обработка и интерпретация исходных геолого-геофизических данных осуществляется на основе широкого применения ЭВМ (БЭСМ-2, -4 и -6); составленные авторами программы предназначены для геолого-геофизических наблюдений в скважинах, но могут быть с успехом использованы и для других материалов. Программы реализуют целый ряд методов МТОЭ: регрессионный, корреляционный, факторный, авторегрессионный, автокорреляционный и спектральный анализы.

Авторы считают своим приятным долгом выразить признательность проф. Е. М. Смехову за поддержку при написании книги. Авторы благодарны также И. П. Макаренко, оказавшему существенную помощь при составлении программ для ЭВМ, и А. П. Хусу, взявшей на себя труд проследить математические выкладки. Особо ценные замечания, способствовавшие более четкому и последовательному изложению материала, были сделаны И. Г. Клушиным, которому авторы рады выразить свою признательность.

Часть первая

ИЗУЧЕНИЕ ОСАДОЧНЫХ ПОРОД

Глава I

ОЦЕНКА КОЛЛЕКТОРСКИХ И ЛИТОЛОГО-ПЕТРОГРАФИЧЕСКИХ ХАРАКТЕРИСТИК

ПОСТАНОВКА И СОДЕРЖАНИЕ ЗАДАЧИ

Оценка коллекторских и литолого-петрографических характеристик горных пород является одной из основных задач, которые возникают при интерпретации геофизических наблюдений в разрезах скважин. Эта задача широко известна, однако ее решение, за редким исключением, нельзя назвать удовлетворительным. Дело в том, что при интерпретации, как правило, широко используются заранее установленные соотношения между интересующими исследователя геологическими характеристиками и измеренными физическими свойствами пород. Очевидно, что, зная такое соответствие в виде либо вербальных (словесных) описаний, либо строгих математических соотношений (что предпочтительнее), гораздо легче восстановить геологические (в частности, коллекторские и литолого-петрографические) характеристики.

Однако в условиях, когда на геофизические показатели влияет большое число геологических характеристик, которые сложным образом взаимосвязаны и, кроме того, обладают случайной изменчивостью, искомые соотношения получить достаточно сложно, так же как и оценить их правильность. Если этому вопросу и уделяется внимание, то не всегда достаточное, обычно он решается без привлечения строгих мер и критериев, отвечающих сложности задачи. Поэтому прежде всего необходимо строго сформулировать задачу и уяснить себе ее содержание.

Чтобы по данным геофизических исследований скважин правильно восстановить коллекторские и литолого-петрографические характеристики горных пород, необходимо знать, какие из них и в какой степени влияют на геофизические измерения. Разумеется, если влияние той или иной характеристики не сказывается, то информацию о ней из геофизических данных извлечь невозможно. Поэтому требуется выяснить, как реагируют показания того или иного геофизического метода на изменение интересующих нас характеристик пород. Но в условиях, когда природные факторы действуют одновременно (такая ситуация при решении подобных задач типична), трудно определить, по каким

именно причинам изменяется геофизическая характеристика. Более того, иногда вообще бывает неясно, изменяется эта характеристика под влиянием интересующих исследователя геологических факторов или ее вариация отражает воздействие многих эффектов случайного характера.

Таким образом, специфика поставленной задачи (установить влияние коллекторских и литолого-петрографических характеристик горных пород на изменчивость геофизических данных) заключается в том, что изучается большое число влияющих факторов и их воздействие осложнено вариациями случайного характера. Заметим, что сами факторы чаще всего имеют численное выражение, но, безусловно, не исключены случаи, когда они могут быть описаны только качественно, например, если рассматривается влияние определенного типа пород (песчаник, известняк, доломит и т. п.) или условий (атмосферных, пластовых), при которых осуществлялась регистрация геофизических данных.

МЕТОДЫ РЕШЕНИЯ И ПРИМЕРЫ ИХ ИСПОЛЬЗОВАНИЯ

Специфике стоящей перед нами задачи отвечает дисперсионный анализ, так как он представляет собой статистический метод обработки и истолкования результатов наблюдений, зависящих от различных одновременно действующих факторов, осложненных эффектами случайного характера. Этот метод широко известен, чаще всего он используется в экспериментальных областях науки и техники (биология, физика, астрономия, металловедение и др.). При обработке геолого-геофизической информации, получаемой по скважинам, этот метод, однако, применяется редко и не во всей полноте, хотя его эффективность и целесообразность, как будет видно из приводимых ниже примеров, не вызывает сомнений.

Метод основан на изучении дисперсии (рассеяния, вариации) исходных характеристик. Все приемы дисперсионного анализа нацелены на то, «чтобы извлечь какую-то существенную информацию из того простого и на первый взгляд, казалось бы, неприятного факта, что результаты наблюдений оказываются рассеянными» [60, с. 69]. Методология математической статистики сумела показать, что изучение дисперсии путем ее разложения на составные элементы — это путь к исследованию сложных многофакторных систем, какими являются геологические объекты, и в частности горные породы.

В рассматриваемой задаче целесообразно различать три основных аспекта [92]: первый — установить факт изменения средних значений одних характеристик (скажем, геофизических) с изменением средних значений других характеристик (например, геологических); второй — найти уравнение, связывающее параметры какой-либо геологической характеристики (коллекторской, литолого-петрографической) горных пород с комплексом

результатов геофизических показаний; третий — решить вопрос о сохранении вида полученного уравнения при изменении одного или нескольких факторов. В соответствии с этим рассмотрим три типа анализа, укладывающихся в одну математическую схему: собственно дисперсионный, регрессионный и ковариационный, отвечающие трем указанным аспектам задачи. В их основе лежит одно и то же математическое представление [92]:

$$y_i = x_{1i}\beta_1 + x_{2i}\beta_2 + \ldots + x_{pi}\beta_p + e_i, \qquad (I.1)$$

где n наблюдаемых случайных величин y_1, y_2, ..., y_n являются линейными комбинациями с p неизвестными постоянными β_1, β_2, ..., β_p; e_1, e_2, ..., e_n — флуктуации наблюдений; $\{x_{ji}\}$ — в зависимости от модификации анализа либо известные постоянные коэффициенты, либо известные переменные.

Введем векторные и матричные обозначения. Представим последовательность исходных наблюдений (например, значения какой-либо геологической характеристики) как вектор $\mathbf{Y}$, параметры $\{\beta_j\}$ как вектор β, флуктуации как вектор $\mathbf{e}$:

$$\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \ldots \\ y_n \end{pmatrix}; \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \ldots \\ \beta_p \end{pmatrix}; \quad \mathbf{e} = \begin{pmatrix} e_1 \\ e_2 \\ \ldots \\ e_n \end{pmatrix}.$$

Введем матрицу

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1n} \\ x_{21} & x_{22} & \ldots & x_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ x_{p1} & x_{p2} & \ldots & x_{pn} \end{pmatrix},$$

элементы каждой строки которой могут быть, например, показаниями отдельного геофизического метода, а совокупность строк — показаниями комплекса геофизических методов.

Тогда множество уравнений (I.1) можно записать в виде матричного выражения

$$\mathbf{Y} = \mathbf{X}'\beta + \mathbf{e};$$

если $\mathrm{M}(\mathbf{e}) = 0$, то

$$\mathrm{M}(\mathbf{Y}) = \eta = \mathbf{X}'\beta,$$

где $\mathbf{X}'$ — матрица, транспонированная с $\mathbf{X}$; M — математическое ожидание.

Если элементы матрицы $\mathbf{X}$ указывают на присутствие или, наоборот, отсутствие влияния анализируемых факторов и принимают соответственно значения 1 или 0, то мы имеем дело с собственно дисперсионным анализом. Если элементы векторов-строк матрицы $\mathbf{X}$ пробегают непрерывный ряд значений тех или иных показаний (например, дискретные данные каротажа), то мы получаем модель регрессионного анализа. В этом случае величина

$\{x_{ji}\}$ называется переменной-предсказателем или регрессором, а $\{y_i\}$ — предсказуемой или регрессируемой переменной. В том случае, когда величины $\{x_{ji}\}$ есть переменные двух типов, имеющие смысл как влияющих факторов (принимают значения 1 или 0), так и переменных-предсказателей, мы имеем дело с ковариационным анализом.

Неизвестные параметры $\{\beta_j\}$ являются отражением некоторых сторон наблюдаемого явления. Конечная цель дисперсионного анализа — получить выводы относительно величин $\{e_i\}$ и некоторых $\{\beta_j\}$. При этом общетеоретические концепции дисперсионного анализа требуют сначала выдвинуть некоторую гипотезу H относительно искомых параметров $\{\beta_j\}$, которая, как правило, записывается в виде равенства функций от оцениваемых параметров модели. Затем разными способами оцениваются дисперсии наблюдений, соответствующие выдвинутым предположениям, и проверяется равенство дисперсий (обычно по F-критерию).

В основе дисперсионного анализа лежат основные предположения Ω [92], а именно:

$$\mathrm{M}(\mathbf{Y}) = \boldsymbol{\eta} = \mathbf{X}'\boldsymbol{\beta} \text{ или } \mathbf{Y} = \mathbf{X}'\boldsymbol{\beta} + \mathbf{e}.$$

Как и ранее, математическое ожидание флуктуаций $\mathrm{M}(\mathbf{e}) = 0$, флуктуации некоррелированы (независимы) и имеют одинаковые дисперсии, т. е. $\mathrm{M}(\mathbf{ee}') = \sigma^2\mathbf{I}$, где $\mathbf{I}$ — единичная матрица.

В этом случае оценкой компонентов вектора $\boldsymbol{\beta}$ служит величина $\mathbf{b} = (\mathbf{XX}')^{-1}\mathbf{XY}$, которую находят из условия минимизации суммы квадратов флуктуаций $\sum_{i=1}^{n} e_i^2$:

$$SS_e = \sum_{i=1}^{n} e_i^2 = (\mathbf{Y} - \mathbf{X}'\mathbf{b})'(\mathbf{Y} - \mathbf{X}'\mathbf{b}).$$

Для проверки гипотезы H к предположениям Ω необходимо добавить налагаемые ею ограничения, так что при выполнении гипотезы мы имеем предположения, которые в символической записи выражаются следующим образом: $\omega = H \cap \Omega$. Коэффициенты $\mathbf{b}_\omega$ в предположениях ω оценивают, минимизируя сумму квадратов отклонений наблюденных значений от их математического ожидания, полученного уже при условии выполнения гипотезы H:

$$SS_\omega = (\mathbf{Y} - \mathbf{X}'\mathbf{b}_\omega)'(\mathbf{Y} - \mathbf{X}'\mathbf{b}_\omega).$$

F-критерий в этих условиях имеет вид

$$F = (\nu_e/\nu_H)(SS_H/SS_e);\ SS_H = SS_e - SS_\omega,$$

где ν_e и ν_H — числа степеней свободы.

Теперь F-критерию можно дать наглядное объяснение. С точностью до постоянной он определяется выражением $1 - SS_\omega/SS_e$.

Величину SS_e можно рассматривать как меру точности подбора оценок $\beta_1, \beta_2, \ldots, \beta_p$ по исходным наблюдениям $y_1, y_2, \ldots, y_n$; чем меньше значение SS_e, тем лучше проведен выбор оценок. Величина же SS_ω характеризует наилучший выбор, который может быть сделан в предположениях $\omega = H \cap \Omega$; она рассматривается как мера согласованности гипотезы H с экспериментальными данными. Таким образом, величина SS_e показывает, насколько мала может быть мера SS_ω, если оставить только основные предположения Ω. Тогда отношение SS_ω/SS_e показывает, насколько хуже выбор при предположениях ω, чем при Ω. Гипотеза H отвергается, если отношение SS_ω/SS_e велико. Практически это сводится к тому, что полученный по наблюдениям F-критерий сравнивается с табличными значениями $F_{\alpha,\, \nu_H,\, \nu_e}$ при уровне значимости α и числе степеней свободы ν_H и ν_e. Гипотеза H отвергается, если $F > F_{\alpha,\, \nu_H,\, \nu_e}$.

Расчет F-критерия можно рассматривать как предварительный способ решения вопроса о целесообразности продолжать обработку наблюдений другими, более сложными, методами. Если гипотеза H по F-критерию отвергается, то исследуются соответствующие функции от параметров β_j, чтобы решить вопрос, из-за каких именно параметров гипотеза не принята.

Сравнением параметров $\beta_1, \beta_2, \ldots, \beta_p$ является линейная функция $\psi = \mathbf{C}'\beta$ с известными постоянными коэффициентами, удовлетворяющими условию $\sum_{j=1}^{p} C_j = 0$ [92]. При этом сравнением можно называть разность любых значений $\beta_j - \beta_i$ или средних некоторых подмножеств. Гипотеза H эквивалентна утверждению, что любые сравнения $\psi = (\psi_1, \psi_2, \ldots, \psi_q)$ равны нулю. Она будет отвергнута лишь в том случае, когда не все истинные сравнения равны нулю. Чтобы судить, значимо ли отличаются оценки сравнений от нуля, надо построить доверительный эллипсоид. Если эллипсоид не накроет точку $\psi = 0$, то гипотезу H мы вправе отбросить.

Случайная величина ψ распределена нормально с вектором средних $\hat{\psi} = \mathbf{C}'\beta$ [где $\mathbf{C} = (c_{ji})$] и ковариационным матрицей оценок $\{\psi_i\}$, определяемой как $\mathbf{\Gamma}_{\hat{\psi}} = \sigma^2\mathbf{A}\mathbf{A}' = \sigma^2\mathbf{B}$. Элементы матрицы $\mathbf{A}$ зависят от оценок $\mathbf{b} = (b_1, b_2, \ldots, b_p)$.

Доверительный эллипсоид дается выражением $(\psi - \hat{\psi})'\mathbf{B}^{-1}(\psi - \hat{\psi})$. Гипотеза H отвергается тогда и только тогда, когда

$$(\psi - \hat{\psi})'\mathbf{B}^{-1}(\psi - \hat{\psi}) > q\sigma^2 F_{\alpha,\, q,\, n-p}.$$

В конечном итоге, получив ту или иную оценку дисперсий и параметров β_i, после соответствующих проверок можно установить, действительно ли предполагаемый фактор (или факторы) влияет на изменчивость определяемой характеристики или это влияние отсутствует.

ГОСУДАРСТВЕННАЯ БИБЛИОТЕКА ... 2

Интерпретации геофизических или геологических данных в соответствии с первым из вышеуказанных аспектов (установить факт изменения средних значений одних характеристик с изменением средних значений других характеристик) отвечает модель собственно дисперсионного анализа [92]. Технику обработки исходных данных и основные возможности метода проиллюстрируем на примерах так называемого двухфакторного дисперсионного анализа. В этом случае оценивается действие двух факторов A и B на какую-либо определяемую характеристику C.

Если фактор A имеет I уровней, а фактор B характеризуется J уровнями, то результаты наблюдений по этим двум факторам могут быть расположены в виде таблицы с двумя входами; I строк таблицы соответствуют уровням фактора A, а J столбцов — уровням фактора B. В (ij)-ю ячейку, расположенную на пересечении i-й строки и j-го столбца, записываются значения параметра C, полученные при одновременном исследовании факторов A и B соответственно на i-м и j-м уровнях. Рассмотрим случай, когда в каждой ячейке имеется одно наблюдение. Тщательный разбор этого случая поможет понять ситуацию, при которой в каждой ячейке находится k наблюдений.

Введем ряд основных понятий и обозначений, совокупность которых определяет математическую модель двухфакторного анализа. Предположим сначала, что в каждой ячейке находятся истинные средние значения η_{ij}. Тогда среднее i-го уровня фактора A (среднее i-й строки) получим как $\eta_{i*} = \left(\sum_j \eta_{ij}\right) \Big| J$, среднее j-го уровня фактора B (среднее j-го столбца)—как $\eta_{*j} = \left(\sum_i \eta_{ij}\right) \Big| I$, генеральное среднее — как $\eta_{**} = \left(\sum_i \sum_j \eta_{ij}\right) \Big| (IJ)$.

Главный эффект i-го уровня фактора A определяется как превышение среднего i-го уровня над генеральным средним:

$$\alpha_i = \eta_{i*} - \eta_{**}.$$

Аналогично главный эффект j-го уровня фактора B

$$\beta_j = \eta_{*j} - \eta_{**}.$$

Главные эффекты α_i и β_j называются также эффектами соответственно i-й строки и j-го столбца. Взаимодействием i-го уровня фактора A и j-го уровня фактора B называется величина

$$\gamma_{ij} = \eta_{ij} - \eta_{i*} - \eta_{*j} + \eta_{**}.$$

Таким образом, модель двухфакторного анализа при истинных средних ячейки принимает вид

$$\eta_{ij} = \eta_{**} + \alpha_i + \beta_j + \gamma_{ij}.$$

Проверяемыми гипотезами являются H_A: все $\alpha_i = 0$; H_B: все $\beta_j = 0$; H_{AB}: все $\gamma_{ij} = 0$. Гипотеза H_A утверждает, что все средние различных уровней фактора A равны, или, иными словами, все уровни фактора A имеют один и тот же эффект и поэтому данный фактор на изменение изучаемой величины влияния не оказывает. Гипотеза B имеет аналогичное содержание. Гипотеза H_{AB} соответствует оценке взаимодействия факторов.

Поскольку в реальных ситуациях в ячейках не имеем истинных средних, то каждое наблюдение

$$y_{ij} = \eta_{ij} + e_{ij},$$

где e_{ij} — флуктуации, которые распределены нормально и независимы.

Для получения корректных критериев и доверительных интервалов главных эффектов в условиях, когда проводится только одно наблюдение в ячейке, целесообразно считать, что взаимодействия факторов нет. В предположениях Ω модель имеет вид

$$y_{ij} = \mu + \alpha_i + \beta_j + e_{ij}.$$

Оценки параметров модели — истинного среднего и главных эффектов — даются следующими выражениями:

$$\hat{\mu} = \left(\sum_i \sum_j y_{ij}\right)/(IJ) = y_{**};\ \alpha_i = y_{i*} - y_{**};\ \beta_j = y_{*j} - y_{**}.$$

Проверяемыми гипотезами являются H_A: все $\alpha_i = 0$ и H_B: все $\beta_j = 0$. Соответствующие им суммы SS_H обозначим через SS_A и SS_B, а степени свободы ν_H — через ν_A и ν_B. Все возможные суммы квадратов отклонений SS, полученные в предположениях Ω и ω, и соответствующие значения ν_A, ν_B, ν_e сведены в табл. I.1. Заметим, что полная, или общая, сумма квадратов представима формулой

$$SS_{\text{п}} = SS_A + SS_B + SS_e,$$

которая и свидетельствует, что общая сумма квадратов отклонений отдельных наблюдений разлагается на составляющие, связанные с соответствующими эффектами.

Рассмотрим пример. Требуется оценить влияние пористости $K_{\text{п}}$ (фактор A) и глинистости $K_{\text{гл}}$ (фактор B) на относительное электрическое сопротивление пород — параметр пористости $P_{\text{п}}$. Все характеристики измерены на образцах керна, отобранных из двух пластов Самотлорского нефтяного месторождения [56].

Сгруппировав соответствующим образом исходные данные (табл. I.2) и выполнив необходимые вычисления, получим $SS_A = 95{,}0$; $SS_B = 3{,}5$; $SS_e = 4{,}5$. Следовательно, по F-критерию гипотеза H_A с уровнем значимости α отбрасывается, если

$$[(J-1)(I-1)/(I-1)]\,(SS_A/SS_e) > F_{\alpha,\,(I-1),\,(I-1)(J-1)}.$$

Таблица I.1

Оценки дисперсий для двухфакторного анализа с одним наблюдением в ячейке

Источник дисперсии	Сумма SS квадратов отклонений	Степень свободы ν
Главные эффекты:		
фактора A	$SS_A = J \sum_i (y_{i*} - y_{**})^2$	$I - 1$
фактора B	$SS_B = I \sum_i (y_{*j} - y_{**})^2$	$J - 1$
Флуктуации	$SS_e = \sum_{ij} \sum (y_{ij} - y_{i*} - y_{*j} + y_{**})^2$	$(I - 1)\ (J - 1)$
Полная сумма квадратов	$SS_{\text{п}} = \sum_i \sum_j (y_{ij} - y_{**})^2$	$IJ - 1$

Таблица I.2

Результаты двухфакторного дисперсионного анализа с одним наблюдением в ячейке (оценка влияния пористости $K_{\text{п}}$ и глинистости $K_{\text{гл}}$ на параметр пористости $P_{\text{п}}$)

Пористость $K_{\text{п}}$ (фактор A), %	Глинистость $K_{\text{гл}}$ (фактор B), %			y_{i*}	Главные эффекты α_i
	1—2	3—4	5—6		
21—22	16	17	19	17,33	3,83
23—24	15	15	15	15,00	1,50
25—26	10	11	10	10,33	—3,17
27—28	10	12	12	11,33	—2,17
y_{*j}	12,75	13,75	14,00		
Генеральное среднее y_{**}				13,50	
Главные эффекты β_j	—0,75	0,25	0,50		

Расчетные значения

$SS_A = 95{,}01$; $SS_B = 3{,}5$; $SS_{\text{п}} = 103$; $SS_e = 4{,}5$.

$F_A = (3-1) \cdot 95{,}01 : 4{,}49 = 42{,}5$; $F_B = (4-1) \cdot 3{,}5 : 4{,}49 = 2{,}34$.

Табличные значения

$F_{0{\cdot}1;\ 3;\ 6} = 3{,}29$; $F_{0{\cdot}1;\ 2;\ 6} = 3{,}46$.

Аналогично следует отбросить гипотезу H_B, если

$$[(J-1)(I-1)/(J-1)]\,(SS_B/SS_e) > F_{\alpha,\ (J-1),\ (I-1)\ (J-1)}.$$

Табличное значение F-критерия для гипотезы H_A (степени свободы 3 и 6, уровень значимости 10%) равно 3,29, следовательно, мы отбрасываем гипотезу о том, что пористость $K_{\text{п}}$ не влияет на параметр пористости $P_{\text{п}}$, так как расчетное значение $F = 42{,}5$. Напротив, гипотезу H_B, по которой глини-

стость $K_{гл}$ не влияет на параметр пористости $P_п$, мы должны считать состоятельной, так как экспериментальное значение $F = 2{,}34$ меньше теоретического, равного 3,46.

Таким образом, в приведенном примере относительное сопротивление зависит лишь от пористости, следовательно, нет необходимости вводить поправки за влияние глинистости.

Теперь пусть в каждой ячейке проведено по k наблюдений. Если через y_{ijk} обозначить k-е наблюдение в (i, j)-й ячейке, то модель примет вид

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk},$$

где величина γ_{ij} является оценкой взаимодействия i-го уровня фактора A и j-го уровня фактора B, тем самым учитывается и взаимодействие между факторами.

В этой модели значения $\{e_{ijk}\}$ также независимы и распределены нормально с дисперсией σ^2.

Оценками истинных средних, главных эффектов и взаимодействия являются

$$\hat{\mu} = y_{***}; \quad \hat{\alpha}_i = y_{i**} - y_{***}; \quad \hat{\beta}_j = y_{*j*} - y_{***};$$

$$\gamma_{ij} = \bar{y}_{ij*} - y_{i**} - y_{*j*} + y_{***}.$$

Разложение полной суммы квадратов отклонений наблюдений от генерального среднего представлены в табл. I.3. Из нее видно, что полная сумма квадратов

$$SS_п = SS_A + SS_B + SS_{AB} + SS_e.$$

Таблица I.3

Оценки дисперсий для двухфакторного анализа с k наблюдениями в ячейке

Источник дисперсии	Сумма SS квадратов отклонений	Степень свободы ν
Главные эффекты:		
фактора A	$SS_A = Jk \sum\limits_i (y_{i**} - y_{***})^2$	$I - 1$
фактора B	$SS_B = Ik \sum\limits_i (y_{*j*} - y_{***})^2$	$J - 1$
Взаимодействия AB	$SS_{AB} = k \sum\limits_i \sum\limits_j (y_{ij*} - y_{i**} - y_{*j*} + y_{***})^2$	$(I-1)\ (J-1)$
Флуктуации	$SS_e = \sum\limits_i \sum\limits_j \sum\limits_k (y_{ijk} - y_{ij*})^2$	$IJ\ (k-1)$
Полная сумма квадратов	$SS_п = \sum\limits_i \sum\limits_j \sum\limits_k (y_{ijk} - y_{***})^2$	$IJk - 1$

Первые три слагаемых этой суммы связаны с рассеянием наблюдений вследствие изменений исследуемых факторов и их взаимодействия. Последнее слагаемое представляет собой сумму квадратов разностей между отдельными наблюдениями в ячейке и средним соответствующей ячейки. Это слагаемое называют суммой квадратов отклонений внутри ячейки, оно характеризует остаточное случайное рассеяние. Сравнивая дисперсию по факторам с остаточной дисперсией, по их отношению судят, насколько рельефно проявляется влияние факторов.

При проверке гипотез H_A, H_B и H_{AB} соответствующий F-критерий имеет вид

$$F_A = \frac{SS_A}{SS_e} \cdot \frac{\nu_e}{\nu_A}; \quad F_B = \frac{SS_B}{SS_e} \cdot \frac{\nu_e}{\nu_B}; \quad F_{AB} = \frac{SS_{AB}}{SS_e} \cdot \frac{\nu_e}{\nu_{AB}}.$$

По F-критерию гипотеза H отвергается с заданным уровнем значимости α, если вычисленное значение F меньше табличного F_{α, ν_i, ν_e}, где ν_i — степени свободы соответственно ν_A, ν_B или ν_{AB}.

Рассмотрим еще два примера. В них также использованы данные по Самотлору [56]. Результаты сведены в табл. I.4 и I.5, сформированные по схеме двухфакторного дисперсионного анализа с k наблюдениями в ячейке.

Таблица I.4

Результаты двухфакторного дисперсионного анализа с тремя наблюдениями в ячейке (оценка влияния пористости $K_п$ и содержания суммы полевых шпатов $K_{п.ш}$ на параметр пористости $P_п$)

Пористость $K_п$ (фактор A), %	Содержание полевых шпатов $K_{п.ш}$ (фактор B), %			y_{i**}	Главные эффекты α_i
	24—34	35—45	46—56		
22—23	19; 21; 21	14; 17; 20	12; 15; 18	17,44	4,22
24—25	13; 13; 15	13; 15; 16	14; 15; 15	14,33	1,11
26—27	10; 11; 13	9; 10; 15	9; 11; 12	11,00	—2,22
28—29	7; 10; 12	9; 10; 11	10; 10; 11	10,00	—3,22
y_{*j*}	13,75	13,25	12,67		
Генеральное среднее y_{***}				13,22	
Главные эффекты β_j	0,53	0,03	—0,55		

Расчетные значения

$$SS_A = 309; \; SS_B = 7{,}0; \; SS_{AB} = 6{,}0; \; SS_п = 414; \; SS_e = 92{,}0.$$

$$F_A = 24{,}0; \; F_B = 0{,}91; \; F_{AB} = 027.$$

Табличные значения

$$F_{0{,}1;\,3;\,24} = 2{,}33; \quad F_{0{,}1;\,2;\,24} = 2{,}54; \quad F_{0{,}1;\,6;\,24} = 2{,}04.$$

Таблица I.5

Результаты двухфакторного дисперсионного анализа с тремя наблюдениями в ячейке (оценка влияния глинистости $K_{гл}$ и содержания суммы полевых шпатов $K_{п.ш}$ на пористость $K_п$)

Содержание суммы полевых шпатов $K_{п.ш}$ (фактор A), %	Глинистость $K_{гл}$ (фактор B), %			y_{i**}	Главные эффекты α_i
	1—2	3—4	5—6		
24—34	28; 27; 24	26; 28; 28	27; 28; 26	26,88	1,36
35—45	28; 26; 23	26; 27; 27	27; 27; 28	26,56	1,04
46—56	26; 23; 23	23; 22; 24	22; 22; 23	23,11	—2,10
y_{*j*}	25,33	25,67	25,55	25,52	
Генеральное среднее y_{***}					
Главные эффекты β_j	—0,19	0,15	0,03		

Расчетные значения

$$SS_A = 52{,}4;\ \bar{S}\bar{S}_B = 0{,}36;\ \bar{S}\bar{S}_{AB} = 36{,}0;\ \bar{S}\bar{S}_п = 125;\ SS_e = 36{,}0.$$

$$F_A = 13{,}0;\ \bar{F}_B = 0{,}09;\ \bar{F}_{AB} = 4{,}5.$$

Табличные значения

$$F_{0{,}1;\ 2;\ 18} = 2{,}62;\quad F_{0{,}1;\ 2;\ 18} = 2{,}62;\quad \bar{F}_{0{,}1;\ 1;\ 18} = 2{,}29.$$

В первом примере ставился вопрос о воздействии содержания суммы полевых шпатов $K_{п.\ ш}$ и пористости $K_п$ на параметр пористости $P_п$. Авторы работы [56] считают, что для образцов с близкой пористостью прослеживается слабое закономерное снижение параметра $P_п$ с ростом количества полевых шпатов.

Вычисления с учетом модели дисперсионного анализа показывают (см. таблицу I.4), что только значение пористости оказывает воздействие на параметр пористости. Следовательно, выводы авторов статьи, основанные на визуальном анализе характеристик, оказались ошибочными. Дисперсионный же анализ застраховывает исследователя от необоснованной интерпретации и неверных выводов.

Во втором примере ставился вопрос, какой из факторов в основном влияет на межзерновую пористость: содержание глинистого материала или общее количество полевых шпатов. Результаты дисперсионного анализа показывают (см. табл. I.5), что пористость существенно зависит от общего количества полевых шпатов и определенного взаимодействия количества полевых шпатов с глинистостью. Последнее обстоятельство крайне интересно. Диапазон изменения глинистости в анализируемом примере не столь широк, чтобы ожидать взаимосвязи пористости с глинистостью. Однако определенное соотношение количества полевых шпатов с глинистостью (на вопрос, какое именно, дисперсионный анализ не дает ответа) может существенно повлиять на изменение пористости. Как видим, такой факт является несколько неожиданным, и в этом направлении полезно было бы провести дополнительные исследования.

Теперь целесообразно обратить внимание на следующее обстоятельство. Во всех разобранных примерах мы приходили к выводу,

основанному на F-критерии, что какую-либо из гипотез — H_A или H_B — следует отвергнуть. После такого предварительного исследования уместно поставить вопрос: какие главные эффекты — α_i или β_j — ответственны за отбрасывание гипотезы H? С этой целью оценим сравнения.

Для двухфакторного дисперсионного анализа с k наблюдениями в ячейке сравнениями являются следующие линейные функции:

$$\hat{\psi}_A = \sum_i c_i \alpha_i \quad \text{(главные эффекты } \alpha_i\text{)};$$

$$\hat{\psi}_B = \sum_j c_j \beta_j \quad \text{(главные эффекты } \beta_j\text{)}$$

с дисперсиями

$$\sigma^2_{\hat{\psi}_A} = S^2 \sum_i c_i^2/(kI);$$

$$\sigma^2_{\hat{\psi}_B} = S^2 \sum_j c_j^2/(kJ),$$

где

$$S^2 = SS_e/[IJ\,(k-1)].$$

Доверительный интервал для ψ дается выражением

$$\hat{\psi} - \hat{S}\sigma_{\hat{\psi}} \leqslant \psi \leqslant \hat{\psi} + \hat{S}\sigma_{\hat{\psi}},$$

где постоянная $\hat{S}$ определяется формулой

$$\hat{S}^2 = (I-1)\,F_{\alpha,\ I-1,\ JI\,(k-1)} \quad \text{(главные эффекты } \alpha_i\text{)};$$

$$\hat{S}^2 = (J-1)\,F_{\beta,\ J-1,\ JI\,(k-1)} \quad \text{(главные эффекты } \beta_j\text{)}.$$

Если доверительный интервал не содержит значения $\psi = 0$, то различия главных эффектов значимы.

Оценим сравнения на примере полученных выше результатов. С помощью F-критерия было выяснено различие главных эффектов или средних значений параметра пористости $P_п$, отвечающих трем уровням содержания полевых шпатов. Теперь следует установить, каким именно уровням отвечают обнаруженные различия. Судя по средним значениям, есть смысл сравнить попарно все три главных эффекта, а также третий с группой, состоящей из первых двух.

При попарном сравнении $c_1 = 1/2;\ c_2 = -1/2;$

$$\sigma^2_{\hat{\psi}} = \frac{SS_e}{IJ\,(k-1)} \sum \frac{c_i^2}{kI} = \frac{36}{18} \cdot \frac{1/4+1/4}{9} = \frac{1}{9}$$

и, значит,

$$\sigma_{\hat{\psi}} = 1/3.$$

При групповом сравнении $c_1 = c_2 = 1/4;\ c_3 = -1/2;$

$$\sigma^2_{\hat{\psi}} = \frac{36}{18} \cdot \frac{1/16+1/16+1/4}{9} = \frac{1}{12};$$

$$\sigma_{\hat{\psi}} = 0{,}29.$$

Величина $\hat{S}^2$ при всех сравнениях одинакова:

$$\hat{S}^2 = (l-1)\,F_{0.1;\ 2;\ 18} = 2 \cdot 2{,}62 = 5{,}24;$$

$$\hat{S} = 2{,}29.$$

Соответственно при попарном сравнении

$$\hat{S}\sigma_{\hat{\psi}} = 0{,}75;$$

при групповом

$$\hat{S}\sigma_{\hat{\psi}} = 0{,}66.$$

Теперь можно вычислить соответствующие сравнения и доверительные интервалы для них.

При сравнении значений параметра пористости для образцов с содержанием полевых шпатов 24—34 и 35—45% (уровни 1 и 2) $\hat{\psi}_{1,\,2} = 0{,}16$. Сравнение уровней 1 и 3; 2 и 3 дает соответственно $\hat{\psi}_{1,\,3} = 1{,}86$ и $\hat{\psi}_{2,\,3} = 1{,}73$. При групповом сравнении $\hat{\psi}_{12,\,3} = 1{,}87$. Вычисляя доверительные интервалы, получаем

$$-0{,}59 \leqslant \hat{\psi}_{1,\,2} \leqslant 0{,}91;\ 1{,}11 \leqslant \hat{\psi}_{1,\,3} \leqslant 2{,}61;$$

$$0{,}98 \leqslant \hat{\psi}_{2,\,3} \leqslant 2{,}48;\ 1{,}21 \leqslant \hat{\psi}_{12,\,3} \leqslant 2{,}53.$$

Таким образом, только доверительный интервал для $\hat{\psi}_{1,\,2}$ содержит значение, равное нулю.

Отсюда можно заключить: в образцах с содержанием полевых шпатов 24—34 и 35—45% уловить различия в значениях параметра пористости на имеющемся материале не удается, но в образцах с содержанием полевых шпатов 46—56% параметр пористости меньше, чем при их содержании 24—34 и 35—45%. Если последние уровни объединить в один (24—45%), то для них среднее значение параметра пористости будет больше, чем в образцах с содержанием полевых шпатов 46—56%.

В заключение рассмотрим примеры использования дисперсионного анализа в случае, когда один из факторов изменяется по качественным уровням. Следует выяснить, насколько существенно воздействуют пластовые условия на такие важные геофизические характеристики, как параметр пористости $P_п$ и скорость распространения упругих колебаний v. Анализу подвергнуты данные *, полученные по образцам терригенных коллекторов Узловой площади (Северный Сахалин) в интервале глубин 1500—2600 м. В атмосферных условиях и в условиях, приближенных к пластовым, были определены пористость $K_п$, параметр пористости $P_п$ и

* Результаты измерений любезно предоставлены Л. М. Марморштейном.

Таблица I.6

Результаты двухфакторного дисперсионного анализа с пятью наблюдениями в ячейке (оценка влияния условий измерения и пористости на параметр пористости $P_п$)

Пористость $K_п$ (фактор A), %	Условия измерений (фактор B)		y_{i**}	Главные эффекты α_i
	Атмосферные	Приближенные к пластовым		
5—9	256; 42; 52; 48; 56	308; 34; 36 47; 52	93,1	46,9
10—14	40; 46; 35; 32; 31	42; 31; 50 76; 48	43,1	—3,1
15—19	36; 42; 22; 30; 31	17; 43; 27; 31; 53	33,2	—13,0
20—24	14; 15; 18 15; 11	19; 22; 16 12; 12	14,4	—30,8
y_{*j*}	43,6	48,8	46,2	
Генеральное среднее y_{***}				
Главные эффекты β_j	—2,6	2,6		

Расчетные значения

$$SS_A = 33\,268;\ SS_B = 270;\ SS_{AB} = 91{,}6;\ SS_п = 126\ 944;$$

$$SS_e = 93\,313.$$

$$F_A = 3{,}8;\ F_B = 0{,}09;\ F_{AB} = 0{,}001.$$

Табличные значения

$$F_{0{,}1;\ 3;\ 32} = 2{,}23;\quad F_{0{,}1;\ 1;\ 32} = 2{,}84;\quad F_{0{,}1;\ 3;\ 32} = 2{,}23.$$

скорость распространения упругих волн v. В табл. I.6 и I.7 приведены результаты, сформированные по схеме двухфакторного анализа. Одним из факторов является пористость пород (как известно, пористость вызывает основную изменчивость анализируемых геофизических характеристик), другим — условия проведения измерений (атмосферые или близкие к пластовым).

Результаты дисперсионного анализа показывают, что пластовые условия не оказывают заметного влияния на параметр пористости и можно считать, что этот параметр зависит только от коэффициента пористости (см. табл. I.6). На скорость же распространения упругих колебаний пластовые условия оказывают существенное воздействие (см. табл. I.7). Под их влиянием скорость упругих волн изменяется даже сильнее, чем под воздействием колебаний пористости в диапазоне 5—24%. Отсюда надо сделать вывод, что при расчете скорости по данным, например, акустического каротажа следует вводить поправки за пластовые условия. Значение поправки может быть получено из модели дисперсионного анализа.

Таблица I.7

Результаты двухфакторного дисперсионного анализа с пятью наблюдениями в ячейке (оценка влияния условий измерения и пористости на скорость распространения упругих колебаний v)

Пористость $K_п$ (фактор A), %	Условия измерений (фактор B)		y_{i**}	Главные эффекты α_i
	Атмосферные	Приближенные к пластовым		
5—9	4,7; 3,8; 9; 4,1; 3,8	4,5; 4,7; 4,2; 5,1; 4,7	4,35	0,61
10—14	3,5; 3,1; 3,7; 3,6; 3,4	4,3; 4,3; 4,3; 4,1; 4,8	3,91	0,17
15—19	3,6; 3,2; 3,3; 3,5; 3,3	4,2; 1,4; 4,1; 3,9; 4,2	3,74	0,00
20—24	2,3; 2,2; 2,4; 2,5; 2,7	3,3; 3,7; 4,0 3,2; 3,3	2,96	—0,78
y_{*j*}	3,33	4,15		
Генеральное среднее y_{***}			3,74	
Главные эффекты β_j	—0,41	0,41		

Расчетные значения

$$SS_A = 10{,}0;\ SS_B = 6{,}7;\ SS_{AB} = 0{,}35;\ \overline{SS}_п = 19{,}4;\ \overline{SS}_e = 2{,}26.$$

$$F_A = 47{,}5;\ F_B = 108{,}0;\ F_{AB} = 1{,}68.$$

Табличные значения

$$F_{0{,}1;\ 3;\ 32} = 2.23;\quad F_{0{,}1;\ 1;\ 32} = 2.84;\quad F_{0{,}1;\ 3;\ 32} = 2{,}23.$$

Регрессионный анализ

Другой путь оценки коллекторских свойств и литолого-петрографических характеристик горных пород основан на использовании уравнений, связывающих эти характеристики с геофизическими данными. Очевидно, что из-за наличия случайных флуктуаций эта связь может быть лишь статистической, а не функциональной. Недоучет этого обстоятельства характерен для традиционных методов интерпретации. Статистическую связь одной величины с остальными и позволяет выяснить регрессионный анализ. Уравнение, выражающее зависимость характеристики y (коллекторские свойства или литолого-петрографические данные) от $x_1, x_2, \ldots, x_p$ (различные геофизические данные), здесь имеет линейный вид. Это значит, что в основе регрессионного анализа лежит все та же линейная модель (I.1), которую теперь можно представить в виде

$$y_i = \beta_0 + \beta_1 (x_{i1} - \bar{x}_1) + \beta_2 (x_{i2} - \bar{x}_2) + \ldots + \beta_p (x_{ip} - \bar{x}_p) + e_i.$$

Иногда характеристику, обозначаемую через y, хотя и не совсем приемлемо для данной ситуации, называют «зависимой» переменной, а $\{x_i\}$ — «независимыми» переменными. Очевидно,

что такая терминология, заимствованная из обычной алгебры, не очень удачна. Действительно, сами величины x_i могут быть зависимыми, что вносит известную путаницу. Поэтому, следуя современной терминологии [43], целесообразно величины x_i называть либо предсказателями, либо регрессорами, а характеристику y — предсказуемой или регрессируемой переменной. Как было принято и ранее, e — наблюдаемые флуктуации.

Введем в рассмотрение матрицу следующего вида:

$$\mathbf{X} = \begin{vmatrix} 1 & (x_{11} - \bar{x}_1) & (x_{12} - \bar{x}_2) & \dots & (x_{1p} - \bar{x}_p) \\ 1 & (x_{21} - \bar{x}_1) & (x_{22} - \bar{x}_2) & \dots & (x_{2p} - \bar{x}_p) \\ \dots & \dots & \dots & \dots & \dots \\ 1 & (x_{n1} - \bar{x}_1) & (x_{n2} - \bar{x}_2) & \dots & (x_{np} - \bar{x}_p) \end{vmatrix}.$$

Тогда совокупность определенных значений переменных-предсказателей (признаков) можно обозначить как вектор

$$\mathbf{x}_i = [1, (x_{i1} - \bar{x}_1), (x_{i2} - \bar{x}_2), \dots, (x_{ip} - \bar{x}_p)].$$

Пусть при каждом фиксированном значении $\mathbf{x}_i$ проведено m_i измерений y_{ij} зависимой (предсказуемой) переменной y; выборочное среднее такой серии обозначим $\bar{y}_i$. Всего измерения выполнены при n значениях $\mathbf{x}_i$ или, как еще говорят, в n точках. Получаем вектор $\mathbf{Y}$ со следующими компонентами: $\mathbf{Y}' = (y_1, y_2, \dots, y_n)$.

Тогда в матричных обозначениях модель принимает уже знакомый нам вид:

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{e}.$$

Цель регрессионного анализа заключается в оценке вектора β и проверке некоторых гипотез относительно параметров β_i.

В основе процедуры оценки параметров β_i лежат пять главных предположений Ω [89].

1. Математическое ожидание величины y_i при заданном значении $\mathbf{x}_i$ является линейной по параметрам функцией

$$\mathrm{M}\{y_i/\mathbf{x}_i\} = \eta_i = \beta_0 + \beta_1 (x_{i1} - \bar{x}_1) + \dots + \beta_p (x_{ip} - \bar{x}_p); \ \mathrm{M}\{\mathbf{e}\} = 0.$$

2. Значения $\mathbf{x}_i$, выбранные для изучения, не являются случайными величинами.

3. Дисперсия $\sigma^2_{e_i}$ величины e_i равна дисперсии $\sigma^2_{y_i}$ величины y_i и может быть постоянной или зависимой от $\mathbf{x}_i$. Причем, как известно, дисперсия выборочного среднего $\sigma^2_{\bar{y}_i}$ связана с дисперсией единичного наблюдения $\sigma^2_{y_{ij}}$ соотношением $\sigma^2_{\bar{y}_i} = \sigma^2_{y_{ij}}/m_i$, где m_i — число повторных измерений регрессируемой переменной при данном значении $\mathbf{x}_i$.

4. Различные измерения величины y_i взаимно независимы или, что то же самое, статистически независимы флуктуации e_i, т. е. $\mathrm{M}(\mathbf{e}\mathbf{e}') = 0$.

5. Условное распределение вероятности y_i при заданном значении x_i нормально относительно математического ожидания:

$$\eta_i = M(y_i/x_i).$$

Если наблюдения в каждой из n точек проведены с дисперсией $D\{e_i\} = \sigma_i^2$, то можно составить ковариационную матрицу $\mathbf{f}$ вида

$$\mathbf{f} = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 & 0 \\ 0 & \sigma_2^2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \sigma_n^2 \end{pmatrix}.$$

Если дисперсии во всех точках x_i одинаковы и равны $\sigma^2_{y_i}$, то

$$\mathbf{f} = \sigma^2_{y_i}\mathbf{I},$$

где $\mathbf{I}$ — единичная матрица.

В таких обозначениях оценкой коэффициентов β являются компоненты вектора $\mathbf{b}$:

$$\mathbf{b} = (\mathbf{X}'\mathbf{f}^{-1}\mathbf{X})^{-1}(\mathbf{X}'\mathbf{f}^{-1}\mathbf{Y})$$

с дисперсионной матрицей

$$D(\mathbf{b}) = (\mathbf{X}'\mathbf{f}^{-1}\mathbf{X})^{-1}$$

или, если $\mathbf{f} = \sigma^2_{y_i}\mathbf{I}$, то

$$D(\mathbf{b}) = \sigma^2_{y_i}(\mathbf{X}'\mathbf{X})^{-1}.$$

Оценка $\mathbf{b}$ найдена из условия минимизации взвешенной суммы квадратов флуктуаций (т. е. по методу наименьших квадратов):

$$SS_e = \sum_{i=1}^{n} f_i^{-1} e_i^2 = \mathbf{e}'\mathbf{f}^{-1}\mathbf{e} = (\mathbf{Y} - \mathbf{X}\mathbf{b})'\mathbf{f}^{-1}(\mathbf{Y} - \mathbf{X}\mathbf{b}).$$

Совместная доверительная область для параметров β_i при заданном уровне значимости α дается уравнением

$$(\beta - \mathbf{b})'\mathbf{X}'\mathbf{f}^{-1}\mathbf{X}(\beta - \mathbf{b}) = S^2_{y_i}(p+1)F_{1-\alpha;\, p+1;\, n-(p+1)},$$

где $S^2_{y_i}$ — оценка дисперсии $\sigma^2_{y_i}$.

Регрессионный анализ как способ оценки коэффициентов β_i является наиболее распространенным статистическим методом при обработке результатов геолого-геофизических исследований. Он широко вошел в практику интерпретации. Уже во многих нефтегазоносных провинциях регрессионные уравнения используют для установления фильтрационных, емкостных и других геологических характеристик горных пород по косвенным данным геофизических исследований скважин [58, 68].

К регрессионному анализу очень тесно примыкает корреляционный, позволяющий выявлять разного рода взаимозависимости между измеряемыми параметрами. Следует отметить, что понятия взаимозависимость и зависимость несколько отличны друг от друга (на это обращают внимание М. Кендалл и А. Стьюарт [43]), хотя четкой грани между ними провести невозможно, ибо прежде чем выполнять собственно регрессионный анализ, как правило, обращаются к исследованию матрицы коэффициентов корреляции.

Однако оценка коэффициентов β_j по имеющимся данным не должна являться самоцелью. При исследовании связи характеристики y с характеристиками $\mathbf{x}$ прежде всего необходимо убедиться, что y действительно выражается через $\mathbf{x}$ с помощью формулы (I.1). Для этого используют одну из модификаций дисперсионного анализа, сводящуюся к тому, что разлагается сумма квадратов отклонений $(y_{ij}-\bar{y})$, где $\bar{y} = \sum_{ij} y_{ij} / \sum_i m_i$ — общее среднее. Соответствующее этому случаю разложение дисперсий представлено в табл. I.8; через $\hat{y}_l$ обозначена оценка η_l по наблюдениям: $\hat{y}_l = \mathbf{b}'\mathbf{x}_l$.

Таблица I.8

Распределение вариаций (дисперсий) относительно среднего $\bar{y}$

Отклонения (источник дисперсии)	Число степеней свободы	Сумма квадратов	Средний квадрат
Между значениями на линии регрессии и общим средним (обусловленным регрессией)	p	$\sum_{i=1}^{n} m_i (\hat{y}_i - \bar{y})^2$	$S_y^2 = \dfrac{\sum m_i (\hat{y}_i - \bar{y})^2}{p}$
От эмпирической линии регрессии	$n-p-1$	$\sum_{i=1}^{n} m_i (\bar{y}_i - \hat{y}_i)^2$	$S_r^2 = \dfrac{\sum m_i (\bar{y}_i - \hat{y}_i)^2}{n-p-1}$
Внутри серий (погрешность эксперимента)	$\sum_{i=1}^{n} m_i - n$	$\sum_{i=1}^{n} \sum_{j=1}^{m_i} (y_{ij} - \bar{y}_i)^2$	$S_e^2 = \dfrac{\sum_{i=1}^{n} \sum_{j=1}^{m_i} (y_{ij} - \bar{y}_i)^2}{\sum m_i - n}$
Общее	$\sum_{i=1}^{n} m_i - 1$	$\sum_{i=1}^{n} \sum_{j=1}^{m_i} (y_{ij} - \bar{y})^2$	—

Выражение $\sum\sum (y_{ij}-\bar{y})^2$ есть мера экспериментальной флуктуации, полученной в каждом отдельном эксперименте, выполненном при различных значениях $\mathbf{x}_i$. Член $\sum m_i (\bar{y}_i - \hat{y}_i)^2$ служит мерой эффективности линейной модели при подгонке экспериментальных данных. Если модель корректна, то S_r^2 будет несмещен-

ной оценкой дисперсии $\sigma^2_{\bar{y}_i}$. Величина S^2_r также является несмещенной оценкой $\sigma^2_{\bar{y}_i}$, она служит мерой рассеяния, вызванного экспериментальной погрешностью, а оценка S^2_r характеризует неточность подгонки (неадекватность модели экспериментальным данным).

Следовательно, чтобы проверить, что линейная модель $\eta = \beta_0 + \beta_1 (x_1 - \bar{x}_1) + \ldots + \beta_p (x_p - \bar{x}_p)$ удовлетворительно описывает экспериментальные данные, надо составить отношение дисперсий S^2_r/S^2_e. Если $S^2_r/S^2_e > F_{1-\alpha,\, \nu_1,\, \nu_2}$, то гипотезу об адекватности модели следует отвергнуть. Тогда необходимо выбрать иную модель, которая лучше согласуется с экспериментальными данными, но вместе с тем не следует забывать и о физико-геологической стороне явления. Если модель адекватна, то дисперсии S^2_r и S^2_e можно объединить, чтобы улучшить оценку величины $\sigma^2_{\bar{y}_i}$. При таком объединении каждая из дисперсий входит с весом, пропорциональным соответствующему числу степеней свободы:

$$S^2_{y_i} = \frac{\sum\sum (y_{ij} - \bar{y}_i)^2 + \sum m_i (\bar{y}_i - \hat{y}_i)^2}{\left(\sum m_i - n\right) + (n - p - 1)} = \frac{\sum\sum (y_{ij} - \hat{y}_i)^2}{\sum m_i - p - 1}.$$

Очевидно, что без повторных данных F-критерий не может быть применен для проверки гипотезы линейности. Вообще регрессионный анализ в изложенной постановке очень трудно использовать в практике геолого-геофизических исследований. При составлении регрессионного уравнения исследователь редко располагает повторными наблюдениями. Кроме того, математический аппарат регрессионного анализа, как было уже отмечено, развит в предположении, что регрессоры x являются неслучайными (детерминированными) величинами. На практике требование детерминированности очень часто не выполняется. Помимо ошибок, возникающих при измерении регрессируемых переменных и регрессоров, имеют место случайные изменения этих переменных, обусловленные самой природой.

В этом случае обычный регрессионный анализ, вообще говоря, неприменим. Здесь нужна структурная зависимость, которая позволяет описывать случайные изменения регрессируемой переменной и регрессоров — изменения, которые не обязательно связаны только с погрешностью измерений. Они могут порождаться (и действительно порождаются) самой структурой связи между переменными. Так, например, между пористостью и сопротивлением водонасыщенных пород существует внутреннее (структурное) ослабление связи, совсем не зависящее от погрешности измерений. Можно очень точно измерить пористость пород и их сопротивление, но это не значит, что строго фиксированным значениям пористости (очень близких по составу образцов) будет соответствовать

одно и то же сопротивление пород. В этом случае говорят не о функциональной связи, а о структурном соотношении. Техника построения структурных соотношений по экспериментальным наблюдениям пока что находится в стадии разработки [43].

Определенные попытки учесть погрешности в регрессионных (контролируемых) переменных на основе метода наименьших квадратов и итерационной техники предприняты В. В. Федоровым [84]. Им же отмечен очень важный факт, а именно: при небольших погрешностях в регрессорах и при достаточно гладких функциях $M(\bar{y}|\mathbf{x})$ использование обычных методов регрессионного анализа не приводит к существенным ошибкам. Поэтому, не касаясь обсуждения алгоритма регрессионного анализа при наличии погрешностей в переменных-регрессорах [84], мы обратим внимание на ту его постановку, которая используется на практике.

Итак, при использовании регрессионного анализа вопреки сделанным ранее предположениям обычно полагают, что в точках $\mathbf{x}_l$ выполнено по одному измерению, дисперсия $\sigma^2_{e_l}$ величины e_l постоянна для всех $\mathbf{x}_l$ (не зависит от $\mathbf{x}$). Тогда оценку вектора β по методу наименьших квадратов дает следующее матричное выражение:

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}.$$

Если элементы столбцов матрицы $\mathbf{X}$ не только центрировать, но и нормировать, т. е. от матрицы $\mathbf{X}$ перейти к матрице

$$\bar{\mathbf{X}} = \begin{pmatrix} \frac{x_{11}-\bar{x}_1}{\sigma_x} & \frac{x_{12}-\bar{x}_2}{\sigma_{x_2}} & \cdots & \frac{x_{1p}-\bar{x}_p}{\sigma_{x_p}} \\ \frac{x_{21}-\bar{x}_1}{\sigma_{x_1}} & \frac{x_{22}-\bar{x}_2}{\sigma_{x_2}} & \cdots & \frac{x_{2p}-\bar{x}_p}{\sigma_{x_p}} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{x_{n1}-\bar{x}_1}{\sigma_{x_1}} & \frac{x_{n2}-\bar{x}_2}{\sigma_{x_2}} & \cdots & \frac{x_{np}-\bar{x}_p}{\sigma_{x_p}} \end{pmatrix},$$

а элементы вектора $\mathbf{Y}$ также центрировать и нормировать, получив

$$\bar{\mathbf{Y}}' = \left(\frac{y_1-\bar{y}}{\sigma_y}, \frac{y_2-\bar{y}}{\sigma_y}, \ldots, \frac{y_n-\bar{y}}{\sigma_y}\right),$$

то выражение для $\mathbf{b}$ примет вид

$$\mathbf{b} = \mathbf{R}^{-1}\mathbf{r}.$$

Это получилось потому, что произведение матриц $\bar{\mathbf{X}}'\bar{\mathbf{X}}$ с центрированными и нормированными вектор-столбцами дает с точностью до постоянного множителя матрицу парных коэффициентов корреляции между регрессорами, т. е. $(n-1)\mathbf{R}$, в то время как $\bar{\mathbf{X}}'\bar{\mathbf{Y}} = (n-1)\mathbf{r}$, где $\mathbf{r}$ есть не что иное, как вектор парных

коэффициентов корреляции между регрессируемой переменной и регрессорами.

Здесь отчетливо видно, что для расчетов коэффициентов регрессии **b** необходимо выполнить обращение корреляционной матрицы **R**. Однако прямое обращение, особенно при большом размере матрицы или при ее элементах, близких к единице (т. е. когда переменные x_i сильно связаны между собой), часто бывает нецелесообразным из-за накопления значительных ошибок, возникающих вследствие того, что определитель матрицы **R** может оказаться близким к нулю. В последнем случае получаются крайне неустойчивые оценки коэффициентов регрессии, т. е. уравнения регрессии могут не отражать реальной связи, а следовательно, **Y** будет «плохо восстанавливаться» по **x**.

Эти достаточно неприятные особенности регрессионного анализа в какой-то мере можно преодолеть, если сначала провести некоторые математические преобразования матрицы **R**, а именно ее ортогонализацию. Такой способ нашел широкое применение при обработке метеорологических, биологических и сельскохозяйственных экспериментальных данных [5, 64, 75]. Нами этот метод также был использован при обработке результатов геолого-геофизических исследований скважин с целью оценки различных типов пористости карбонатных пород [12, 36, 46]. Обстоятельная речь об этом методе пойдет в следующей главе, ибо способ ортогонализации корреляционной матрицы естественно вытекает из теории метода главных компонент (факторного анализа).

Выше мы уже отмечали, что при выявлении соотношений между интересующими нас характеристиками горных пород и геофизическими данными нельзя сводить задачу лишь к вычислению коэффициентов $\{\beta_j\}$. Важно выяснить, какие из геофизических показателей несут информацию о восстанавливаемом свойстве породы, а какие нет. В связи с этим возникает проблема отсеивания или, наоборот, включения тех или иных переменных-регрессоров в модель, т. е. проблема построения модели. От правильности выбора регрессоров зависит устойчивость регрессионного уравнения, а значит, и верность определения искомой характеристики. При построении регрессионных уравнений в практических применениях этот вопрос обычно или обходится стороной, или решается по наитию.

Естественно, что исходя из предметных, содержательных, физических и геологических представлений сначала следует составить перечень тех характеристик, которые могут оказать влияние на предсказуемую переменную, а уж затем решать вопрос о том, какие из перечисленных характеристик должны быть включены в модель и в каком виде (результаты непосредственных наблюдений или какие-либо функции от них: логарифмы, обратные величины и т. д.). На первый взгляд кажется, что если нас интересует задача предсказания одной величины (y) по другим, то заманчиво было бы вычислить парные коэффициенты корреля-

ции между y и теми величинами, которые предполагается включить в регрессионное уравнение. Затем те переменные, которые с регрессируемой (y) образуют нулевые или малые коэффициенты корреляции, следует отбросить. Тем самым, казалось бы, вопрос об информативности регрессоров решен. Подобным образом задача об отсеивании и информативности регрессоров решается у ряда исследователей, например в программе «Оракул» М. Д. Белопина, однако это неверно.

К сожалению, подобного рода процедура может привести к серьезному заблуждению. Совершенно справедливо, что множество парных коэффициентов корреляции полностью определяет всю совокупность так называемых частных коэффициентов корреляции, которые отражают меру взаимозависимости между двумя величинами при фиксированных значениях остальных и на которые как раз и следует ориентироваться при формировании матрицы X. Действительно, если одна величина коррелирована с другой, то это может быть отражением того факта, что они обе коррелированы с некоторой третьей величиной или с совокупностью величин, — ситуация, типичная в геолого-геофизических исследованиях. Если действуют побочные факторы, связь между двумя переменными может в определенной мере ослабнуть, однако если побочные факторы фиксированы (устранены), сила связи между этими переменными может существенно возрасти.

В качестве примера определим выражение для расчета частных коэффициентов корреляции. Если имеется матрица $\mathbf{R}$ парных коэффициентов корреляции размером $p \times p$, то оценка парного частного коэффициента корреляции $r_{ij\cdot(k, l, \ldots, p)}$ (в скобках приводятся номера фиксированных переменных) дается формулой

$$r_{ij\cdot(k,l,\ldots,p)} = -|\mathbf{R}_{ij}|\,(|\mathbf{R}_{ii}|\,|\mathbf{R}_{jj}|)^{-1/2},$$

где $\mathbf{R}_{ij}$ — алгебраическое дополнение для элемента r_{ij} в определителе

$$|\mathbf{R}| = \begin{vmatrix} 1 & r_{12} & r_{13} & \ldots & r_{1p} \\ r_{21} & 1 & r_{23} & \ldots & r_{2p} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ r_{p1} & r_{p2} & r_{p3} & \ldots & 1 \end{vmatrix}.$$

Так, для трех переменных x_1, x_2, x_3 с матрицей парных коэффициентов корреляции

$$\mathbf{R} = \begin{pmatrix} 1 & r_{12} & r_{13} \\ r_{21} & 1 & r_{23} \\ r_{31} & r_{32} & 1 \end{pmatrix}$$

частный коэффициент корреляции между первой и второй переменной при фиксированной третьей дается выражением

$$r_{12\cdot(3)} = \frac{-|\mathbf{R}_{12}|}{(|\mathbf{R}_{11}|\,|\mathbf{R}_{22}|)^{1/2}} = \frac{r_{12} - r_{13}r_{23}}{\left[\left(1-r_{13}^2\right)\left(1-r_{23}^2\right)\right]^{1/2}}.$$

Возьмем следующую симметричную матрицу парных коэффициентов корреляции, характеризующую взаимозависимость между пористостью и данными геофизических методов * [58].

	$K_п$	$\Delta U_{ПС}$	$J_{НГМ}$	$J_{ГМ}$	$\rho_{ПЗ}$	$\rho_{ГЗ}$	$h_п$
$K_п$	1,00						
$\Delta U_{ПС}$	0,79	1.00					
$J_{НГМ}$	—0,43	—0,04	1,00				
$J_{ГМ}$	—0.51	—0,64	—0,37	1.00			
$\rho_{ПЗ}$	—0,29	—0,40	0,50	—0,75	1,00		
$\rho_{ГЗ}$	—0,35	—0,22	0,80	—0,38	0,72	1,00	
$h_п$	0,39	0,01	—0,47	—0,15	—0,33	—0,06	1,00

Эта матрица получена в результате обработки геолого-геофизического материала, характеризующего пашийские слои девона на Туймазинском месторождении. Анализируемые пласты (их число равно 31) являются нефтеносными, они вскрыты при бурении на нефтеэмульсионном растворе. Пласты представлены мелкозернистыми кварцевыми песчаниками и алевролитами, достаточно однородными по составу. Известно, что данные НГМ и метода ПС контролируют изменение пористости и глинистости, а данные ГМ — в основном изменение глинистости.

Значения парных коэффициентов корреляции не очень высокие. Рассчитаем частные коэффициенты корреляции между пористостью и данными НГМ при фиксированной глинистости:

$$r_{13\cdot(4)} = (r_{13} - r_{14}r_{34})\left[(1 - r_{14}^2)(1 - r_{34}^2)\right]^{-1/2} = -0{,}78,$$

и между пористостью и данными метода ПС при фиксированной глинистости (при фиксированных данных ГМ):

$$r_{12\cdot(4)} = (r_{12} - r_{14}r_{24})\left[(1 - r_{14}^2)(1 - r_{24}^2)\right]^{-1/2} = 0{,}70.$$

Как видим, частный коэффициент корреляции между пористостью и данными НГМ резко возрос по сравнению с парным коэффициентом, в то время как между пористостью и данными метода ПС коэффициент остался практически на прежнем уровне. Если бы при формировании матрицы регрессоров мы пользовались информацией о парных коэффициентах корреляции, то данные НГМ, вероятно, были бы отброшены и тем самым была бы допущена серьезная ошибка.

---

* В книге встречаются следующие методы каротажа: АМ — акустический, БМ — боковой, БЭЗ — бокового электрозондирования, ГЗ — градиент-зонда, ГМ — гамма-, ИМ — индукционный, КС — сопротивления (кажущегося), МБК — микробоковой, МГЗ — микроградиент-зонда, МПЗ — микропотенциал-зонда, НГМ — нейтронный гамма- (НГМ-т — то же, нейтроны тепловой энергии), ПЗ — потенциал-зонда, ПС — потенциалов собственной поляризации (естественного электрического поля), УЭС — удельного электрического сопротивления.

Для отсеивания тех или иных регрессоров из составленного списка или из полной модели может быть полезной одна из форм дисперсионного анализа [89], которая сводится к проверке значимости одного или нескольких регрессоров матрицы $\mathbf{X}$, описывающей полную модель. Иными словами, проверяется гипотеза, что соответствующие компоненты вектора β равны нулю ($\beta_i = 0$). В частности, можно проверить наличие регрессии, т. е. все $\beta_i = 0$, кроме β_0. Для этого вычисляют сумму квадратов, соответствующую исключению одного, двух, трех и т. д. регрессоров из модели, первоначально содержащей все регрессоры; таким образом осуществляется разложение суммы квадратов разностей между значениями, предсказанными по уравнению регрессии, и средним: $\sum(\bar{y}-\hat{y}_i)^2$. К сумме квадратов, выделенной из $\sum(\bar{y}-\hat{y}_i)^2$, можно применить F-критерий. В табл. I.9 приведен алгоритм расщепления $\sum(\bar{y}-\hat{y}_i)^2$ и попутно указаны соответствующие степени свободы и дисперсии, которые используются в F-критерии.

Таблица I.

Дисперсионный анализ при последовательном отсеивании регрессоров

Источник дисперсии (рассеяния)	Число степеней свободы	Сумма квадратов	Средний квадрат
Исключается x_1 (т. е. $b_1=0$)	1	$\Delta SS_1 = b_1^2/c_{11}$; $\Delta SS_1 = \Delta SS_1 - 0 = \Delta SS_1$	$\frac{b^2}{c_{11}} = S^2$
Исключается x_2 (т. е. $b_2=0$) после исключения x_1	1	$\Delta SS_{1+2} = \mathbf{b}_2\,\mathbf{C}^{-1}\mathbf{b}_2$; $\Delta SS_2 = \Delta SS_{1+2} - \Delta SS_1$	$\frac{\Delta SS_2}{1}$
Исключается x_3 (т. е. $b_3=0$) после исключения x_1 и x_2 (и т. д.)	1	$\Delta SS_{1+2+3} = \mathbf{b}_3'\mathbf{C}^{-1}\mathbf{b}_3$; $\Delta SS_3 = \Delta SS_{1+2+3} - \Delta SS_{1+2}$	$\frac{\Delta SS_3}{1}$
Неполная сумма	p	$\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$	—
Исключается свободный член b на последнем шаге	1	$\sum_{i=1}^{n}(\bar{y}-0)^2 = n\bar{y}^2$	$\frac{n\bar{y}^2}{1}$
Отклонение от линии регрессии	$n-p-1$	$\sum_{i=1}^{n}(y_i-\hat{y}_i)^2 = (\mathbf{Y}-\mathbf{Xb})'\,\mathbf{f}^{-1}\,(\mathbf{Y}-\mathbf{Xb})$	$S_2^2 = \frac{\sum(y_i-\hat{y}_i)^2}{n-p-1}$
Полная сумма	n	$\sum_{i=1}^{n} y_i^2$	—

Поясним таблицу. Приведенная в таблице матрица имеет вид $\mathbf{C} = (\mathbf{X}'\mathbf{X})^{-1}$, ее элементы обозначены c_{ij}. Исключение одного из членов модели равнозначно условию $\beta_k = 0$. Тогда разность между суммой квадратов $\sum (y_i - \bar{y})^2$ для первоначальной модели и модели с $\beta_k = 0$ равна $\Delta SS = b_k^2/c_{kk}$. Величину ΔSS часто называют суммой квадратов для переменных x_k, скорректированной на все другие регрессоры, или суммой квадратов при заданных других регрессорах.

Сумма квадратов для любой группы из q регрессоров, скорректированная на все другие регрессоры, вычисляется по формуле

$$\Delta SS = \mathbf{b}_q' \mathbf{C}_q^{-1} \mathbf{b}_q,$$

где $\mathbf{b}_q$ — вектор, составленный только из выбранной группы параметров b_k; $\mathbf{C}_q$ — матрица из соответствующих элементов c_{jk}.

Очевидно, что матричное выражение вида $\mathbf{b}_q' \mathbf{C}_q^{-1} \mathbf{b}_q$ представляет собой ту часть дисперсии, которая извлекается данными регрессорами из дисперсии регрессируемой переменной y.

Таблица дисперсионного анализа показывает, что исключение каждой дополнительной переменной дает новую совместную сумму ΔSS, вычитая из которой предыдущую сумму, можно найти остаточный эффект удаления этого регрессора. Сумма всех величин ΔSS, вычисленных таким способом для каждого этапа, равна сумме квадратов $\sum_{i=1}^{n} (y_i - \bar{y})^2$.

Чтобы проверить значимость некоторого регрессора (или их группы), можно применить F-критерий для отношения дисперсий $S^2/S_{y_i}^2$, где величина S^2 определяется следующим образом.

ΔSS	ν	S^2
b_k^2/c_{kk}	1	$\Delta SS_k = b_k^2/c_{kk}$
$\mathbf{b}_q' \mathbf{C}_q^{-1} \mathbf{b}$	q	$\mathbf{b}_q' \mathbf{C}_q^{-1} \mathbf{b}_q/q = \Delta SS_{k_1+k_2+\ldots+k_q}/q$

При интерпретации результатов предложенным способом надо быть в какой-то мере осторожными. Вклад, вносимый в модель регрессором x_k, в действительности может зависеть от некоторого регрессора x_n, который даже и не наблюдается, но сильно коррелирован с x_k. Поэтому здесь имеет значение порядок исключения регрессоров, в зависимости от этого вообще могут получаться разные суммы ΔSS_l.

Рассмотрим вопрос о построении регрессионной модели на примере оценки пористости пашийских слоев девона Туймазин-

ского месторождения по данным различных геофизических методов. В качестве исходных возьмем регрессоры, приведенные выше в виде корреляционной матрицы. По-прежнему будем оперировать с нормированными значениями регрессируемой переменной (пористость) и регрессоров (данные геофизических методов).

Оцепим коэффициенты $\mathbf{b}$ при нормированных регрессорах по формуле $\mathbf{b} = \mathbf{R}^{-1}\mathbf{r} = \mathbf{C}\mathbf{r}$, где $\mathbf{R}$ — матрица, полученная из исходной путем вычеркивания первой строки (и первого столбца). Очевидно, что элементы строки (столбца) представляют собой компоненты вектора $\mathbf{r}$ коэффициентов корреляции между регрессируемой переменной и регрессорами. При этом матрица, обратная корреляционной матрице $\mathbf{R}$, равна

$$\mathbf{C} = \mathbf{R}^{-1} = \begin{vmatrix} 0{,}749 & -2{,}093 & -0{,}410 & -1{,}121 & 2{,}414 & -1{,}278 \\ -2{,}093 & 4{,}958 & -1{,}543 & -0{,}852 & -0{,}305 & 1{,}580 \\ -0{,}410 & -1{,}543 & -0{,}337 & -1{,}458 & 1{,}998 & -1{,}133 \\ -1{,}121 & -0{,}852 & -1{,}458 & -0{,}519 & 0{,}209 & -0{,}766 \\ 2{,}414 & -0{,}305 & 1{,}997 & 0{,}209 & 5{,}502 & -1{,}349 \\ -1{,}278 & 1{,}580 & -1{,}133 & -0{,}766 & -1{,}349 & 1{,}252 \end{vmatrix}.$$

Отсюда вектор $\mathbf{b}$ имеет следующие компоненты: $\mathbf{b}' = (0{,}682, -0{,}628, -0{,}207, -0{,}006, 0{,}227, 0{,}071)$, а уравнение регрессии принимает вид

$$K_{\text{п}} = 0{,}682\,\Delta U_{\text{пс}} - 0{,}628 J_{\text{нгм}} - 0{,}207 J_{\text{гм}} - \\ - 0{,}006\rho_{\text{пз}} + 0{,}227\rho_{\text{гз}} + 0{,}071 h_{\text{п}}.$$

Возникает естественный вопрос: какие из регрессоров следует отсеять, т. е. при каких регрессорах коэффициенты β_j пренебрежимо малы и их можно считать равными пулю? Будем решать этот вопрос по только что изложенной процедуре. Напомним, что F-критерий строится как отношение

$$F = \frac{\nu_e}{\nu_H} \cdot \frac{SS_\omega - SS_e}{SS_e}.$$

В данном случае предположения Ω сводятся к тому, что используется полная модель (6 регрессоров), а предположения ω заключаются в том, что те или иные члены модели незначимы, т. е. некоторые $\beta_i = 0$. Нетрудно показать, что для полной модели $SS_e = \sum K_{\text{п}}^2 - (n-1)\mathbf{r}_6'\mathbf{b}_6$, а для усеченной $SS_\omega = \sum K_{\text{п}}^2 - (n-1)\mathbf{r}_k'\mathbf{b}_k$, где k — число членов, включенных в модель. Таким образом, член $\mathbf{r}_k'\mathbf{b}_k$ показывает ту долю дисперсии $K_{\text{п}}$, которая исчерпывается регрессией на k переменных, включенных в модель; разность же $SS_\omega - SS_e = (n-1)(\mathbf{r}_6'\mathbf{b}_6 - \mathbf{r}_k'\mathbf{b}_k)$.

Последовательно вводя в модель все 6 членов, затем 5; 4 и т. д., получаем

$$\mathbf{b}'_6 = (0{,}682,\ -0{,}628,\ -0{,}207,\ -0{,}006,\ 0{,}227,\ 0{,}071);$$

$$\mathbf{b}'_5 = (0{,}757,\ -0{,}718,\ -0{,}142,\ 0{,}046,\ 0{,}305);$$

$$\mathbf{b}'_4 = (0{,}678,\ -0{,}524,\ -0{,}200,\ 0{,}092);$$

$$\mathbf{b}'_3 = (0{,}529,\ -0{,}547,\ -0{,}375);$$

$$\mathbf{b}'_2 = (0{,}79,\ -0{,}43);$$

$$\mathbf{b}'_1 = (0{,}79).$$

Ниже приведены соответствующие расчетные и табличные значения F-критерия при последовательном отбрасывании членов модели.

Исключаемые из модели члены	$F_{\text{расчет}}$	$F_{\text{табл}}$
$h_{\text{п}}$	0,919	2,92
$\rho_{\text{гз}}$, $h_{\text{п}}$	2,48	2,53
$\rho_{\text{пз}}$, $\rho_{\text{гз}}$, $h_{\text{п}}$	1,22	2,32
$J_{\text{гм}}$, $\rho_{\text{пз}}$, $\rho_{\text{гз}}$, $h_{\text{п}}$	2,57	2,18
$J_{\text{нгм}}$, $J_{\text{гм}}$, $\rho_{\text{пз}}$, $\rho_{\text{гз}}$, $h_{\text{п}}$	8,82	2,09
$\Delta U_{\text{пс}}$, $J_{\text{нгм}}$, $J_{\text{гм}}$, $\rho_{\text{пз}}$, $\rho_{\text{гз}}$, $h_{\text{п}}$	26,47	2,02

Видно, что значимый вклад в модель вносят только первые три члена, ибо исключение трех последних к изменению дисперсии (в статистическом смысле) не приводит ($F_{\text{расчет}} < F_{\text{табл}}$). Таким образом, первоначальная модель

$$K_{\text{п}} = 0{,}682\,\Delta U_{\text{пс}} - 0{,}628 J_{\text{нгм}} - 0{,}207 J_{\text{гм}} -$$

$$- 0{,}006\rho_{\text{пз}} + 0{,}227\rho_{\text{гз}} + 0{,}071 h_{\text{п}}$$

без ущерба может быть заменена моделью

$$K_{\text{п}} = 0{,}529\,\Delta U_{\text{пс}} - 0{,}547 J_{\text{нгм}} - 0{,}375 J_{\text{гм}}.$$

Расчеты показывают, что первые три члена модели исчерпывают 84,4% дисперсии $K_{\text{п}}$, первые два — 80,8, а один первый — 62,4, в то время как все шесть — 86,4. Соответственно, если один первый член исчерпывает 62,4% дисперсии $K_{\text{п}}$, то второй — 18,4, а третий — 3,6. Отсюда можно по F-критерию проверить значимость уже не всех трех оставшихся членов модели, а каждого в отдельности.

Испытываемая переменная	$F_{\text{расчет}}$	Табличные значения $F_{\text{табл}}$ (при уровне значимости 10%)
$J_{\text{гм}}$	6,61	2,92
$J_{\text{нгм}}$	33,82	2,92
$\Delta U_{\text{пс}}$	114,71	2,92

Видно, что каждый член модели приводит к значимому уменьшению дисперсии K_{nn} и следовательно, должен быть включен как регрессор в модель.

Таким образом, в данном случае практически только результаты ГМ, НГМ и ПС несут информацию о пористости. Вопрос о выборе наилучшей модели можно поставить и иначе, а именно: в каких регрессорах лучше всего отражаются значения предсказуемой переменной? Мы уже видели, что частный коэффициент корреляции несет в себе подобную информацию. Поэтому, принимая решение, какой из регрессоров ввести в модель, целесообразно использовать частный коэффициент корреляции между y и каждой из многих переменных x_j.

На этом основан так называемый шаговый регрессионный метод. Он состоит в последовательном включении каждого из регрессоров в первоначальную линейную модель и проведении на каждом этапе проверки, является ли добавляемая переменная значимой. Сначала к модели добавляется регрессор с наибольшим частным коэффициентом корреляции $r_{y|x_j(x)}$, т. е. тот, который вызывает наибольшее уменьшение суммы квадратов остатков при условии, что это уменьшение является значимым. Для этого необходимо вычислить частный коэффициент корреляции

$$r_{y|x_j(x_i)} = (b_j/S_{b_j})\,(b_j/S_{b_j} + n - p - 1)^{-1/2}$$

и проверить его значимость по t-критерию:

$$t = r_{y|x_j(x_i)}\left[\left(1 - r^2_{y|x_j(x_i)}\right)/\nu\right]^{-1/2}.$$

Затем можно добавить в модель второй регрессор, который имеет наибольший частный коэффициент корреляции среди оставшихся и т. д. В формуле для t-критерия число ν для первой переменной равно $n-2$, для второй $n-3$ и т. д., т. е. число степеней свободы равно числу наблюдений минус число членов, связанных корреляцией.

Надо отметить, что шаговая регрессия эффективна в том случае, когда переменные-регрессоры ортогональны (независимы), т. е. порядок включения или отсеивания регрессоров не играет роли. При использовании неортогональных переменных различные исходные модели могут привести к различным окончательным уравнениям. Может оказаться, что со статистической точки зрения некоторые модели получатся адекватными друг другу и каждая из них будет иметь право на существование. Какую из этих моделей следует считать лучшей — вопрос уже не статистический, а практический.

В этом случае надо ориентироваться на ряд физических или иных априорных сведений, лежащих в основе рассматриваемого явления. Именно априорные данные в некоторых ситуациях могут точно указать, какая существует форма связи между предсказу-

емой переменной и регрессорами. Например, хорошо известен факт уменьшения межзерновой пористости с глубиной. Закон этот, вообще говоря, очень сложный. Но иногда экспериментальные данные показывают, что уменьшение можно считать линейным или экспоненциальным; более того, статистические выводы могут установить адекватность этих двух моделей. Но с физической точки зрения экспоненциальная модель более логична и содержательна, так как ожидать, что на больших глубинах можно встретить такой порог, где осадочные породы имеют точно нулевые значения пористости, абсурдно. Значения пористости могут быть близки к нулю, но не равны ему, что и отражает экспоненциальная модель. Однако при решении частных задач, если требуется получить определенные оперативные выводы, в качестве первого приближения можно довольствоваться линейной моделью (например, когда рассматривается узкий стратиграфический диапазон ограниченного региона, где пористость далека от нулевых значений). Таким образом, иногда выбор модели может зависеть от условий, которые налагаются при решении той или иной задачи.

Ковариационный анализ

При восстановлении коллекторских свойств пород и определении их литолого-петрографических характеристик по геофизическим данным встречается достаточное число задач, своеобразие которых будет ясно из следующего примера. Известно, что для пород различного литологического состава уравнения регрессии, связывающие межзерновую пористость $K_п$ с параметром пористости $P_п$, нередко имеют различные константы. Так, известное уравнение Арчи $P_п = bK_п^{-a}$ для многих пород сохраняет одно и то же алгебраическое выражение, но параметры b и a принимают разные значения в зависимости от типа пород и структуры их порового пространства. В билогарифмическом масштабе формула Арчи имеет вид уравнения прямой

$$\ln P_п = \ln b - a \ln K_п.$$

При обработке образцов, принадлежащих породам разного литологического состава, часто возникает вопрос: можно ли для исследования всей совокупности образцов использовать уравнение Арчи с постоянными коэффициентами b и a, т. е. считать, что в данном конкретном случае эти коэффициенты не зависят от литологического типа пород, или это предположение в корне неверно и для каждого типа пород надо определять свои параметры b_i и a_i. Таким образом, здесь ставится вопрос о влиянии на параметр пористости как количественно заданного фактора (пористость, связанная с параметром пористости соответствующим уравнением регрессии), так и качественного (литологический состав пород, влияние которого проявляется в том, что изменяются параметры уравнения).

Очевидно, что в терминах методов анализа этот же вопрос можно сформулировать и так: при наблюдаемом разбросе (рассеянии) точек вокруг некоторых прямых регрессии, относящихся к разным (по литологическому составу) группам образцов, можно ли считать эти прямые эквивалентными или они отличаются друг от друга? Такие задачи возникают и при числе регрессоров больше одного, в качестве «дискретного» фактора могут фигурировать кроме литологического состава другие параметры (термобарические условия, географическая или стратиграфическая принадлежность выборки).

Выше уже отмечалось, что дисперсионный анализ используется тогда, когда все факторы исследуются качественно. В регрессионном анализе все факторы являются количественными. В случаях, отвечающих рассмотренной ситуации, когда одна часть факторов (переменных) задается качественно, а другая — количественно, целесообразно применять ковариационный анализ.

Рассмотрим модель ковариационного анализа. Математическое ожидание наблюдений $M(y_i)$ равно сумме двух выражений: первое обычно рассматривается в дисперсионном анализе, а второе — в регрессионном (линейная комбинация значений регрессоров с коэффициентами). Первое выражение представляет собой сумму

$$\sum_{j=1}^{p} x_{jl}\beta_j,$$

где $\{x_{jl}\}$ — элементы матрицы $\mathbf{X}$, которые принимают значения 0 либо 1; $\{\beta_j\}$ — эффекты, рассматриваемые в дисперсионном анализе.

Второе выражение должно быть суммой вида

$$\sum_{j=1}^{h} z_{jl}\gamma_j,$$

где $\{z_{jl}\}$ — значения регрессоров.

Таким образом, в матричных обозначениях модель ковариационного анализа запишется в следующем виде [92]:

$$\mathbf{M}(y) = \mathbf{X}'\boldsymbol{\beta} + \mathbf{Z}'\boldsymbol{\gamma}.$$

Ответ на задачу ковариационного анализа можно получить, решив соответствующую задачу дисперсионного анализа. Обозначим основные предположения задачи ковариационного анализа через $\bar{\Omega}$, а символ Ω сохраним для основных предположений соответствующей задачи дисперсионного анализа [92].

Введем $(h \times h)$-матрицу $\mathbf{M}_{\bar{\Omega}}$ и $(h \times 1)$-вектор $\mathbf{m}_{\bar{\Omega}}$, так что (ij)-м элементом матрицы $\mathbf{M}_{\bar{\Omega}}$ является $m_{z_i z_j \bar{\Omega}}$, а $(i1)$-й элемент вектора $\mathbf{m}_{\bar{\Omega}}$ равен $m_{z_i y \bar{\Omega}}$. Вычисляются эти элементы следующим образом.

Определяется выражение для SS_e в обычном дисперсионном анализе, т. е. в предположениях Ω. Затем сумма квадратов представляется как сумма произведений двух одинаковых сомножителей. В одном из них переменные заменяются на z_i, а в другом на z_j, если надо получить элемент $m_{z_i z_j \Omega}$, или на y, если нужен элемент $m_{z_i y \Omega}$. Так, в двухфакторном анализе с одним наблюдением в ячейке

$$SS_e = SS_\Omega = \sum_i \sum_j (y_{ij} - y_{i*} - y_{*j} + y_{**})^2.$$

Эту сумму можно переписать как

$$SS_e = SS_\Omega = \sum_i \sum_j y_{ij}^2 - J \sum_i y_{i*}^2 - I \sum_j y_{*j}^2 + IJ y_{**}^2.$$

Тогда элемент $m_{zy\Omega}$ (где $z = z_i$) равен

$$m_{zy\Omega} = \sum_i \sum_j z_{ij} y_{ij} - J \sum_i z_{i*} y_{i*} - I \sum_i z_{*i} y_{*i} + IJ z_{**} y_{**}.$$

Определив матрицы $\mathbf{M}_\Omega$ и вектор $\mathbf{m}_\Omega$, можно найти вектор коэффициентов регрессии $\gamma_{\tilde{\Omega}}$ из решения уравнения

$$\mathbf{M}_\Omega \boldsymbol{\gamma}_{\tilde{\Omega}} = \mathbf{m}_\Omega.$$

В предположениях $\tilde{\Omega}$ ковариационная матрица $\gamma_{\tilde{\Omega}}$ равна $S^2\mathbf{M}^{-1}$, где $S^2 = \overline{SS}_{\tilde{\Omega}} = SS_e/\nu_e$. Величина

$$SS_e = SS_{\tilde{\Omega}} = SS_\Omega - \mathbf{m}'_\Omega \boldsymbol{\gamma}_{\tilde{\Omega}}.$$

Число степеней свободы $\nu_e = n - r + h$, где n — общее число наблюдений; r — ранг матрицы $\mathbf{X}$; h — число регрессоров.

Если проверяется гипотеза H_γ, состоящая в том, что $\gamma = 0$, то $\Omega = H_\gamma \cap \tilde{\Omega}$ и числитель статистики F равен

$$SS_H = SS_\Omega - SS_{\tilde{\Omega}} = \mathbf{m}'_\Omega \boldsymbol{\gamma}_{\tilde{\Omega}}.$$

При вычислении коэффициента $\beta_{\tilde{\Omega}}$ в предположениях $\tilde{\Omega}$ необходимо иметь в виду, что в задаче дисперсионного анализа каждое значение $\beta_{i\Omega}$ ($i = 1, 2, \ldots, p$) является известной линейной формой предсказуемой (регрессируемой) переменной y, например:

$$\beta_{i\Omega} = l_i(y).$$

Образуем такую же линейную форму для каждого регрессора, а именно: $l_i(z_j)$ ($j = 1, 2, \ldots, h$).

Тогда при ковариационном анализе в предположениях $\tilde{\Omega}$ оценка

$$\beta_{i\tilde{\Omega}} = l_i(y) - \sum_{j=1}^{h} \gamma_{i\tilde{\Omega}} l_i(z_j).$$

Член $\gamma_{i\bar{\Omega}} l_i(z_j)$ естественно назвать поправкой к $\hat{\beta}_{i\bar{\Omega}}$ от регрессии по z_j.

Если проверяется гипотеза H_β, состоящая в том, что $\beta_1 = \hat{\beta}_{\bar{\Omega}} = \ldots = \beta_p$, то предположения дисперсионного и ковариационного анализов имеют вид

$$\omega = H_\beta \cap \Omega \quad \text{и} \quad \bar{\omega} = H_\beta \cap \bar{\Omega}.$$

Пусть через $\mathbf{M}_\omega$ обозначена $(h \times h)$-матрица, элемент (ij) которой равен $m_{z_i z_j \omega}$, а через $\mathbf{m}_\omega$ обозначен $(h \times 1)$-вектор с элементом $(i1)$, равным $m_{z_i y \omega}$. Тогда

$$\mathbf{M}_\omega \gamma_{\bar{\omega}} = \mathbf{m}_\omega.$$

В этих условиях

$$SS_{\bar{\omega}} = SS_\omega - \mathbf{m}'_\omega \gamma_{\bar{\omega}}.$$

Сумма квадратов числителя статистики F для проверки гипотезы H_β в ковариационном анализе вычисляется как $SS_{\bar{\omega}} - SS_{\bar{\Omega}}$, а сумма квадратов знаменателя равна $SS_{\bar{\Omega}}$, при этом числами степеней свободы являются q и $n-r+h$, где q — число линейно независимых ограничений на коэффициенты $\{\beta_i\}$.

Рассмотрим частный пример модели ковариационного анализа, часто встречающийся в практике.

Для случая однофакторного анализа с одним регрессором основные предположения $\bar{\Omega}$ задаются в форме

$$\bar{\Omega}: \begin{cases} y_{ij} = \beta_i + \gamma z_{ij} + e_{ij} \quad (i = 1, 2, \ldots, I; \quad j = 1, 2, \ldots, J_i); \\ \{e_{ij}\} \text{ независимы и имеют распределение } N(0, \sigma^2). \end{cases}$$

Соответствующий дисперсионный анализ в предположениях Ω, по которым $\gamma = 0$, дает суммы SS_Ω (SS_e) в виде

$$SS_\Omega = \sum_i \sum_j (y_{ij} - y_{i*})^2$$

или

$$SS_\Omega = \sum_i \sum_j y_{ij}^2 - \sum_i J_i y_{i*}^2 = m_{yy\Omega}.$$

В этих условиях вместо матрицы $\mathbf{M}_\Omega$ имеем просто величину

$$m_{zz\Omega} = \sum_i \sum_j z_{ij}^2 - \sum_i J_i z_{i*}^2,$$

а вместо вектора $\mathbf{m}_\Omega$ одну его компоненту

$$m_{zy\Omega} = \sum_i \sum_j z_{ij} y_{ij} - \sum_i J_i z_{i*} y_{i*}.$$

Следовательно, уравнение $\mathbf{M}_\Omega \gamma_{\bar{\Omega}} = \mathbf{m}_\Omega$ сводится к равенству $m_{zz\Omega} \gamma_{\bar{\Omega}} = m_{zy\Omega}$, откуда $\gamma_{\bar{\Omega}} = m_{zy\Omega}/m_{zz\Omega}$. Теперь $SS_{\bar{\Omega}} = SS_\Omega -$

$- m_{zy\Omega}\gamma_{\tilde{\Omega}} = m_{yy\Omega} - m^2_{zy\Omega}/m_{zz\Omega}$. Число степеней свободы равно $n-I-1$, где $n = \sum_i j_i$.

Оценкой β_i в предположениях Ω является

$$\beta_{i\Omega} = y_{i*}.$$

Поправка регрессии по z равна $\gamma_{\tilde{\Omega}} z_{i*}$. таким образом, оценкой β_i в предположениях $\tilde{\Omega}$ служит

$$\beta_{i\tilde{\Omega}} = y_{i*} - \gamma_{\tilde{\Omega}} z_{i*}.$$

Если проверяется гипотеза $H_\gamma : \gamma = 0$, то для построения доверительного интервала γ необходимо найти дисперсию $\mathrm{D}(\gamma_{\tilde{\Omega}})$. Формулу для нее можно получить из общего выражения ковариационной матрицы $S^2\mathbf{M}_\Omega^{-1}$:

$$\mathrm{D}(\gamma_{\tilde{\Omega}}) = \frac{S^2}{m_{zz\Omega}} - \frac{S^2}{\sum_i \sum_j (z_{ij} - z_{i*})^2},$$

где $S^2 = SS_{\tilde{\Omega}}/(n-I-1)$.

Доверительный интервал для γ дается выражением

$$\gamma_{\tilde{\Omega}} - t\sqrt{\mathrm{D}(\gamma_{\tilde{\Omega}})} \leqslant \gamma \leqslant \gamma_{\tilde{\Omega}} + t\sqrt{\mathrm{D}(\gamma_{\tilde{\Omega}})},$$

где t — статистика t-распределения с $n-I-1$ степенями свободы.

Если проверяется гипотеза H_β, которая состоит в том, что $\beta_1 = \beta_2 = \ldots = \beta_I$, то предположения ω заключаются в том, что $H_\beta \cap \Omega$. В соответствующей задаче дисперсионного анализа

$$SS_\omega = \sum_i \sum_j (y_{ij} - y_{**})^2$$

или

$$SS_\omega = \sum_i \sum_j y_{ij}^2 - n y_{**}^2,$$

где $y_{**} = \frac{1}{n} \sum_i \sum_j y_{ij}$.

В предположении ω аналогично Ω получим

$$m_{yy\omega} = \sum_i \sum_j y_{ij}^2 - n y_{**}^2;$$

$$m_{zz\omega} = \sum_i \sum_j z_{ij}^2 - n z_{**}^2;$$

$$m_{zy\omega} = \sum_i \sum_j z_{ij} y_{ij} - n z_{**} y_{**};$$

$$SS_{\tilde{\omega}} = m_{yy\omega} - m^2_{zy\omega}/m_{zz\omega}.$$

Теперь

$$F = [(n-I-1)/(I-1)]\left[(SS_{\tilde{\omega}} - SS_{\tilde{\Omega}})/SS_{\tilde{\Omega}}\right].$$

Если гипотеза H_β была отвергнута, то можно определить, из-за каких сравнений коэффициентов $\{\beta_i\}$ это произошло. Оценка сравнения

$$\psi_{\tilde{\Omega}} = \sum_i c_i y_{i*} - \gamma_{\tilde{\Omega}} \sum_i c_i z_{i*},$$

что вытекает из условия $\beta_{i\tilde{\Omega}} = y_{i*} - \gamma_{\tilde{\Omega}} z_{i*}$. Соответственно дисперсия оценки

$$D(\psi_{\tilde{\Omega}}) = \sum_i c_i^2 D(y_{i*}) + \left(\sum_i c_i z_{i*}\right)^2 D(\gamma_{\tilde{\Omega}}) =$$

$$= S^2 \left[\sum_i \frac{c_i^2}{J_i} + \frac{\left(\sum_i c_i z_{i*}\right)^2}{\sum_i \sum_j (z_{ij} - z_{i*})^2} \right].$$

Доверительный интервал для ψ определяется неравенством

$$\hat{\psi} - S\sqrt{D(\psi_{\tilde{\Omega}})} \leqslant \psi \leqslant \hat{\psi} + S\sqrt{D(\psi_{\tilde{\Omega}})},$$

где постоянная

$$S = \sqrt{(I-1) F_{\alpha, I-1, n-I-1}}.$$

Приведем примеры, исходные данные для которых заимствованы из работы [83]. Исследуем зависимость скорости распространения упругих колебаний от литологического состава пород и их пористости.

Материалом служат измерения интервального времени продольной волны (акустический каротаж), проведенные в песчаниках и алевролитах разного состава, относящихся к каменноугольным отложениям северо-западной части Днепровско-Донецкой впадины. Одновременно была определена пористость образцов. Всего обработано 226 наблюдений.

Общее уравнение регрессии интервального времени ΔT продольной волны на пористость $K_п$ для объединенной выборки (песчаники и алевролиты, $n = 226$) имеет вид

$$\Delta T = 3{,}36 K_п + 194.8$$

со стандартом $\hat{S} = 21{,}5$.

При разделении песчаников и алевролитов на три группы — глинистые ($J_1 = 69$), малоглинистые ($J_2 = 117$) и карбонатные ($J_3 = 40$) — получены для каждой группы следующие уравнения:

$$\Delta T_1 = 5.32 K_п + 211.2; \quad \Delta T_2 = 3.99 K_п + 181.5; \quad \Delta T_3 = 3.21 K_п + 165.2.$$

Стандартные отклонения соответственно равны:

$$\hat{S}_1 = 24.98; \quad \hat{S}_2 = 13.1; \quad \hat{S}_3 = 7.9.$$

Возникает вопрос, описываются отмеченные связи в песчаниках разного состава одним и тем же уравнением регрессии (прямые регрессии совпадают) или разными, т. е. влияет ли состав на характер связи. Это вопрос существенный, так как если прямые регрессии совпадают, то уравнения можно построить для объединенной выборки песчаников, не дробя их на мелкие литологически разнородные группы.

Для ответа на этот вопрос необходимо сначала проверить гипотезу H_β: $\beta_1 = \beta_2 = \beta_3$ (коэффициенты регрессии γ_i считаются равными) — и в слу-

чае, если она будет отвергнута, определить, из-за каких коэффициентов β_i это произошло (т. е. для каких пород значения β_i больше, а для каких меньше). В данном случае

$$SS_{\omega} = 104\,006{,}25; \quad SS_{\Omega} = 64\,772{,}36,$$

тогда

$$F = \frac{226-3-1}{3-1} \cdot \frac{104\,006{,}25 - 64\,772{,}36}{64\,772{,}36} = 67{,}26.$$

Табличное значение F-критерия меньше: $F_{0{,}1;\,2;\,222} = 2{,}30$. Следовательно, различие в положении прямых регрессии существует, гипотеза H_{β} отвергается. Выяснив, что прямые регрессии не совпадают, определим, какая из них расположена выше, а какие ниже. Для примера возьмем данные по малоглинистым и карбонатным песчаникам. Уравнения регрессии для них характеризуются значениями $\beta_2 = 181{,}5$ и $\beta_3 = 165{,}2$. Судя по ним, прямая регрессии для глинистых песчаников расположена выше, чем для карбонатных. Проверим это предположение. Для этого вычислим сравнение

$$\psi_{\Omega} = \sum c_i \beta_i = (1/2) \cdot 181{,}5 - (1/2) \cdot 165{,}2 = 8{,}15.$$

В данном случае

$$\sum c_i z_{i*} = -0{,}8; \quad \sum\sum (z_{ij} - z_{i*})^2 = 5403{,}21.$$

Дисперсия сравнения

$$D(\psi_{\Omega}) = \frac{64\,772{,}36}{226-3-1}\left(\frac{1/4}{117} + \frac{1/4}{40} + \frac{0{,}64}{5403{,}21}\right) = 2{,}33$$

и, следовательно, $\sqrt{D(\psi_{\Omega})} = 1{,}53$.

Тогда доверительный интервал для ψ при

$$S = \sqrt{(I-1)\,F_{0{,}1;\,2;\,222}} = \sqrt{2 \cdot 2{,}30} = 2{,}15$$

выразится неравенством

$$8{,}15 - 2{,}15 \cdot 1{,}53 \leqslant \psi \leqslant 8{,}15 + 2{,}15 \cdot 1{,}53;$$

$$4{,}86 \leqslant \psi \leqslant 11{,}4.$$

Таким образом, $\beta_2 > \beta_3$.

Отсюда можно заключить, что интерпретацию данных акустического каротажа следует проводить на основе уравнений регрессии, построенных отдельно для карбонатных песчаников и песчаников с преимущественно глинистым и слабоглинистым цементом.

Очевидно, что аналогичный анализ необходимо выполнять и для уравнений регрессии, связывающих пористость с параметром пористости. Хотя известно, что разный литологический состав пород приводит к разным уравнениям регрессии, но в конкретных случаях подробный анализ обычно не проводится. Теоретические уравнения (типа уравнения Арчи) не всегда правомерны, ибо они, не учитывая обязательно присутствующую рассеянность (разброс) вокруг линии регрессии, не дают права делать однозначное заключение: можно ли в конкретной ситуации пользоваться осредненным уравнением регрессии или для каждой литологически отличающейся группы пород следует строить «свое» уравнение.

Итак, мы рассмотрели решение задач, связанных с оценкой коллекторских и литолого-петрографических характеристик горных пород по данным геофизических исследований скважин. На первый взгляд, задача — установление соответствия между геологическими и геофизическими характеристиками — является как бы общеизвестной, она решается достаточно часто. Однако в предлагаемой постановке вскрылись специфические стороны этой задачи, на которые при традиционном ее решении не обращается должного внимания. Так, нами не только показана необходимость учитывать две характерные особенности интерпретируемого материала: большое число одновременно действующих факторов и наличие случайных флуктуаций геологической природы, но и объяснено, как этот учет осуществлять.

Эти обстоятельства потребовали привлечения особых методов, отвечающих указанной специфике. Ими оказались многомерные дисперсионный, регрессионный и ковариационный анализы. Каждый из этих методов пригоден для решения ряда геологических задач, несколько отличных друг от друга с содержательной точки зрения, хотя любую из них можно рассматривать как задачу установления связи между геологическими (коллекторскими, литолого-петрографическими и др.) характеристиками пород и геофизическими данными. Связь в каждой из этих задач понимается по-разному. В одном случае выясняется изменение поинтервальных средних значений одних характеристик при изменении средних значений других; во втором составляется уравнение, описывающее зависимость одной характеристики от другой (или от совокупности других) и определяются коэффициенты такого уравнения; в третьем решается вопрос о сохранении выявленных связей, например, в породах разного состава, и тем самым выясняется, как сильно влияет состав (или ряд других факторов) на геофизические переменные.

Такой подход приводит к правильной, более полной и эффективной интерпретации и, что, пожалуй, самое главное, к самоконтролю. Как показывают рассмотренные примеры, недоучет отмеченной специфики задачи может стать причиной неверного результата. Особенно остро это проявляется при решении вопроса о том, какие из геофизических данных несут информацию об интересующей исследователя геологической характеристике, а какие нет. В рамках традиционных методов вопрос об информативности геофизических параметров, отражающих влияние большого числа геологических факторов, которые в свою очередь осложнены эффектами случайного характера, решен быть не может. В данной книге, по-видимому, впервые указан конкретный путь для решения этого вопроса.

Глава II

ВЫЯВЛЕНИЕ ФАКТОРОВ, ОПРЕДЕЛЯЮЩИХ ИЗМЕНЧИВОСТЬ СОСТАВА И ФИЗИЧЕСКИХ СВОЙСТВ ОСАДОЧНЫХ ПОРОД

ПОСТАНОВКА И СОДЕРЖАНИЕ ЗАДАЧИ

В предыдущей главе рассматривалась оценка по геофизическим данным геологических характеристик горных пород. Эти характеристики, вызывающие изменчивость геофизических данных, были названы факторами. Теперь, не исключая прежнего понимания и смысла термина «фактор», мы вкладываем в него более широкое содержание, в ряде случаев лучше отвечающее существу дела. Так, под фактором могут подразумеваться различные геологические процессы (седиментация, первичные и вторичные преобразования пород и др.), которые в принципе и определяют состав и свойства пород, слагающих разрез. Поэтому при интерпретации геофизических данных мы будем стараться извлекать из них информацию о природных факторах, обусловивших изменчивость любых регистрируемых по скважине геолого-геофизических характеристик.

Очевидно, что такая задача является в какой-то мере новой и своеобразной как по содержанию, так и по результатам. Если в предыдущей главе исследовались факторы, пределы и диапазон изменения которых были известны, то теперь влияющие факторы нам не известны. Заданы, как правило, лишь вариации регистрируемых свойств объекта, вызванные действием искомых факторов. Задача как раз и сводится к выявлению этих факторов. При этом преследуются две основные цели. Во-первых, необходимо найти и установить «причины» (факторы), которыми можно объяснить «следствия» — взаимосвязанное поведение контролируемых геолого-геофизических данных, характеризующих исследуемый объект. Во-вторых, требуется количественно оценить степень влияния каждого из выделенных факторов на изменчивость исходных показателей. Очевидно, что такая постановка задачи представляет совершенно в новом свете интерпретацию геофизических данных.

ОСНОВЫ КОМПОНЕНТНОГО АНАЛИЗА

При поиске метода, решающего данную задачу, необходимо иметь в виду, что здесь приходится анализировать большой комплекс геолого-геофизических показателей (характеристик), сложным образом переплетенных и взаимосвязанных между собой. Очевидно, что связь между этими характеристиками, как правило, не однозначная (функциональная), а стохастическая (вероятностная). В такой ситуации при минимуме имеющейся информации

требуется сначала получить сведения о силе и характере взаимозависимости геологических и геофизических данных, установить смысл этих зависимостей и затем извлечь информацию об источниках (причинах) изменчивости исходных данных. При этом большую систему геологических и геофизических показателей, характеризующих разрез, надо свести к значительно меньшему числу других показателей (факторов). Ответ на все эти вопросы дает метод главных компонент (факторный анализ), целью которого и является экономное описание больших массивов экспериментальных данных, сжатие и свертка исходной информации с последующим ее расщеплением на составные части, поддающиеся более определенной и углубленной интерпретации.

Как и ранее, введем в рассмотрение матрицу наблюдений

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{pmatrix}.$$

Каждый столбец матрицы представляет последовательность измерений того или иного свойства в координатах геологического объекта (например, значения пористости, данных НГМ, БМ и др.). Вектор-строка матрицы, обозначаемая через $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip})$, дает значения измеренных свойств в фиксированной точке объекта. Весь анализ проводится с матрицей X. Здесь уже нет регрессируемых переменных и регрессоров. Все переменные, входящие в матрицу X, являются как бы равноправными.

Представление о факторах вырабатывается на основе изучения взаимосвязанного изменения переменных, образующих матрицу X. Считается, что зависимость между вектор-столбцами (переменными) матрицы X обусловлена тем фактом, что все они представляют собой линейную комбинацию некоторых независимых (ортогональных) переменных $\mathbf{z} = (z_1, z_2, \dots, z_p)$. В этом смысле переменные z_i есть «причины», а исходные переменные x_i — «следствия». Переменные z_i называются главными компонентами или факторами.

Модель j-й главной компоненты можно записать в следующем виде:

$$z_{ji} = w_{j1}x_{i1} + w_{j2}x_{i2} + \dots + w_{jp}x_{ip}$$

или, если ввести в рассмотрение вектор коэффициентов $\mathbf{w}_j = (w_{j1}, w_{j2}, \dots, w_{jp})$, то

$$z_{ji} = \mathbf{w}_j'\mathbf{x}_i.$$

Поскольку каждая главная компонента z_j представляет собой линейную комбинацию исходных переменных $\mathbf{x} = (x_1, x_2, \dots, x_p)$ со своими коэффициентами $\mathbf{w}_j$, то в итоге для описания всех компонент $\mathbf{z} = (z_1, z_2, \dots, z_n)$ будем иметь p векторов $\mathbf{w}_j$, совокуп-

ность которых образует матрицу W. Каждая строка этой матрицы соответствует вектору $\mathbf{w}_j$, относящемуся к данной компоненте z_j.

Для проведения анализа переменные x_i нормируют и центрируют. В дальнейшем будем считать, что каждый элемент матрицы X есть

$$(x_{ij} - \bar{x}_j)/\sigma_{x_j},$$

где

$$\bar{x}_j = \frac{1}{n}\sum_i x_{ij}; \quad \sigma^2_{x_j} = \frac{1}{n-1}\sum (x_{ij} - \bar{x}_j)^2.$$

В соответствии с переменными z введем матрицу Z. Задача, таким образом, заключается в определении элементов матрицы

$$\mathbf{Z} = \begin{pmatrix} z_{11} & z_{12} & \cdots & z_{1p} \\ z_{21} & z_{22} & \cdots & z_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ z_{n1} & z_{n2} & \cdots & z_{np} \end{pmatrix}.$$

Следовательно, математически метод главных компонент сводится к замене системы зависимых переменных x системой независимых (ортогональных) переменных z путем линейного преобразования

$$\mathbf{Z} = \mathbf{X}\mathbf{U}\Lambda^{1/2},$$

где U — матрица, составленная из собственных векторов матрицы парных коэффициентов корреляции $\mathbf{R} = [1/(n-1)]\mathbf{X}'\mathbf{X}$ (элементы матрицы R, как известно, отражают силу линейной связи между любой парой исходных вектор-столбцов матрицы X).

Для этого осуществляется разложение матрицы R по собственным числам и собственным векторам, т. е. матрица R представляется в виде

$$\mathbf{R} = \mathbf{U}\Lambda\mathbf{U}',$$

где Λ — диагональная матрица, элементами которой являются собственные числа λ_i.

Анализу подвергаются сами главные компоненты $\{z_i\}$, которые в той или иной степени присутствуют во всех анализируемых переменных, и, следовательно, их совокупность определяет каждую исходную переменную в отдельности.

Часто главные компоненты называют факторами. Действительно, если исходные переменные контролируют суперпозицию ряда геологических процессов, то каждый фактор в какой-то мере определяет согласованное (взаимосвязанное) поведение контролируемых признаков в координатах исследуемого объекта. В этом случае каждый выделенный фактор может соответствовать определенному геологическому процессу. Однако не всегда фактор имеет генетический смысл. Более того, при геофизических исследованиях тот или иной вскрытый фактор (главная компонента)

может быть следствием используемой системы (комплекса) измерений, когда одно и то же свойство с разной степенью «выраженности» и детальности проявляется в данных различных методов.

Смысл каждого фактора устанавливается на основе анализа факторных нагрузок, определяемых уравнением

$$\mathbf{W} = \mathbf{U}\Lambda^{-1/2}\mathbf{R} = \mathbf{U}\Lambda^{1/2}.$$

Элементы матрицы $\mathbf{W}$ представляют собой коэффициенты корреляции между главными компонентами и исходными переменными. Именно такая простая интерпретация факторных нагрузок позволяет вскрыть смысловое содержание главных компонент (факторов).

Кроме главных компонент и факторных нагрузок немаловажную информацию несут в себе собственные числа, т. е. элементы матрицы $\Lambda = \mathbf{U}'\mathbf{R}\mathbf{U}$. Они отражают меру связи между соответствующей компонентой и совокупностью исходных эмпирических переменных. Доля полной дисперсии переменных, которая может быть объяснена действием соответствующего фактора, называется весом фактора и определяется отношением $\lambda_i \Big/ \sum_{j=1}^{p} \lambda_j$. Факторы выделяются в порядке убывания весов. Это значит, что первый фактор ответствен за формирование наиболее тесных связей между рассматриваемыми эмпирическими переменными в одной группе; второй — за связи в другой группе (это как раз и позволяет применять метод главных компонент при классификации результатов наблюдений).

Иными словами, переменные z_i выделяются так, что последовательно максимизируется их дисперсия $\mathbf{D}(\mathbf{z}) = (\mathbf{XW})^2 = (\mathbf{XW})'\mathbf{XW} = \mathbf{W}'\mathbf{X}'\mathbf{XW} = \mathbf{W}'\mathbf{RW}$, чем и определяется набор коэффициентов $\mathbf{w}_k = (w_{1k}, w_{2k}, \ldots, w_{pk})$. При этом на коэффициенты $\mathbf{w}_k$ накладывается условие $\mathbf{w}_k'\mathbf{w}_k = 1$. Эти коэффициенты и представляют собой нагрузки на компоненты $\{z_i\}$ исходной переменной x_k. Поскольку сумма квадратов нагрузок на все главные компоненты равна единице, то можно определить те компоненты, действием которых в большей степени может быть объяснена изменчивость данной переменной x_k. Ими являются компоненты, для которых сумма $\sum_i w_{ik}^2$ близка к единице.

В связи с тем что главные компоненты в той или иной степени присутствуют во всех анализируемых переменных x, метод главных компонент в принципе можно рассматривать как метод «расщепления» интегральной информации на ряд элементарных составных элементов. Сначала, ориентируясь на максимум изменчивости (дисперсии) исходных геолого-геофизических переменных, из их системы выделяют общность «первого порядка» (первая главная компонента). Затем, исключая эту общность, можно аналогичным путем выделить общность «второго порядка» и т. д.

Важно подчеркнуть, что метод позволяет не только построить главные компоненты, отражающие изменение того или иного фактора в координатах пространства (в частности, по разрезу осадочной толщи), но и оценить влияние каждого фактора или их суммы на формирование изменчивости входящих в анализ эмпирических переменных. Это осуществляется путем следующего преобразования:

$$\mathbf{X} = \mathbf{ZW}, \tag{II.1}$$

т. е. можно восстановить любую эмпирическую переменную через выделенные факторы (главные компоненты).

Факторы, входящие в исходную переменную с незначительным весом (нагрузка при соответствующем факторе тождественно равна нулю), целесообразно вообще не включать при восстановлении переменных. Если исходная переменная будет восстановлена только через значимые факторы, а незначимые (вероятно, здесь случайные) будут отброшены, то такая переменная станет свободной от ряда компонент случайного характера и ее интерпретация, как покажет разбор ряда практических примеров, может быть осуществлена намного полнее и эффективнее.

Таким образом, факторный анализ имеет и то преимущество, что он позволяет от p коррелированных переменных перейти к значительно меньшему числу q некоррелированных (ортогональных) переменных $z_1, z_2, \ldots, z_q$. При этом очевидно, что в общем случае число главных компонент равно числу p исходных признаков. Однако несколько первых главных компонент, как правило, учитывают большую часть общей изменчивости (дисперсии) этих признаков. Первые q членов разложения $z_1, z_2, \ldots, z_q$ извлекают из полной вариации исходных переменных, образующих матрицу $\mathbf{X}$, долю $\varepsilon_q = \sum_{j=1}^{q} \lambda_j \Big/ \sum_{j=1}^{p} \lambda_j$. С возрастанием числа q доля ε_q увеличивается и в случае сильной корреляции исходных переменных быстро приближается к единице. Поэтому для некоторого q, при котором величина ε_q близка к единице, ряд $z_1, z_2, \ldots$, образующий столбцы матрицы $\mathbf{Z}$, с достаточной для практики точностью можно ограничить (оборвать). Очевидно, что слагаемые с малыми весами λ_j очень слабо влияют на значения исходных переменных (признаков).

Возможность ограничиться рассмотрением лишь ряда факторов (главных компонент) и исключить из анализа факторы с достаточно малыми весами λ_j позволяет значительно сократить число исходных переменных и без существенной потери информации об изменчивости признаков рассматривать изменчивость небольшого числа факторов. Практика использования факторной модели при интерпретации комплекса геолого-геофизических данных, характеризующих разрезы осадочных толщ, показывает, что во многих случаях достаточно двух-трех доминирующих ортогональных

компонент (факторов), чтобы система (комплекс) изучаемых переменных была описана на 70—80%.

Так как элементы матрицы W определяются с точностью до поворота в факторном пространстве [87], то в некоторых случаях целесообразно осуществить поворот факторов, что позволит придать им более точное смысловое содержание. Например, иногда выполняют ортогональное вращение факторной матрицы W так, что величина

$$\sum_j \left[q \sum_i (w_{ij}^2/h_i^2)^2 - \left(\sum_i w_{ij}^2/h_i^2 \right)^2 \right]$$

выбирается максимальной (здесь $i = 1, 2, \ldots, p$ — номера переменных; $j = 1, 2, \ldots, q$ — факторы; w_{ij} — нагрузка для i-й переменной j-го фактора; h_i^2 — суммарная нагрузка i-й переменной, определяемая как $h_i^2 = \sum_{j=1}^{q} w_{ij}$.

Однако подобное вращение можно видоизменить, если ориентироваться на смысловое содержание фактора. Так, если требуется придать точное смысловое содержание выделяемому фактору и при этом известно, что один из анализируемых признаков контролирует именно этот фактор, то очевидно, что нагрузка, связывающая указанный фактор с признаком, должна быть максимальной, в то время как нагрузки при этом признаке на другие факторы должны быть равны нулю. Следовательно, если требуется осуществить поворот двух факторов, то необходимо пользоваться формулой

$$W^* = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} W,$$

где W — матрица нагрузок до вращения; W* — после вращения.

Элементы матрицы вращения равны:

$$\cos\varphi = w_{1j}/\sqrt{w_{1j}^2 + w_{2j}^2};$$

$$\sin\varphi = w_{2j}/\sqrt{w_{1j}^2 + w_{2j}^2},$$

где w_{ij} — нагрузка при переменной, на которую ориентируется вращение.

Для трех факторов аналогично имеем

$$W^* = \begin{pmatrix} \cos\psi\cos\varphi & -\sin\varphi & -\sin\psi\cos\varphi \\ \cos\varphi\sin\varphi & \cos\varphi & -\sin\psi\sin\varphi \\ \sin\psi & 0 & \cos\psi \end{pmatrix} W;$$

элементы матрицы вращения задаются формулами

$$\cos\varphi = w_{1j}/\sqrt{w_{1j}^2 + w_{2j}^2}; \quad \sin\varphi = w_{2j}/\sqrt{w_{1j}^2 + w_{2j}^2},$$

$$\cos\psi = \sqrt{w_{1j}^2 + w_{2j}^2}/\sqrt{w_{1j}^2 + w_{2j}^2 + w_{3j}^2};$$

$$\sin\psi = w_{3j}/\sqrt{w_{1j}^2 + w_{2j}^2 + w_{3j}^2}.$$

Итак, метод главных компонент позволяет выделить и оценить факторы, определяющие изменчивость и взаимосвязь геолого-геофизических данных, характеризующих исследуемый геологический объект. Выполняется это путем свертки исходной информации: система, состоящая из большого числа взаимосвязанных показателей, сводится к значительно меньшему числу независимых (ортогональных) компонент (факторов). Очевидно, что меньшее число переменных можно проинтерпретировать намного полнее и эффективнее. В отечественной геологической практике факторный анализ впервые был использован А. Б. Вистелиусом и М. А. Романовой [20]. При истолковании непосредственно геолого-геофизических показателей этот метод применен в ряде работ [12, 36, 46].

ИЗУЧЕНИЕ РАЗРЕЗОВ КАРБОНАТНЫХ ТОЛЩ

При рассмотрении практического применения метода главных компонент основное внимание уделим изучению характеристик, отражающих разрезы карбонатных пород. Как известно, они являются наиболее сложными объектами и характеризующие их геолого-геофизические данные трудно поддаются анализу, многие традиционные методы интерпретации здесь оказываются малоэффективными. Нами отобраны геолого-геофизические показатели, образующие один и тот же комплекс, но описывающие разрезы карбонатных пород, которые формировались в разных геологических условиях. Исследованные разрезы отличаются друг от друга условиями накопления, литологическим составом, возрастом, характером насыщения и свойствами насыщающего флюида и т. п.

Изложение результатов целесообразно начать с разрезов позднемелового возраста Северо-Восточного Предкавказья и Южного Дагестана. Эти разрезы наиболее полно освещены как керновым материалом, тщательно обработанным лабораторными и литолого-петрографическими методами, так и данными достаточно обширного комплекса геофизических измерений.

Северо-Восточное Предкавказье и Южный Дагестан

Скв. 30-Карабулак. Исходным материалом для использования компонентной модели послужили лабораторные и литолого-петрографические определения, проведенные по керну и в шлифах.

Были оценены следующие показатели (параметры): 1) пористость открытая $K_п$, определенная по керну; 2) пористость вторичная блоковая $K_{п.бл}$ (пустоты в органических остатках, цементе), определенная по большим шлифам; 3) пористость вторичная общая $K_{п.общ}$ (т. е. блоковая плюс обусловленная пустотами расширения, развитыми по трещинам), определенная по большим шлифам; 4) плотность эффективных трещин $T_{эф}$, определенная по большим шлифам, 5) содержание $CaCO_3$ в породе (карбонатность), определенное по керну лабораторным методом; 6) общее количество органических остатков, определенное по шлифам; 7) степень эпигенетической перекристаллизации, определенная по шлифам.

Эти данные были дополнены геофизическими показателями, измеренными в тех точках разреза, к которым приурочены исследованные образцы керна. Комплекс геофизических измерений состоит из данных: 1) радиометрии (НГМ); 2) электрометрии (БЭЗ при двух промывочных жидкостях с сопротивлением 1,3 и 3,2 Ом·м, ПС).

Сравнительно полная освещенность карбонатных отложений набором геолого-геофизических показателей позволяет решить две важные задачи.

Первая задача — выявление ведущих геологических факторов (причин), определяющих формирование различных типов пористости, и оценка роли каждого из этих факторов. Решение этой задачи позволяет также выяснить, каким образом те или иные свойства (параметры) пород взаимосвязаны между собой и каково их значение в образовании коллекторов. Заметим, что вопросы, касающиеся роли отдельных факторов в формировании коллекторских свойств пород позднемелового возраста Восточного Предкавказья, затрагиваются уже давно [11, 21, 23, 76]. На некоторые из этих вопросов получены однозначные ответы [76]; точки зрения на другие (например, на роль трещин в формировании коллекторов [9, 76]) все больше сближаются; третьи (например, влияние примеси глинистого материала в известняках на их емкостные свойства) остаются дискуссионными.

Вторая задача — выявление геологических факторов, определяющих изменчивость регистрируемых по скважине геофизических показателей, измеренных различными геофизическими методами. При этом особое внимание обращено на то, как влияют различные типы пористости карбонатных пород на геофизические показатели и насколько интенсивность проявления емкостных характеристик больше (или меньше) интенсивности других геологических эффектов.

Визуальные наблюдения, изучение пород в керне по многим скважинам показали, что позднемеловые отложения Восточного Предкавказья представлены однообразной толщей светлых, в разной степени (чаще незначительной) глинистых тонкослоистых известняков. Это полностью согласуется и с наблюдениями, про-

веденными непосредственно в скв. 30-Карабулак. Так, карбонатность пород равна 95,3% при стандартном отклонении 5,62%, т. е. диапазон изменения карбонатности (глинистости) крайне незначителен и, следовательно, литологический состав пород однообразен.

Структура известняков зависит от содержания в них органических остатков, которое колеблется от 5 до 70—80%. Для образцов исследуемой скважины содержание органических остатков составляет 58,5% при стандартном отклонении 21,5%. Органические остатки представлены в основном одно-камерными (сферы) и многокамерными (фораминиферы) образованиями, детритом и обломками макрофауны. Породообразующим и цементирующим веществом является мелко- и тонкозернистый кальцит.

Известняки сравнительно мало изменены постседиментационными процессами. Наиболее заметны перекристаллизация и выщелачивание. Перекристаллизация, как показывает анализ шлифов, проходила преимущественно по органическим остаткам, реже по цементу (в основном при диагенезе и незначительно при эпигенезе). По мнению Л. П. Гмид и С. Ш. Леви [23], такая избирательность перекристаллизации объясняется более свободным движением поровых вод в относительно более пористой среде, какой явились агрегаты, образованные органическими остатками. В целом степень диагенетической перекристаллизации различных типов пород, как показали исследования многочисленных шлифов [11, 21], изменяется от 12 до 45%; эпигенетической — от 5 до 17%. Наиболее перекристаллизованными являются органогенно-обломочные разности, но их доля в разрезе весьма незначительна. Для всего разреза верхнего мела скв. 30-Карабулак среднее значение степени эпигенетической перекристаллизации равно 3,9% при стандартном отклонении 4,3%.

Находящиеся в тектонически активном районе породы позднемелового возраста были подвержены интенсивной трещиноватости. Для них характерна высокая плотность трещин, как залеченных (преимущественно кальцитом), так и эффективных. Например, среднее значение эффективной трещиноватости исследуемых отложений скв. 30-Карабулак составляет 75 $м^{-1}$ при почти таком же стандартном отклонении — 58 $м^{-1}$, т. е. в некоторых интервалах разреза плотность трещин может достигать 150 $м^{-1}$.

Трещиноватость, как известно, является важнейшим свойством рассматриваемых пород, она определяет их высокую проницаемость. Типичные кавернозные разности пород в верхнемеловом разрезе не имеют широкого распространения. Обычно вторичные пустоты не превышают по своим размерам крупных пор (0,1—0,5 мм). Пустоты выщелачивания (как свидетельствует анализ шлифов) приурочены к органическим остаткам. Они также наблюдаются в цементе, связаны с открытыми и с частично или полностью залеченными трещинами. Наиболее значительные

пустоты (до 1 мм) отмечаются в органогенно-обломочных разностях известняков.

Поры, развитые в ограниченных трещинами блоках, обусловливают вторичную блоковую пористость (0,5—1,0%). В рассматриваемой скважине средняя пористость этого типа равна 0,93% при стандартном отклонении 0,83%, следовательно, лишь на редких участках разреза блоковая пористость может превосходить 1,8%. Пустоты расширения в трещинах определяют пористость того же порядка, что и каверны в блоках: от 0,5 до 2%. В скв. 30-Карабулак пустоты расширения открытых трещин обусловливают в среднем по разрезу пористость 1,05%. Вторичная пористость, образованная блоковой и трещинной, составляет 2,05% при стандартном отклонении 1,5%.

Пористость $K_п$ известняков, определяемая по керну в лаборатории, как будет показано ниже, отражает все виды пористости, за исключением, быть может, собственно трещинной. Средние значения открытой пористости по многим скважинам равны 3—5% и редко превосходят 7—10%. Для исследуемой скважины средняя пористость равна 3,63% при стандартном отклонении 1,58%, т. е. диапазон изменения открытой пористости во всем интервале исследований крайне мал.

Для выделения геологических факторов, определяющих формирование пористости (см. формулировку первой задачи), с помощью факторного (компонентного) анализа по исходным показателям была получена матрица нагрузок, первые три столбца которой представлены в табл. II.1. Матрица дана после вращения факторов, в приведенных координатах элементы матрицы приобретают наиболее отчетливое смысловое содержание. Истолкование матрицы показывает, что система, состоящая из семи геологических показателей, эффективно может быть заменена и описана тремя факторами. Их сумма извлекает из исходных показателей более 70% информации.

Первый фактор (вес 30%) свидетельствует о том, что определенная доля пористости обнаруживает тенденцию к уменьшению по мере увеличения количества органических остатков, повышения степени эпигенетической перекристаллизации пород и в какой-то мере увеличения карбонатности. Эта часть пористости не является вторичной, поскольку названные свойства пород не оказывают воздействия на вторичную блоковую $K_{п.бл}$ и вторичную общую $K_{п.общ}$ пористость — нагрузки при этих показателях равны нулю. Следовательно, есть основание полагать, что это сказывается седиментационная пористость. Она определяется первичными межзерновыми пустотами, которые своим происхождением обязаны процессам седиментации.

Анализ шлифов показывает, что органические остатки по сравнению с цементирующей массой кальцита подверглись большей перекристаллизации как в стадии диагенеза, так и позднее. Перекристаллизация привела к частичному сокращению объема

Таблица II.1

Результаты компонентного анализа (скв. 30-Карабулак) (первая задача)

Переменная	Факторы			Среднее значение переменной	Стандартное отклонение
	Первый	Второй	Третий		
Степень эпигенетической перекристаллизации	0,86	0,00	0,17	8,9	4,30
Содержание органических остатков	0,84	0,00	0,07	58,5	21,48
Карбонатность	0,28	—0,60	0,54	95,3	5,62
Пористость:					
в блоках (по шлифам) $K_{п.\,бл}$	0,00	0,75	—0,29	0,93	0,83
в блоках плюс трещинная $K_{п.\,общ}$	0,00	0,74	0,40	2,05	1,05
по лабораторным данным $K_п$	—0,44	0,42	0,63	3,63	1,58
Плотность эффективных трещин $T_{эф}$	0,65	—0,58	0,15	75,52	57,97
Вес факторов, %	30,0	28,3	13,4	—	—

первичных (седиментационных) пор в результате роста кристаллов цементирующего кальцита, по главным образом в результате перекристаллизации самих остатков. Это обусловило обратную зависимость между степенью эпигенетической перекристаллизации и открытой пористостью.

Органические остатки и перекристаллизация определили в конечном итоге неоднородность структуры карбонатных пород. Это, по всей видимости, положительно повлияло на образование эффективной трещиноватости, о чем свидетельствуют большие однозначные нагрузки при названных показателях в первом факторе. Увеличению трещиноватости могло способствовать также и некоторое возрастание карбонатности известняков. Но это обстоятельство наиболее отчетливо проявилось во втором факторе.

Следовательно, истолкование нагрузок первого фактора свидетельствует о том, что этот фактор отражает в основном явления седиментационного характера и позволяет установить, каким образом структурные первичные особенности пород влияют на образование первичной пористости. Переработка структуры пород под воздействием процессов эпигенетической перекристаллизации может частично сократить межзерновую пористость, но увеличения вторичной блоковой пористости в этом случае может и не происходить.

Второй фактор (вес 28,3%) указывает на то, что уменьшение содержания $CaCO_3$ в породе влечет за собой увеличение всех типов пористости. Колебание карбонатности в известняках позднемелового возраста связано преимущественно с изменением содержания в них глинистой (терригенной) примеси, что,

как известно, может отражать обстановку осадконакопления. Следовательно, наличие прямой связи между глинистостью и пористостью $K_п$ свидетельствует о том, что выделяемая и контролируемая вторым фактором часть пористости тоже является первичной межзерновой пористостью седиментационного происхождения.

Может показаться странным, что два самых значительных фактора связываются с процессами седиментации. Но, как известно, геологические, и в частности седиментационные, процессы — это сложные явления, обусловленные многими причинами (физико-географической обстановкой, изменением рельефа дна и т. п.), что и оттеняется выявленными факторами.

На первый взгляд кажется противоречивой обратная зависимость между содержанием $CaCO_3$ в породе и вторичной пористостью, обусловленной выщелачиванием, а также общей пористостью. Незначительное увеличение глинистости в данном случае не приводит к торможению процессов выщелачивания, а наоборот, вызвав, как мы убедились ранее, увеличение седиментационной пористости, в конце концов создает, вероятно, благоприятную обстановку для активной фильтрации агрессивных растворов. Следовательно, в чистых разностях известняков фильтрационный фактор и агрессивность растворов при образовании вторичных пустот выступают на первый план. При этом очевидно, что чем выше первичная (межзерновая) пористость пород, тем вероятнее ожидать увеличения вторичной пористости блоков.

Анализ нагрузок во втором факторе свидетельствует также о том, что увеличение карбонатности известняков положительно влияет на трещиноватость. Объясняется это явление повышенной хрупкостью чистых известняков, которые больше подвержены трещиноватости, чем их глинистые (наиболее пластичные) разности.

Таким образом, второй фактор позволяет выявить влияние вещественного состава пород на их первичную и вторичную пористость, в том числе и на трещиноватость. При этом изменения первичной (межзерновой) пористости и плотности трещин непосредственно зависят от содержания $CaCO_3$ в породе, в то время как образование вторичной пористости связано с карбонатностью, вероятно, лишь косвенно. По мере увеличения седиментационной пористости улучшаются фильтрационные способности пород и, следовательно, создается благоприятная обстановка для процессов выщелачивания.

Третий фактор (вес 13,4%) отчетливо отражает двойственное влияние содержания $CaCO_3$ на разные типы пористости. Прямая связь с суммарной вторичной пористостью указывает на приуроченность части пустот к минеральным (выполненным кальцитом) и к открытым трещинам, что подтверждается и наблюдением в шлифах. Интенсивная в трещинах фильтрация, как правило, резко замедляется в пористой среде, где может осаждаться и, следовательно, запечатывать первичные и вторичные пустоты

растворенный и вынесенный из трещин $CaCO_3$. Зависимость между открытой (определенной по керну) и суммарной вторичной пористостью, прослеживающаяся в двух последних факторах, свидетельствует о том, что открытая пористость отражает все отмеченные типы пустот, т. е. и первичные и вторичные. Таким образом, третий фактор, по всей видимости, контролирует процессы выщелачивания.

Кратко просуммируем выводы, полученные при интерпретации матрицы нагрузок.

Емкость известняков позднемелового возраста складывается из первичных межзерновых и вторичных пустот, развитых в блоках и приуроченных к трещинам. Открытая пористость, определенная лабораторным путем, включает все виды пористости (за исключением, быть может, трещинной) и является суммарной. Доля вторичной пористости невелика, о чем свидетельствует анализ средних значений изучаемых типов пористости. Лишь в отдельных случаях она может достигать 1,5—2%. Основную долю пористости составляет первичная седиментационного происхождения. В тех интервалах, где она больше, отмечается увеличение и каверновой (вторичной) пористости. Перекристаллизация сокращает первичную пористость (но не очень существенно) и вместе с тем практически не влияет на образование вторичных пустот. Трещиноватость тесно связана со структурой и вещественным составом пород.

Для решения второй сформулированной выше задачи достаточно подвергнуть анализу лишь те данные, которые приведены в табл. II.2, так как в данной ситуации на геофизические показатели оказывают влияние лишь различные типы пористости пород и содержание глинистого материала. В общем случае на геофизические показатели действует большее число геологических особенностей (табл. II.3).

По сути дела, табл. II.3 есть не что иное, как гипотетическая факторная матрица (точнее, модель характера взаимосвязей), если соблюдается условие, что указанные в таблице геологические особенности не коррелированы между собой. Но в реальной ситуации это условие, как явствует из только что рассмотренного примера, не выполняется.

Можно заранее предположить, что если существует корреляция внутри системы геологических особенностей, то она непременно оставит своеобразный отпечаток на значении элементов фактической матрицы нагрузок (факторной матрицы). Это приведет к тому, что в реальном случае выделяемые из системы геофизических показателей факторы (главные компоненты) не будут точно отражать изменение межзерновой пористости или изменение глинистости и т. п. Тем не менее при сравнении гипотетической матрицы с реальной матрицей нагрузок по совпадению характера взаимосвязей, дополненному сведениями об их силе (тесноте), можно определить, какой именно фактор выделяется из анализируемой

Таблица II.2

Результаты компонентного анализа (скв. 30-Карабулак) (вторая задача)

Переменная	Факторы			Среднее значение переменной	Стандартное отклонение
	Первый	Второй	Третий		
Пористость:					
по лабораторным данным $K_п$	0,71	0,00	—0,05	3,62	1,54
в блоках плюс трещинная $K_{п.\,общ}$	0,21	0,49	0,48	2,24	2,37
Карбонатность	—0,18	—0,27	0,63	94,56	5,02
$J_{НГМ}$	—0,98	—0,13	—0,05	1,74	0,11
$\Delta U_{ПС}$	—0,23	—0,68	0,13	0,24	6,56
$\Delta\rho_{КС}$	—0,65	0,27	—0,40	8,57	21,72
$\ln T_{эф}$	—0,27	0,57	0,35	4,01	0,60
$\ln \rho_8$	—0,87	0,37	0,05	3,96	0,57
$\ln \rho_4$	—0,82	—0,03	0,14	3,74	0,60
$\ln \rho_2$ $\rho_0 = 1{,}3$ Ом·м	—0,85	—0,28	0,13	2,96	0,33
$\ln \rho_1$	—0,70	—0,24	0,18	2,17	0,22
$\ln \rho_{0,5}$	—0,92	—0,29	0,05	3,90	0,69
$\ln \rho_4$ $\rho_0 = 3{,}2$ Ом·м	—0,85	—0,30	0,18	4,02	0,64
$\ln \rho_2$	—0,90	—0,25	0,06	3,96	0,54
Вес факторов, %	51,7	12,6	5,4	—	—

Таблица II.3

Взаимосвязи между геологическими характеристиками и геофизическими данными

Геофизические данные	Геологические признаки (факторы)					
	Пористость			Глинистость	Карбонатность	Содержание доломита
	межзерновая	каверновая	трещинная			
$\rho_{КС}$, БМ, МБК, БЭЗ, ИМ	—/0	—/0	—/0	—/—	+/+	+/+
$\Delta U_{ПС}$	—/0	—/0	—/0	+/+	0/0	0/0
$J_{НГМ}$	—/—	—/—	—/—	—/—	+/+	+/+
$J_{ГМ}$	0/0	0/0	0/0	+/+	—/—	—/—
$\Delta t_{АМ}$	+/+	+/+	+/+	+/+	—/—	—/—
$\Delta\rho_{КС}$	+/+	+/+	+/+	0/0	0/0	0/0

Примечание. Знак «плюс» означает прямую связь, «минус» — обратную, «нуль» — неопределенную или ее отсутствие; в числителе — данные для воды, в знаменателе — для нефти.

системы геофизических показателей, выяснить, какой характер связи в реальном случае присущ неопределенной ситуации. Так как характер связей в табл. II-3 определен для разных флюидов, то сравнение этой таблицы с реальной матрицей может в какой-то степени помочь установить тип заполняющего флюида в радиусе действия зондов.

Заметим, что прежде чем получить матрицу нагрузок, приведенную в табл. II.2, показатели методов электрометрии были прологарифмированы. Мы исходили из известного предположения, что пористость, содержание нерастворимого остатка (карбонатность) и результаты НГМ имеют линейную связь с данными электрометрии, выраженными в полулогарифмическом масштабе (этого предположения мы придерживались и в последующих примерах).

Перейдем к анализу матрицы нагрузок (см. табл. II.2); отметим, что элементы нагрузок даны также после вращения факторов.

Первый фактор, объясняющий более 50% общей дисперсии (изменчивости) признаков, можно истолковать как показатель, контролирующий изменение объема жидкости, которая заполняет все сообщающиеся пустоты пород. Эти пустоты образуют в основном ту пористость, которая определяется по керну (лабораторным методом). Следовательно, можно считать, что этот фактор отражает в основном межзерновую пористость пород $K_п$. Такой вывод следует из анализа значений и знаков нагрузок при показателях пористости $K_п$, электро- и радиометрии. Так, пористость отрицательно связана с данными этих методов. Последнее обстоятельство и сравнение фактической матрицы нагрузок с гипотетической (см. табл. II.3) позволяет заключить, что жидкость, заполняющая пустоты прискважинной зоны в радиусе действия зондов (БЭЗ, НГМ), является водой. Точнее, это фильтрат промывочной жидкости, так как результаты опробования скважины свидетельствуют, что нефтепроявления в верхнемеловом комплексе отмечаются по всему стволу.

Совместный анализ средних значений геофизических показателей и нагрузок с соответствующими знаками в первом факторе убеждает в том, что повторная смена промывочной жидкости с минерализацией, выше первоначальной, изменяя в целом электрическое сопротивление горных пород, изменяет и силу (тесноту) связи между открытой пористостью и кажущимся удельным электрическим сопротивлением. Изменение при этом сопротивления самих пород показывает, что они пропитываются повторным фильтратом промывочной жидкости. Отметим, что зона, затрагиваемая электроизмерительными зондами с большими размерами, претерпевает меньшие изменения.

Таким образом, процесс пропитки блоков породы происходит интенсивно при насыщении их как нефтью, так и фильтратом промывочной жидкости. Этот факт хорошо согласуется с результатами экспериментов и с теоретическими расчетами многих отечественных и зарубежных исследователей, изучавших механизм

противоточной капиллярной пропитки водой гидрофильных пористых блоков трещиноватых пород [4, 94]. Установлено, что меньшая проницаемость нефтенасыщенных блоков пород (капилляры малого радиуса) влечет за собой более эффективную их пропитку. Подобные опыты с целью изучения процесса обводнения карбонатных коллекторов при эксплуатации месторождений проводились и непосредственно на образцах Карабулак-Ачалукского месторождения [4].

По всей вероятности, процесс пропитки идет следующим образом. Из скважины фильтрат поступает в трещины, а затем через стенки трещин — в блоки пород. В конечном итоге первоначальный флюид, заполнявший пустотное пространство, либо смешивается с фильтратом промывочной жидкости (если это была минерализованная вода или фильтрат жидкости с другой минерализацией), либо оттесняется за зону проникновения нового фильтрата (в случае насыщения блоков нефтью). Таким образом, следует считать, что не только трещины, но и пустоты блоков пород (сумма пор межзернового и кавернового типа) заполнены либо фильтратом, либо смесью. Кроме того, тесная отрицательная линейная связь пористости блоков с логарифмом удельного сопротивления показывает, что, вероятно, еще в процессе бурения начинает формироваться зона проникновения фильтрата промывочной жидкости, причем ширина этой зоны превышает радиус действия электрометрического зонда с $AO = 8$ м.

Мы уже отмечали, что связь между пористостью $K_п$ и показаниями электрометрии тем теснее, чем выше сопротивление промывочной жидкости (т. е. чем ниже ее минерализация). Так, теснота связи пористости с данными электрометрии (зонд с $AO = 2$ м) при сопротивлении промывочной жидкости $\rho_0 = 1{,}3$ Ом·м равна —0,85, в то время как при $\rho_0 = 3{,}2$ Ом·м ее значение составляет —0,9, т. е. при насыщении трещин слабоминерализованным раствором усиливается влияние межзерновых и каверновых пустот. При высокой минерализации возрастает роль трещин, и поэтому линейная связь между рассматриваемыми характеристиками ослабевает. Заметим, что аналогичный вывод — возрастание роли трещин в изменении удельного электрического сопротивления при насыщении породы высокоминерализованным раствором — вытекает из многих теоретических исследований [79].

Второй фактор (вес 12,6%) отражает влияние пористости вторичного происхождения (высокие нагрузки при переменных $K_{п.общ}$ и $T_{эф}$). Эта пористость также вызывает изменение геофизических показателей, но связь ее с последними (за исключением ПС) не столь тесна. По нагрузкам видно, что роль трещин несколько выше, когда они заполнены высокоминерализованным раствором. Однако нельзя забывать, что соленый раствор, заполняя трещины, обязательно проникает и в некоторые притрещинные пустоты (в частности, в поры межзернового типа). Формирование притрещинной зоны непременно обусловит фиктив-

ную и несколько большую раскрытость трещин. Таким образом, наблюдаемая связь во втором факторе обязана, вероятно, не только густоте трещин, но и их раскрытости, притом фиктивно увеличенной.

Возникновение притрещинной зоны и, как следствие, появление фиктивной раскрытости трещин — факт весьма существенный; его особенно важно учитывать при количественной оценке трещинной пористости на основе данных метода «двух растворов» [62, 63, 70], в противном случае можно получить весьма завышенные значения трещинной пористости.

Третий фактор (вес 5,4%) отражает влияние карбонатности на интенсивность трещиноватости, но в геофизических характеристиках этот фактор не дает ощутимого эффекта. Лишь разность удельных электрических сопротивлений $\Delta\rho$, измеренных при разных сопротивлениях промывочной жидкости, свидетельствует о том, что чем выше трещиноватость, тем ниже значения этой разности. Можно сделать следующее, весьма вероятное, предположение: чем выше трещиноватость, тем интенсивнее происходит пропитка блоков пород и тем меньше при прочих равных условиях разность между сопротивлениями одних и тех же пород, заполненных флюидом разной минерализации (разного удельного электрического сопротивления).

Таким образом, система (комплекс) из 15 геологических и геофизических показателей с помощью компонентной модели может быть заменена тремя факторами (главными компонентами), которые обеспечивают около 70% общей изменчивости анализируемых показателей. Первый, доминирующий, фактор показывает, что изменчивость геофизических характеристик (методы БЭЗ и НГМ) обусловлена изменением объема фильтрата промывочной жидкости, заполняющей блоковые пустоты (межзерновые поры и часть каверн). Исключив влияние первого фактора, мы пришли к выводу, что второй фактор контролирует изменение каверновой пористости и эффективной трещиноватости. Этот фактор обусловливает значительно меньшую (по сравнению с первым) долю общей изменчивости геофизических показателей (методы БЭЗ, НГМ, ПС). Отметим, что чем выше пористость указанного типа, тем ниже значения геофизических характеристик. Третий фактор, хотя и связан с трещиноватостью, но отражения в геофизических данных почти не находит. Следовательно, собственно трещиноватость дает очень слабый эффект, и поэтому относительное ее изменение по разрезу скважины геофизическими показателями фиксируется неоднозначно.

Анализ средних значений и стандартных отклонений электрометрических показателей (наряду с нагрузками) позволяет сделать вывод, что при бурении в коллекторе, первоначально заполненном нефтью, практически мгновенно формируется зона проникновения фильтрата промывочной жидкости, чему способствуют процессы капиллярной пропитки. Изменив минерализацию промывочной жидкости (уменьшив ее почти в 3 раза), мы заметили

эффект смешения двух фильтратов. Но процесс их смешения и проникновения в пласт, по-видимому, идет не столь быстро. Однако новый фильтрат все же заполняет трещины, каверны, развитые по трещинам, а также притрещинные пустоты. Это обстоятельство заставляет критически относиться к оценке трещинной пористости методом «двух растворов» (в публикациях уже отмечался факт завышения трещинной пористости, рассчитанной по указанному методу [59, 76]).

Скв. 154-Карабулак-Ачалуки. Пробуренная в контуре нефтяной залежи скважина вскрыла карбонатные отложения верхнего мела в интервале 2204—2459 м. Каждая фиксированная точка разреза была охарактеризована данными сравнительно большого комплекса геофизических исследований. Анализу были подвергнуты диаграммы (предварительно квантованные, шаг 0,5 м нормальной мощности разреза) БЭЗ, БМ (при двух промывочных жидкостях с сопротивлением 0,22 и 1,38 Ом·м), НГМ, НГМ-т, АМ и ИМ. Представительный вынос керна (60—70%) позволил равномерно описать интервал 2236—2458 м лабораторными определениями межзерновой пористости $K_п$ и содержания нерастворимого остатка $C_{н.о}$. Не приводя детального геологического описания данного разреза, поскольку он практически такой же, как и по скв. 30-Карабулак, отметим только крайнее однообразие литологического состава пород. Среднее значение содержания нерастворимого остатка по разрезу составляет 4,94% при стандартном отклонении 4,29%.

Изменение коллекторских свойств карбонатного разреза обусловлено значением и типом пористости. Как уже отмечалось, для рассматриваемого разреза характерен смешанный тип пористости: в разных сочетаниях присутствуют межзерновые, трещинные и каверновые пустоты. Среднее значение межзерновой пористости равно 5,66% при стандартном отклонении 2,57%, т. е. колебание пористости по разрезу имеет сравнительно малый диапазон. Трещинная и каверновая пористость прямыми методами в данном случае не оценивалась, но судя по результатам исследований других скважин наиболее интенсивные зоны трещиноватости приурочены к маастрихтскому и кампанскому ярусам.

В соответствии с компонентной моделью была получена матрица нагрузок (табл. II.4), откуда видно, что на долю трех выделенных факторов приходится более 70% полной дисперсии (изменчивости) анализируемой системы показателей. Если сравнить элементы этой матрицы и матрицы, полученной для скв. 30-Карабулак, то можно сделать вывод об их практической идентичности, откуда следует и аналогичная интерпретация.

Первый фактор (вес 54,3%) контролирует изменение объема минерализованного раствора (фильтрата промывочной жидкости) в породе и, как следствие, открытую пористость $K_п$. Для этой скважины более отчетливо проявляется небольшое колебание нагрузок при переменных электрометрии (БЭЗ и БМ).

Таблица II.4

Результаты компонентного анализа (скв. 154-Карабулак-Ачалуки)

Переменная	Факторы			Среднее значение переменной	Стандартное отклонение
	Первый	Второй	Третий		
Пористость $K_п$	0,65	0,00	0,02	5,66	2,57
Содержание нерастворимого остатка $C_{н.о}$	0,06	0,50	0,24	4,94	4,29
$J_{НГМ}$	—0,77	—0,47	0,15	1,18	0,19
$J_{НГМ-т}$	—0,78	—0,39	0,14	2,95	0,39
$\Delta t_{АМ}$	0,60	0,30	0,18	5,51	0,07
$\ln \rho_8$ ($\rho_0 = 1{,}38$ Ом·м)	—0,87	—0,40	—0,03	3,28	0,74
$\ln \rho_4$ ($\rho_0 = 1{,}38$ Ом·м)	—0,81	—0,30	—0,06	3,67	0,69
$\ln \rho_2$ ($\rho_0 = 1{,}38$ Ом·м)	—0,82	—0,11	—0,10	3,61	0,79
$\ln \rho_1$ ($\rho_0 = 1{,}38$ Ом·м)	—0,80	—0,40	—0,22	3,30	0,51
$\ln \rho_{0.4}$ ($\rho_0 = 1{,}38$ Ом·м)	—0,70	—0,35	—0,22	2,36	0,33
$\ln \rho_{БМ}$ ($\rho_0 = 1{,}38$ Ом·м)	—0,78	—0,37	—0,27	3,01	0,69
$\ln \rho_2$ ($\rho_0 = 0{,}22$ Ом·м)	—0,79	—0,36	0,06	3,13	0,69
$\ln \rho_4$ ($\rho_0 = 0{,}22$ Ом·м)	—0,88	—0,42	0,02	3,39	0,78
$\ln \rho_1$ ($\rho_0 = 0{,}22$ Ом·м)	—0,87	—0,26	—0,29	2,67	0,39
$\ln \rho_{0.4}$ ($\rho_0 = 0{,}22$ Ом·м)	—0,82	—0,21	—0,38	1,34	0,26
$\ln \rho_{БМ}$ ($\rho_0 = 0{,}22$ Ом·м)	—0,82	—0,32	—0,40	2,97	0,82
$\ln \rho_{ПМ}$ ($\rho_0 = 0{,}22$ Ом·м)	—0,45	—0,23	—0,41	5,26	0,03
$\Delta\rho$	—0,34	0,07	0,53	1,50	12,99
Вес факторов, %	54,3	10,8	6,3	—	—

Это происходит, вероятно, либо из-за влияния границы раздела двух различных по физическим свойствам пластов, либо из-за изменения в радиальном направлении электрических свойств флюида. Эти обстоятельства могут по-разному сказаться при измерении кажущегося удельного сопротивления зондами разных размеров.

Необходимо обратить внимание на тот факт (см. также анализ предыдущей матрицы), что в первом факторе нагрузка при переменной $K_п$ несколько ниже, чем при геофизических переменных (показаниях различных зондов БЭЗ, НГМ и НГМ-т). Эффект «раскорреляции» (уменьшение тесноты линейной связи) возникает не только из-за неудовлетворительной привязки результатов анализа керна (значений пористости) к геофизическим данным. Возможной причиной может быть и непредставительный вынос керна из интервалов высокопористых пород.

Но вероятнее всего, что геофизические показатели способны отражать не только ту часть пустотного пространства, которая контролируется лабораторными определениями по керну, но и другие пустоты. Во-первых, нельзя пренебрегать влиянием на геофизические показатели таких пустот, которые из-за малых

размеров испытуемого образца не могут быть зарегистрированы при определении межзерновой пористости. Эти пустоты, без сомнения, можно отнести к каверновому типу. Во-вторых, гидродинамически изолированные поры могут быть взаимосвязаны электрически благодаря сообщающимся тончайшим капиллярам, заполненным молекулярно-связанной водой. Эти пустоты в значительной степени воздействуют на показания как электрометрических, так и радиоактивных методов. Данные выводы свидетельствуют о том, что первый фактор кроме межзерновой пористости отражает и каверновую.

Если последнее утверждение справедливо, то высокие нагрузки при переменных различных зондов БЭЗ в какой-то мере могут свидетельствовать, что форма связи результатов электрометрии с пористостью, состоящей из суммы межзерновых и каверновых пустот, близка к форме уравнения Арчи. Иными словами, присутствие каверновых пустот, электрически взаимосвязанных, не искажает известной зависимости Арчи, как это иногда полагают (дискуссия по этому вопросу была поднята в работе И. И. Горюнова [26]). Кстати, этот же вывод был получен в результате прямых экспериментальных исследований на образцах карбонатных пород, проведенных Т. С. Изотовой [40].

Предположение, что каверновые пустоты воздействуют на электрометрические показатели так же, как и межзерновые поры, может быть подтверждено и другим образом. Характер изменения сопротивления, как известно, зависит не только от значения пористости, но и от структуры норового пространства. Одной из характеристик этой структуры является отношение размера отдельной пустотности (поры, каверны) к полному объему породы, затронутой исследованием. Так, при крайне малых разносах электродов в зонде (микрозонд) вследствие того, что радиус его действия (объем исследования) относительно мал, некоторые достаточно крупные по размерам пустоты (каверны) могут оказывать такое влияние на сопротивление, которое не согласуется с формой уравнения Арчи.

Очевидно, что по мере увеличения разносов, а следовательно, и объема изучаемой среды эти же поры будут играть роль пустот межзернового типа. Элементарные статистические расчеты показывают, что сообщающиеся пустоты объемом 1 дм^3, случайно разбросанные в 1 м^3 породы (примерно такой объем породы затрагивается 4-метровым градиент-зондом) и создающие пористость 5%, вызывают такое же изменение удельного сопротивления, как и поры межзернового типа при том же значении пористости.

Таким образом, та пористость, которая определяет изменчивость и взаимосвязь геофизических показателей и выделяется первым фактором, по-видимому, складывается из следующих элементов: 1) пористости межзернового типа, объединяющей как гидродинамически, так и электрически связанные пустоты; 2) пористости кавернового типа, которая имеет тенденцию к увеличению

по мере роста межзерновой пористости. Каверны, определяющие пористость, вероятно, развиты и внутри блоков, и по трещинам.

Второй фактор (вес 10.8%) в большей степени связан с содержанием нерастворимого остатка (сравнительно высокая нагрузка при этой переменной). Не исключено, что этот фактор контролирует также и изменение объема каверновых пустот. Действительно, при анализе геологических данных по скв. 30-Карабулак было показано, что повышение содержания терригенного материала (глинистости) положительно сказывается на увеличении межзерновой пористости. Этот процесс в свою очередь создает благоприятные условия для усиленной фильтрации агрессивных минерализованных вод, которые способствуют образованию пустот выщелачивания. Та незначительная глинистость, которая фиксируется в верхнемеловом разрезе, очевидно, не может существенно изменить межзерновую пористость и, следовательно, ослабить или усилить процессы выщелачивания. Таким образом, изменение каверновой пористости (наряду с изменением содержания нерастворимого остатка), по-видимому, и вызывает дополнительную вариацию (дисперсию) геофизических показателей, чем и определяется тот факт, что они контролируются вторым фактором.

Выделяемые из исходной системы переменных последующие факторы крайне слабы и практически равноценны. Отражают они, очевидно, случайные и разные по природе геологические причины, произвольным образом проявляющиеся по разрезу.

Третий фактор, хотя его вес и мал (6.3%), все же заслуживает определенного внимания. Как видно из анализа нагрузки на $\Delta\rho$, этот фактор как будто бы характеризует изменение трещинной пористости. Значительная отрицательная нагрузка при показателе ИМ лишний раз свидетельствует о справедливости этого предположения (как известно, данные ИМ весьма чувствительны к изменению трещинной пористости).

Теперь целесообразно остановиться на геолого-генетической трактовке выделенных факторов.

Выше было показано, что пористость пород, определяемая по керну лабораторным методом, отражает пористость различных типов (за исключением собственно трещинной) и формируется в основном под влиянием процессов седиментации. Так как первый фактор контролирует изменение именно данной пористости, то его можно определить как седиментационный. Содержание нерастворимого остатка в карбонатных породах может являться, как известно, характеристикой их фациальной изменчивости. Поэтому второй фактор обоснованно трактовать как фациальный. Эти факторы, выделяемые из системы геолого-геофизических показателей, ортогональны друг другу, и, следовательно, соответствующие им процессы могли протекать независимо.

Чтобы наглядно продемонстрировать развитие седиментационного и фациального факторов во времени (по разрезу), следует построить соответствующие кривые главных компонент (рис. II.1),

здесь же показано изменение и третьего (тектонического) фактора, отражающего поведение трещиноватости. Имея три фактора с известным геологическим смыслом, которые объясняют изменчивость комплекса геолого-геофизических переменных более чем на 70%, мы можем по ним восстановить каждую исходную переменную. Для этого следует использовать матричное соотношение (II.1). Такая процедура позволяет исключить из исходных переменных влияние необъяснимых (случайных) факторов. Кроме того,

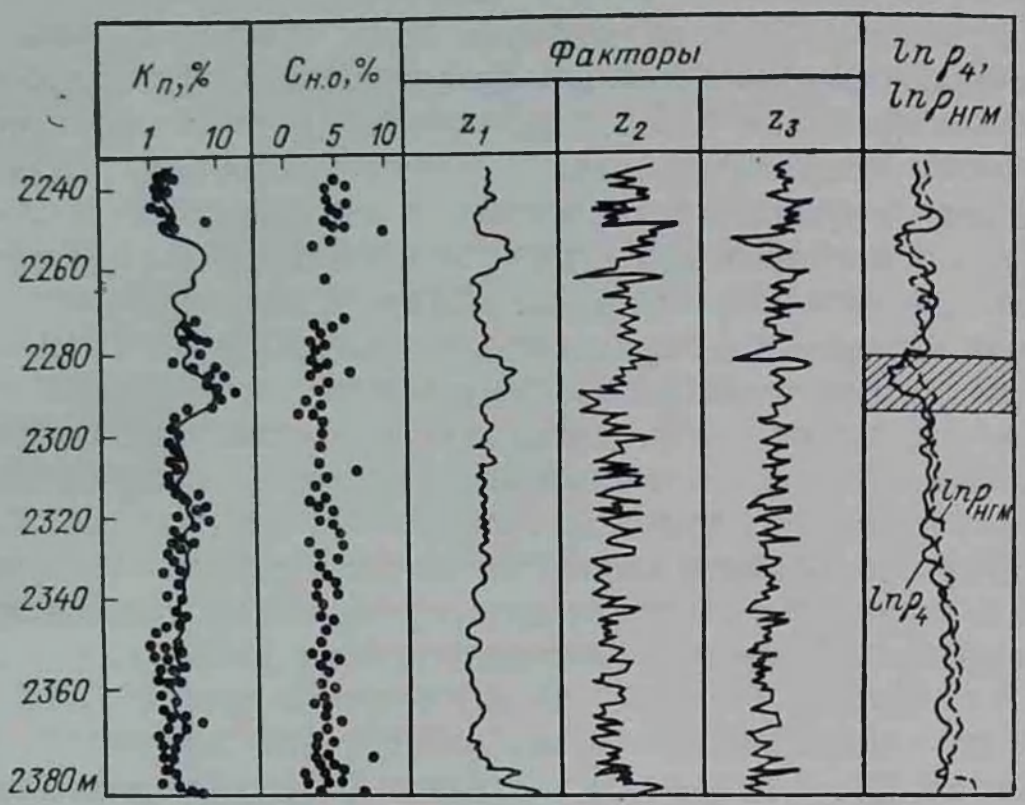


Рис. II.1. Геолого-геофизические переменные и выделенные из них ортогональные компоненты (факторы) в скв. 154-Карабулак-Ачалуки.

Точками показаны фактические значения переменных, кривыми — восстановленные по факторам; заштрихована зона интенсивного развития вторичной (трещинной) пористости.

тот или иной конкретный фактор входит в восстанавливаемую кривую с определенным весом (соответствующей нагрузкой). Это обстоятельство очень важно, ибо геофизические показатели разных методов по-разному отражают влияние одного и того же фактора. Так, из табл. II.4 видно, что при восстановлении кривой НГМ и кривой 4-метрового градиент-зонда ($\rho_0 = 0{,}22$ Ом·м) седиментационный фактор войдет примерно с одинаковым весом (одинаковыми нагрузками), то же самое можно сказать и о фациальном факторе; фактор же трещиноватости войдет в указанные переменные с нагрузками, разными по знаку и абсолютному значению. Это в известной мере справедливо, так как теоретически доказано [25, 27, 61], что удельное электрическое сопротивление чувствительнее реагирует на изменение трещинной пористости, чем данные НГМ.

Сравнивая восстановленные значения переменных НГМ и КС (эти значения нормированы, т. е. изменение кривых разных методов происходит в одном и том же масштабе), по их расхождению можно установить и проследить изменение типа пористости в исследуемом разрезе.

1. Если в интервале наблюдения нормированные и восстановленные по факторам переменные совпадают, то этот интервал характеризуется межзерновым (либо межзерновым и каверновым) типом пористости.

2. Если значения нормированной переменной НГМ в некотором интервале наблюдения больше, чем значения нормированной переменной КС, то этот интервал характеризуется повышенной (превышающей фоновую) трещиноватостью.

3. Если значения нормированной переменной КС, наоборот, превосходят нормированные значения НГМ, то возможны две ситуации: либо интервал наблюдения характеризуется пустотами, которые электрически изолированы друг от друга, либо в этом интервале часть пустот (разного типа) заполнена флюидом высокого сопротивления.

Зона повышенной трещиноватости, выделенная указанным способом (см. рис. II.1), приурочена к тем же стратиграфическим диапазонам, где аналогичные зоны установлены по данным прямых исследований (по методу больших шлифов).

Скв. 33-Заманкул. Находящаяся за контуром нефтеносности скважина также вскрыла разрез верхнемеловых карбонатных

Таблица II.5

Результаты компонентного анализа (скв. 33-Заманкул)

Переменная	Факторы		Среднее значение переменной	Стандартное отклонение
	Первый	Второй		
Пористость $K_п$	0,65	0,00	4,24	2,31
Пористость по шлифам $K_{шл}$	0,21	0,04	0,45	1,22
Проницаемость трещинная по шлифам	—0,05	0,67	4,99	8,53
$T_{эф}$	—0,26	0,75	24,57	29,18
T_{min}	0,36	0,08	31,87	80,94
Карбонатность	0,11	—0,50	94,41	4,12
$J_{НГМ}$	—0,48	—0,73	1,63	0,09
$J_{ГМ}$	—0,37	0,69	7,94	2,92
$\Delta U_{ПС}$	0,91	—0,05	0,44	9,96
$\ln \rho_8$	—0,71	—0,44	1,92	0,86
$\ln \rho_4$	—0,80	—0,46	2,26	0,76
$\ln \rho_2$	—0,66	—0,64	2,72	0,60
$\ln \rho_1$	—0,67	—0,69	3,03	0,53
$\ln \rho_{0,4}$	—0,71	—0,60	3,12	0,41
Вес факторов, %	31,4	28,0	—	—

($\rho_0 = 3{,}3$ Ом·м)

отложений (интервал 2159—2468 м). По керну проведены такие же геологические исследования, как и в скв. 30-Карабулак. Комплекс же геофизических исследований был сильно сокращен, но тем не менее результаты анализа геолого-геофизических данных при помощи компонентной модели (табл. II.5) свидетельствуют о том,

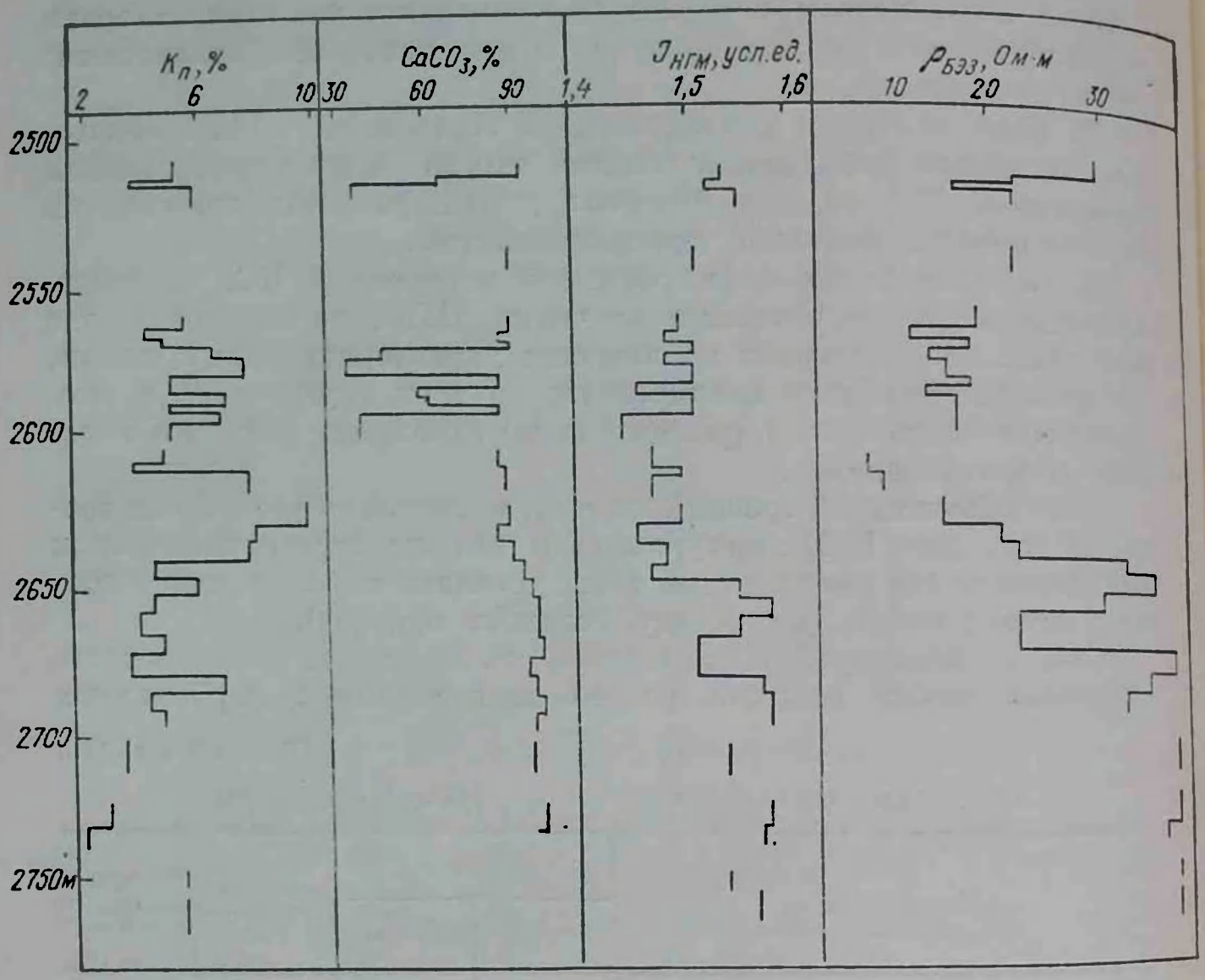
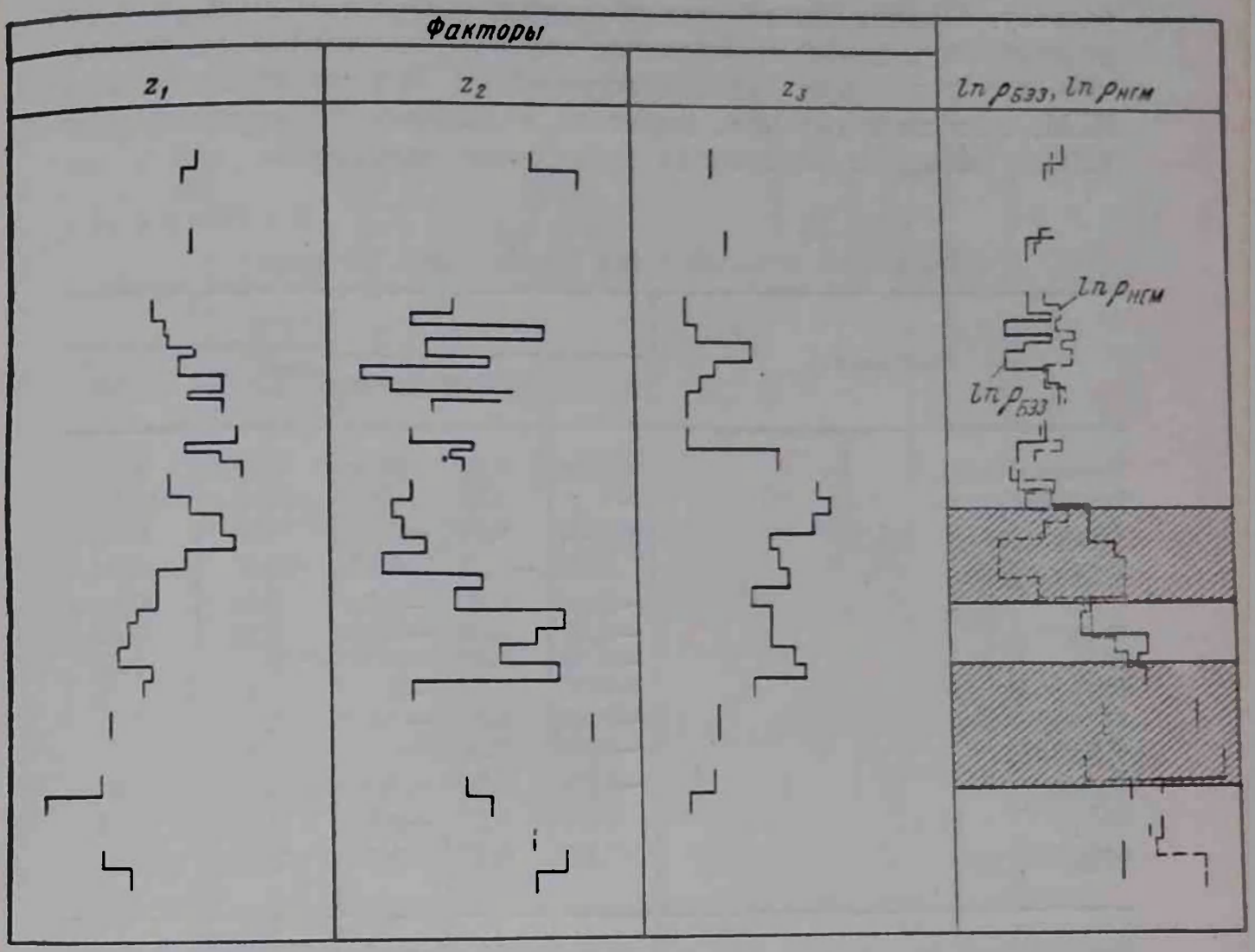


Рис. II.2. Геолого-геофизические переменные и выделенные
В первых четырех колонках даны исходные значения характеристик, в последней — восста
щинной)

что выделенные первые два фактора (значения нагрузок даны после поворота факторов) имеют аналогичный геологический смысл.

Скв. 15-Ялама. В разрезе скважина вскрыла породы верхнего мела на глубине 2116—2761 м. В отличие от ранее рассмотренных этот разрез находится в несколько иных фациальных условиях, что определило соответствующий литологический состав пород. Практически чистые разности известняков, развитые в Чечено-Ингушетии, здесь замещаются более глинистыми. Средняя карбонатность разреза составляет 83,39% при стандартном отклонении 14,9% (в предыдущих разрезах эти характеристики приближались соответственно к 95 и 5%).

Более резкая изменчивость пород по содержанию глинистого вещества позволила проследить его влияние как на коллекторские свойства пород (межзерновую пористость), так и на геофизические показатели. Сразу отметим, что по характеру изменения межзерновой пористости (рис. II.2) рассматриваемый разрез аналогичен



из них ортогональные компоненты (факторы) в скв. 15-Ялама.
новленные по факторам; заштрихованы зоны интенсивного развития вторичной (тре-пористости.

предыдущим. Среднее значение пористости, определенной по керну, составляет 5,79% при сравнительно низком стандартном отклонении, равном 1,47%.

Комплекс геофизических методов, проведенный в этой скважине, включает методы электрометрии (БЭЗ, МПЗ, МГЗ, ПС) и радиометрии (НГМ, ГМ). Дополнительно вычислен показатель приращения $\Delta\rho$ между данными МГЗ и МПЗ. По расхождению в показаниях разных зондов, как известно, можно качественно судить о зонах трещиноватости.

Вся совокупность геологических и геофизических показателей была проанализирована на основе компонентной модели. Интер-

претация матрицы нагрузок (табл. II.6) и трактовка факторов (после соответствующего их поворота) в общих чертах такая же, как и в ранее разобранных примерах. Первый фактор отражает объем жидкости, заполняющей пустоты кавернового и межзернового типа, второй связан с карбонатностью и каверновыми пустотами, третий отражает, вероятно, изменение трещиноватости по разрезу. Причем интенсивность трещиноватости в этом разрезе, по-видимому, довольно высокая, так как отмечаются значительные нагрузки при электрометрических переменных. Данные НГМ, судя по нагрузкам, вероятно, осложнены влиянием относительно большого количества глинистого материала.

Таблица II.6

Результаты компонентного анализа (скв. 15-Ялама)

Переменная	Факторы			Среднее значение переменной	Стандартное отклонение
	Первый	Второй	Третий		
Пористость $K_п$	0,82	0,00	0,00	5,79	1,47
Карбонатность	—0,31	0,45	0,16	83,39	14,9
$J_{НГМ}$	—0,54	0,50	—0,49	1,53	0,04
$J_{ГМ}$	0,12	—0,58	0,63	10,24	0,60
$\ln \rho_8$	—0,50	0,46	—0,50	2,08	0,52
$\ln \rho_4$	—0,50	0,46	—0,50	2,28	0,60
$\ln \rho_2$	—0,51	0,47	—0,51	2,53	0,56
$\ln \rho_1$	—0,57	0,41	—0,61	2,52	0,42
$\ln \rho_{0,5}$	—0,71	0,33	—0,53	1,90	0,30
$\ln \rho_{МПЗ}$	—0,39	0,08	—0,80	1,52	0,63
$\ln \rho_{МГЗ}$	—0,30	0,05	—0,90	1,09	0,92
$\Delta U_{ПС}$	0,16	0,73	—0,21	0,69	2,97
$\Delta \rho_{МГЗ-МПЗ}$	0,00	0,15	0,82	2,19	4,7
Вес факторов, %	22,5	17,5	32,6	—	—

Выделив факторы известного геологического смысла, мы и в этом примере попытались выявить в разрезе зоны повышенной трещиноватости способом восстановления по факторам исходных переменных НГМ и КС. По расхождению восстановленных кривых были намечены две такие зоны.

Прибалтика

Карбонатные отложения ордовика, вскрытые скв. 38-Дурбе и скв. 3-Ремте, хорошо охарактеризованы равномерно отобранным керном (иногда чаще чем через 1 м). Это обстоятельство позволило детально осветить разрезы кривыми изменения межзерновой пористости $K_п$ и плотности γ (рис. II.3). В скв. 38-Дурбе, кроме того, была проведена количественная оценка трещинной

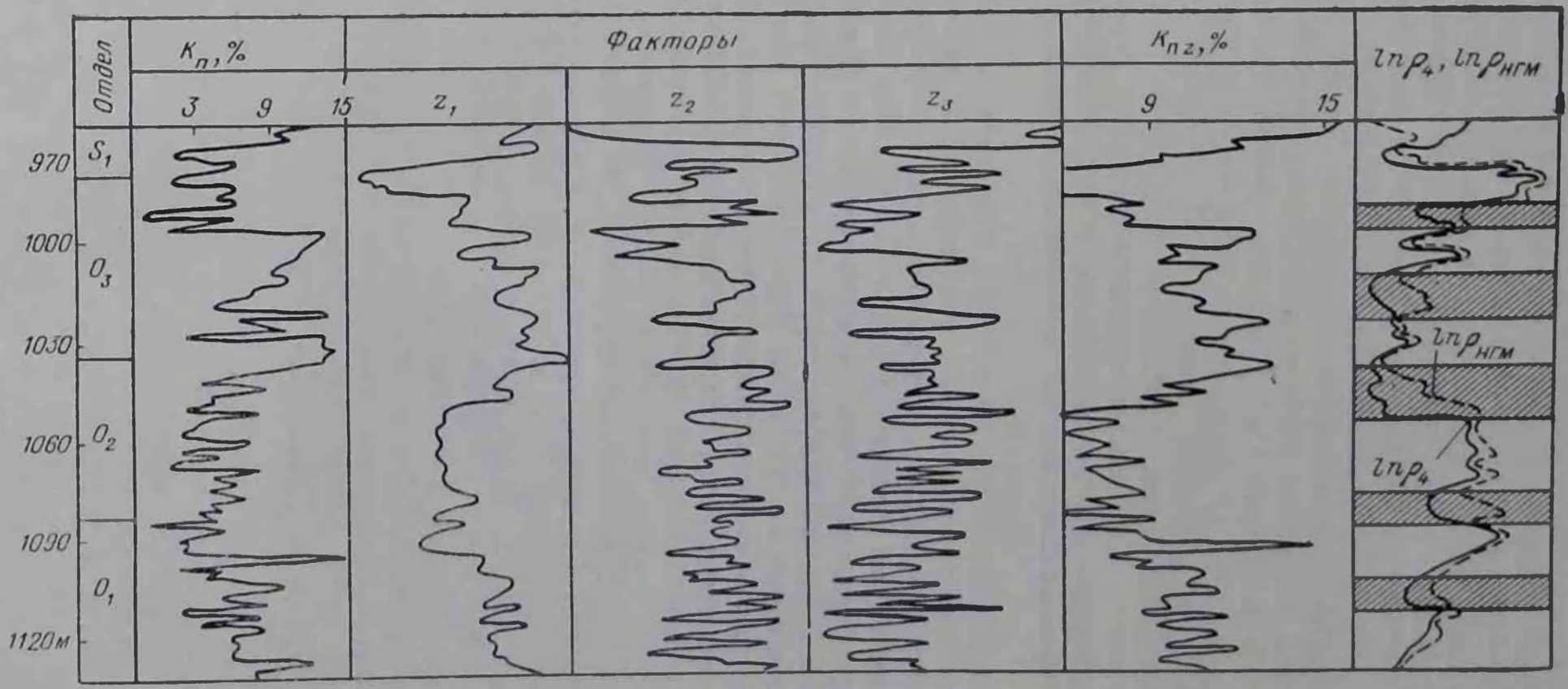


Рис. II.3. Геолого-геофизические переменные и выделенные из них ортогональные компоненты (факторы) в скв. 38-Дурбе.

Заштрихованы зоны интенсивного развития вторичной (трещинной) пористости.

пористости по шлифам, что позволило наметить зоны с повышенными значениями коллекторских свойств.

Скв. 38-Дурбе. Расположенная в прогибе, разделяющем Дурбенскую и Лепайскую брахиантиклинальные складки, скважина вскрыла ордовикские и силурийские отложения. Силурийская глинисто-мергельная толща расценивается как экранирующая для пород-коллекторов ордовика, в частности для горизонта органогенно-обломочных известняков. Этот горизонт (985—995 м) характеризуется незначительной межзерновой пористостью (в среднем 5%), причем вторичная межзерновая и каверновая пористость достигает 3—4% (в среднем 1%). Плотность трещин сравнительно невелика (20—30 $м^{-1}$), но раскрытость их несколько увеличена, что определяет относительно повышенные расчетные значения трещинной проницаемости; трещинная пористость составляет сотые доли процента. Тип коллектора этого горизонта можно определить как каверново-порово-трещинный. Известняки отличаются пятнистым нефтенасыщением по вторичным порам, кавернам и трещинам.

Нижележащие глинистые известняки характеризуются более высокими значениями открытой пористости (до 15%), интенсивной трещиноватостью (до 100 $м^{-1}$); трещинная пористость в них достигает 0,1%, трещинная проницаемость — 2,5 мД. Еще ниже залегает пачка, представленная чередованием глинистых известняков, глин и мергелей. Породы слаботрещиноваты, трещины преимущественно горизонтальные. Эта пачка рассматривается как локальная покрышка для органогенно-обломочных глинистых известняков среднего ордовика. Коллекторские свойства последних характеризуются довольно низкими значениями параметров: межзерновая пористость 4—6%, вторичная 1—2%, трещинная проницаемость 0,5—1 мД; трещинная пористость в редких случаях превосходит 0,05%. По порам, трещинам и кавернам отмечается пятнистое нефтенасыщение. Тип коллектора можно определить как порово-трещинный. Подстилаются нефтенасыщенные породы пачкой сильноглинистых алевритистых известняков, межзерновая пористость которых достигает 11%.

Отложения нижнего ордовика представлены в основном мергелями и глинами с прослоями комковатых известняков. В этом горизонте отмечается тенденция к увеличению межзерновой пористости с глубиной. Трещинная пористость несколько повышенная и достигает 0.1%.

Характеризуя разрез ордовика в целом, можно отметить, что распределение литологического состава пород по глубине подчинено полициклической зависимости. При этом глинистые разности отличаются относительно более высокими значениями межзерновой пористости и трещиноватости по сравнению с чистыми разностями известняков. Кроме того, с глубиной уменьшается содержание нерастворимого остатка в породе (от 49,8 до 18,3%) и окиси магния (7.18 до 0,83%).

Коллекторские характеристики, определенные по керну, дополнены геофизическими показателями (БЭЗ, ГМ, НГМ). Для выяснения связей между геологическими данными и геофизическими показателями целесообразно обратиться к компонентной модели (табл. II.7). Она позволяет установить, какие именно геологические факторы определяют изменчивость геофизических показателей. Комплекс, состоящий из 10 исходных показателей, можно заменить тремя факторами (главными компонентами), сумма которых объясняет общую изменчивость системы взаимосвязанных показателей более чем на 75%.

Таблица II.7

Результаты компонентного анализа (скв. 38-Дурбе)

Переменная	Факторы			Среднее значение переменной	Стандартное отклонение
	Первый	Второй	Третий		
Пористость $K_п$	0,60	—0,44	0,06	9,61	3,96
Пористость $K_{п.\,тр}$	—0,20	0,06	—0,67	0,03	0,07
Плотность γ	—0,46	0,66	—0,45	2,50	0,02
$J_{НГМ}$	—0,78	—0,01	0,35	$7{,}59 \cdot 10^3$	$1{,}18 \cdot 10^3$
$J_{ГМ}$	0,77	—0,27	0,24	$7{,}42 \cdot 10^3$	$1{,}40 \cdot 10^3$
$\ln \rho_4$	—0,79	—0,25	—0,10	1,93	0,62
$\ln \rho_2$	—0,90	—0,13	0,01	2,06	0,65
$\ln \rho_1$	—0,92	—0,17	0,05	2,40	0,70
$\ln \rho_{0,5}$	—0,91	—0,12	0,10	2,39	0,69
$\ln \rho_{0,25}$	—0,87	0,06	0,09	1,99	0,63
Вес факторов, %	57,2	8,6	8,9	—	—

Первый фактор (вес 57,2%) определяет основную долю изменчивости входящих в анализ переменных. Этот фактор контролирует изменение объема жидкости, заполняющей пустоты пород, и, как следствие, межзерновую пористость. Нагрузки и знаки при них, а также средние удельные электрические сопротивления свидетельствуют о том, что пустоты пород насыщены фильтратом промывочной жидкости (такой характер насыщения отмечался и в разрезах Северо-Восточного Предкавказья).

В отличие от ранее разобранных примеров обращает на себя внимание достаточно жесткая связь пористости (блоковой) с показаниями ГМ. Если идти от известного факта, что данные ГМ отражают изменение глинистости по разрезу, то в данной ситуации ее увеличение тесно связано с повышением пористости. По результатам химического анализа керна [65] содержание глинистой фракции может достигать 50%.

Установлено, что увеличение содержания глинистого материала в породе отрицательно сказывается на фильтрационных

способностях коллектора с межзерновым типом пористости. Не случайно поэтому в ордовикских отложениях Прибалтики отмечается ухудшение фильтрационных свойств пород по мере увеличения их межзерновой пористости [41]. Однако это не значит, что сильноглинистые известняки не могут быть хорошими коллекторами, способными отдавать флюид; это свойство, очевидно, определяется наличием или отсутствием трещин.

Второй фактор (вес 8,6%) также определяет изменение по разрезу части межзерновой пористости, не зависящей, однако, от той пористости, которая контролируется первым фактором. Одинаковые знаки при нагрузках переменных $\ln \rho$ и $K_п$ указывают на то, что межзерновые поры, отражаемые вторым фактором, либо заполнены флюидом высокого сопротивления (свечение керна под люминесцентной лампой отмечается по всему разрезу), либо не существовали до экстрагирования и сушки образцов, т. е. образовались искусственно.

Третий фактор (вес 8,9%) контролирует изменение объема трещинных пустот и, вероятно, каверн, развитых по трещинам. В заметной мере он определяет изменчивость плотности пород и радиометрических показаний (НГМ и ГМ). Сопротивление пород при высоких значениях межзерновой пористости почти не зависит от трещиноватости, что согласуется с теоретическими расчетами (конечно, при условии равенства сопротивлений флюидов, заполняющих блоки пород и межблоковое пространство).

Попытка выделить зоны трещиноватости по расхождению восстановленных по факторам исходных показателей (электро- и радиометрии) отображена на рис. II.3. Как видно, определенные расхождения между кривыми имеются, что свидетельствует о наличии зон с трещиноватостью, способной вызвать аномальное

Таблица II.8

Результаты компонентного анализа (скв. 3-Ремте)

Переменная	Факторы		Среднее значение переменной	Стандартное отклонение
	Первый	Второй		
Пористость $K_п$	0,82	—0,46	11,97	6,18
Плотность γ	—0,75	0,53	2,39	0,19
$J_{ГМ}$	0,79	—0,36	$6{,}01 \cdot 10^3$	$2{,}75 \cdot 10^3$
$\ln \rho_4$	—0,85	—0,33	1,92	0,70
$\ln \rho_2$	—0,76	—0,38	2,22	0,82
$\ln \rho_1$	—0,88	—0,32	2,26	0,83
$\ln \rho_{0,5}$	—0,92	—0,08	2,41	0,78
$\ln \rho_{0,3}$	—0,92	—0,14	2,27	0,67
Вес факторов, %	70,2	12,5	—	—

изменение в электрометрических показаниях. Некоторым из этих зон присуща относительно большая трещинная емкость, что установлено по методу больших шлифов.

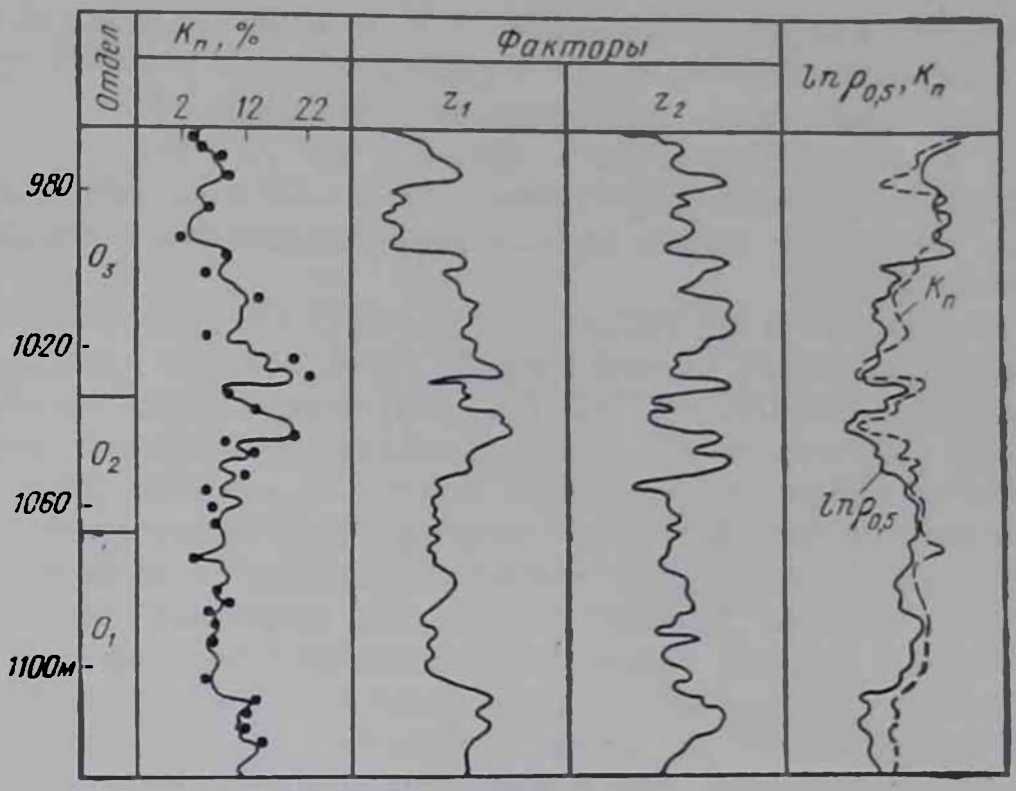


Рис. II.4. Геолого-геофизические переменные и выделенные из них ортогональные компоненты (факторы) в скв. 3-Ремте. Точками показаны фактические значения $K_п$, кривыми — значения переменных, восстановленные по факторам (для $K_п$ дано зеркальное изображение).

Скв. 3-Ремте. Вскрытый ею геологический разрез аналогичен пройденному скв. 38-Дурбе. По характеру поведения главных компонент в разрезе (табл. II.8, рис. II.4) легко выявляются особенности его строения.

Тимано-Печорская область

Объектом исследования служило одно из месторождений, на котором подавляющая часть карбонатных коллекторов обладает низкими значениями межзерновой пористости (до 6%) и в отличие от ранее анализируемых коллекторов содержит газоконденсат. Продуктивная толща сложена преимущественно известняками, значительно меньше распространены доломиты и мергели (последние только в нижней перми).

Данные петрографических исследований [23] показывают, что каменноугольные и пермские отложения были подвергнуты вторичным изменениям под воздействием процессов доломитизации, сульфатизации, окремнения и т. п. Доломитизация может способствовать образованию эффективных пустот, хотя иногда увеличения пористости не отмечается. Сульфатизация и окремнение часто

уменьшают эффективную пористость пород. По результатам петрографических исследований, проведенных Л. П. Гмид, каменноугольные отложения характеризуются низким содержанием нерастворимого остатка (1,5—2%), в то время как в пермских отложениях оно может достигать 35% и более (глинистые известняки, мергели). Во всех типах пород интенсивно развиты трещины и стилолитовые швы, плотность их достигает 300 $м^{-1}$. Относительно высокая раскрытость трещин (до 70 мкм) способствует фильтрации флюидов (проницаемость 15—20 мД). Все это показывает, что среди пород залежи преобладает порово-трещинный тип коллекторов.

Известно, что в тех интервалах исследуемого разреза, где межзерновая пористость низкая (до 6%), истолкование данных геофизических методов (НГМ, БЭЗ, АМ, БМ) даже с целью оценки межзерновой пористости существенно осложнено. Может казаться, что это обстоятельство связано именно с низкими значениями пористости: ее воздействие на геофизические характеристики становится настолько слабым, что влиянием других свойств затушевывается. Однако сравнивая средние значения межзерновой пористости и средние содержания глинистого материала в карбонатных породах данного месторождения и Карабулак-Ачалуки, мы не видим явного различия. Привлечем метод главных компонент, чтобы вскрыть, какие причины и каким образом определяют изменчивость геофизических показателей. Матрицы нагрузок были получены по 9 скважинам отмеченной площади, но мы обсудим результаты по скв. 28 и 33. В скв. 33 интервал исследования (3372—3442 м) относится к продуктивной зоне (газоконденсат), а в скв. 28 (3437—3542 м) — к водоносной.

Скв. 33-Вуктыл. Комплексом лабораторных исследований определены по керну межзерновая пористость $K_п$, плотность γ, межзерновая проницаемость, содержание остаточной воды, отнесенное к объему пор (пересчитано затем к объему породы). Каждая точка разреза, в которой был отобран керн, охарактеризована геофизическими данными методов БЭЗ, БМ, НГМ, ГМ. Параметры трещиноватости, оцененные по методу больших шлифов, не включены в факторный анализ (табл. II.9), так как плотность трещин и их раскрытость оказались постоянными по всему разрезу.

Первый фактор (вес 25,7%) можно определить как фактор межзерновой пористости. То обстоятельство, что в электрометрических данных он не получает должного отражения, свидетельствует о заполнении пустот породы флюидом высокого сопротивления. Кроме того, можно отметить следующие интересные частные зависимости. При увеличении количества остаточной воды, отнесенного к объему пор, ухудшаются фильтрационные способности породы. Аналогичные выводы содержатся в работах других исследователей, например А. А. Ханина [86]. Данные НГМ контролируются содержанием остаточной воды, отнесенной ко всему

Таблица II.9

Результаты компонентного анализа (скв. 33-Вуктыл)

Переменная	Факторы		Среднее значение переменной	Стандартное отклонение
	Первый	Второй		
Пористость $K_п$	0,97	0,00	6,01	4,36
Плотность γ	—0,88	—0,09	2,64	0,121
Содержание остаточной воды:				
в единичном объеме пор	—0,50	0,25	39,99	19,26
в единичном объеме породы	0,50	0,20	2,05	1,66
Проницаемость	0,81	—0,22	—1,381	3,585
$J_{НГМ}$	0,52	0,15	3 291	700.5
$J_{ГМ}$	—0,30	0,27	3 867	1 310
$\ln \rho_{БМ}$	—0,23	—0,61	4,61	1,31
$\ln \rho_{0.4}$	0,05	—0,43	2,78	0,32
$\ln \rho_1$	0,09	—0,80	4,12	0,30
$\ln \rho_2$	0,19	—0,94	5,34	0,39
$\ln \rho_4$	0,13	—0,95	6.01	0,64
$\ln \rho_8$	0,00	—0,83	4,23	0,70
Вес факторов, %	25,7	30,1	—	—

объему породы. По-видимому, этим параметром определяется и изменчивость данных БМ.

Интерпретация первого фактора указывает на отсутствие проникновения промывочной жидкости в пустоты блоков пород. По-видимому, наличие газа в коллекторе препятствует внедрению фильтрата. Это обстоятельство и порождает непривычные связи пористости с геофизическими показателями БЭЗ, НГМ, ГМ. Поэтому оценивать параметры коллекторов по геофизическим данным в подобных разрезах нецелесообразно.

Второй фактор (вес 30,1%) отражает изменчивость электрометрических показателей. Но геологическую причину, которая ее вызывает, установить в данном случае невозможно, так как влияние остаточной воды на эти показатели настолько мало, что им нельзя объяснить наблюдаемую изменчивость. Вполне возможно, что сказываются колебания глинистости карбонатных пород.

Скв. 28-Вуктыл. Интерпретация нагрузок, полученных при изучении геолого-геофизических данных, характеризующих интервал 3437—3542 м, не требует специальных пояснений (табл. II.10). Указанный интервал насыщен водой, что определило обычные формы взаимосвязей. Правда, теснота взаимосвязи межзерновой пористости $K_п$ с данными электрометрии и особенно НГМ очень слабая. Вероятно, это объясняется низкими значениями $K_п$.

Таблица II.10

Результаты компонентного анализа (скв. 28-Вуктыл)

Переменная	Факторы		Среднее значение переменной	Стандартное отклонение
	Первый	Второй		
Пористость $K_п$	0,67	0,00	1,74	1,89
$J_{ГМ}$	0,24	—0,10	$28{,}18 \cdot 10^2$	$14{,}07 \cdot 10^2$
$J_{НГМ}$	—0,09	—0,26	$84{,}75 \cdot 10^2$	$12{,}36 \cdot 10^2$
$\ln \rho_{7,5}$	—0,57	—0,69	4,00	0,37
$\ln \rho_{1,9}$	—0,71	—0,63	5,47	0,94
$\ln \rho_{0,5}$	—0,85	—0,47	7,68	0,60
$\ln \rho_{0,4}$	—0,72	—0,57	2,35	0,49
Вес факторов, %	36,9	21,4	—	—

Изложенные результаты не исчерпывают возможностей компонентного анализа. Он может быть использован для решения чисто прагматических задач, аналогичных тем, которые рассматривались в главе I. Компонентный анализ позволяет в определенной степени усовершенствовать регрессионный анализ и воспроизвести его в несколько ином варианте, который лучше отвечает существу задачи. Если раньше мы использовали модель вида $\mathbf{Y} = \mathbf{X}\beta + \mathrm{e}$, то теперь можем перейти к модели $\mathbf{Y} = \mathbf{Z}\Phi + \mathrm{e}$, т. е. строить регрессию на главные компоненты [5, 75] (Φ — вектор коэффициентов регрессии). В этом случае коэффициенты регрессии Φ при главных компонентах оказываются более устойчивыми, чем коэффициенты β_j, а сходимость по ортогональным компонентам, как правило, достигается намного быстрее. Поэтому перенос на новый объект уравнения регрессии, найденного на другом объекте, при использовании главных компонент более оправдан. Коэффициенты β_j также можно выразить через Φ, при этом они окажутся более устойчивыми, так как будут отсеяны компоненты z_j, определяющие лишь случайные, плохо контролируемые флуктуации исходных показателей. Найденные таким образом коэффициенты β более пригодны для расчетов неизвестных значений y.

В ранее принятых нами обозначениях $\Phi = \Lambda^{-1/2}\mathbf{U}r$; $\beta = \mathbf{U}\Lambda^{-1/2}\Phi$, где $\mathbf{r}$ — вектор коэффициентов корреляции между предсказуемой переменной y и исходными переменными, образующими матрицу $\mathbf{X}$. Оптимальную размерность вектора Φ (оптимальное число главных компонент) можно определить по значению множественного коэффициента корреляции

$$R(y, z_1, z_2, \ldots, z_q) = \sqrt{\sum_{i=1}^{q} r^2(y, z_i)},$$

где $r(y, z_l) = \Phi_l \sqrt{\lambda_l}$ — парный коэффициент корреляции между переменной y и соответствующим вектор-столбцом матрицы Z.

Существуют номограммы, по которым для данного выборочного множественного коэффициента корреляции $R(y, z_1, z_2, \ldots, z_q)$ выборки с объемом n и числом переменных q можно определить минимальное значение истинного коэффициента множественной корреляции при заданном уровне значимости.

При использовании метода главных компонент было показано, что первый фактор (или сумма двух первых факторов), извлеченный из комплекса геолого-геофизических показателей, характеризующих карбонатные разрезы верхнего мела (Северо-Восточное Предкавказье и Южный Дагестан) и ордовика (Прибалтика), тесно связан с изменением блоковой пористости в разрезе. Привлекая эту зависимость, целесообразно оценить блоковую пористость лишь по данным комплекса геофизических исследований скважин с помощью регрессионной модели, учитывающей главные компоненты (факторы).

Пример 1. На основании геофизических данных, характеризующих разрез верхнего мела в скв. 154-Карабулак-Ачалуки, намеренно без учета измеренных значений (признаков) пористости $K_п$ и нерастворимого остатка $C_{н.о}$ согласно процедурам регрессионной модели (по главным компонентам) была получена матрица нагрузок. Ее сравнение с матрицей, приведенной в табл. II.4, показало, что значения нагрузок при входящих в анализ геофизических показателях в первом и во втором факторах обеих матриц имеют различия лишь в третьем знаке. Отсюда следует идентичность факторов, выделяемых из смешанных геолого-геофизических сведений и только из геофизических данных. Но поскольку первый фактор контролирует в основном межзерновую пористость, а второй — каверновую, то очевидно, что их сумма отражает изменение блоковой пористости по разрезу.

Найдя регрессионные коэффициенты Φ_1 и Φ_2 при двух первых факторах, запишем уравнение, связывающее блоковую пористость с главными компонентами, которые, что очень важно, получены только по геофизическим данным:

$$K_{п.бл} = \sigma_{K_{п.ол}} (\Phi_1 z_1 + \Phi_2 z_2) + \overline{K}_{п.бл},$$

где $\overline{K}_{п.бл}$ — среднее значение блоковой пористости; $\sigma_{K_{п.бл}}$ — ее стандартное отклонение.

Очевидно, что подходящей оценкой для $\overline{K}_{п.бл}$ и $\sigma_{K_{п.бл}}$ являются значения, приведенные в табл. II.4, которые соответственно равны 5,66 и 2,57.

На рис. II.1 приведена кривая изменения по разрезу блоковой пористости, рассчитанной с помощью уравнения регрессии (по главным компонентам). Сравнивая ее с фактическими изменениями пористости $K_п$, определенной по керну, можно убедиться в близости этих данных. Более того, в интервале 2250—2270 м вынос керна оказался крайне низким, и в этом промежутке не было количественных определений пористости. Однако по регрессионному уравнению на основе лишь геофизической информации эти значения пористости были восстановлены без особого труда. Отсюда следует, что подобный подход является плодотворным при количественной оценке геологических (прямых) характеристик по геофизическим (косвенным) наблюдениям в тех интервалах разреза, где наблюдений по каким-либо причинам не было.

Пример 2. Выявленная устойчивость смысла первых факторов, выделяемых из геолого-геофизических данных по одновозрастным разрезам различных скважин исследуемого региона, позволяет оценивать пористость по геофизическим данным при отсутствии керна. Для этого также используем регрессионный анализ с учетом главных компонент. Интересно сравнить результаты традиционных методов оценки блоковой пористости в карбонатных коллекторах с результатами, полученными на основе факторного анализа и регрессионного анализа по главным компонентам.

Для этого обратимся к работе А. М. Нечая [63], в которой даны определения блоковой пористости методом БЭЗ вдоль всего ствола скважины, вскрывающей верхнемеловые отложения Карабулак-Ачалукской площади. В этой работе приведены значения параметра пористости $P_п$ и кажущегося удельного сопротивления ρ' и ρ'' интерпретируемых пластов, измеренного методом БМ при двух промывочных жидкостях разного сопротивления (ρ_0' и ρ_0''), а также показания НГМ.

Весь комплекс показаний был подвергнут компонентному (факторному) анализу. Матрицы нагрузок (два первых фактора) приведены в табл. II.11,

Таблица II.11

Результаты компонентного анализа (исходные геолого-геофизические данные заимствованы из работы [63])

Переменная	Матрица нагрузок I		Матрица нагрузок II		Среднее значение переменной	Стандартное отклонение
	Факторы					
	Первый	Второй	Первый	Второй		
$K_{п.\,бл}$	0,91	0,09	—	—	3,58	2,29
$\ln \rho'$	—0,96	0,16	—0,96	0,19	3,54	0,96
$\ln \rho''$	—0,89	0,25	—0,90	0,28	3,27	0,88
$\ln P_п$	—0,88	0,15	—0,87	0,17	2,15	0,94
$J_{НГМ}$	—0,58	—0,76	—0,60	—0,75	3,69	1,01
Вес факторов, %	73,2	13,8	71,3	17,5	—	—

причем матрица I соответствует пяти показателям, а матрица II — четырем (признак $K_{п.\,бл}$ исключен из анализа). Интерпретация нагрузок и сравнение этих двух матриц позволяют сделать следующее заключение. Первая главная компонента (первый фактор) контролирует блоковую пористость. Вторая компонента (второй фактор), как видно, не связана с пористостью. Значительная нагрузка во втором факторе при переменной НГМ может свидетельствовать о том, что этот фактор контролирует изменение содержания в породе глинистого материала.

После исключения из анализа признака «блоковая пористость» нагрузки при оставшихся переменных не претерпели существенного изменения. Иными словами, смысл первого фактора, выделенного только из геофизических данных, остался прежним; то же происходит и со вторым фактором. Отсюда следует, что оценка блоковой пористости может быть успешно выполнена с помощью регрессионной модели по методу главных компонент, причем можно воспользоваться только первой компонентой.

По расчетам $\Phi_1 = 0{,}87$; следовательно, уравнение, связывающее блоковую пористость с главной компонентой, имеет вид

$$K_{п.\,бл} = 2{,}29 \cdot 0{,}87 z_1 + 3{,}58.$$

Очевидно, что первая главная компонента выражается через геофизические признаки следующим образом:

$$z_1 = -0.96 \ln \rho' - 0.90 \ln \rho'' - 0.87 \ln P_п - 0.60 J_{НГМ},$$

признаки центрированы и нормированы (стандартизированы).

Пример 3. Теперь обратимся к скв. 833-Малгобек (Северо-Восточное Предкавказье), вскрывшей также отложения верхнемелового комплекса, который охарактеризован здесь только геофизическими измерениями (БЭЗ и НГМ). Матрица нагрузок (первые два фактора) приведена в табл. II.12. Оценка пористости в этой скважине осуществлена на основе уравнения, полученного в результате использования регрессионной модели по главным компонентам:

$$K_{п.\,бл} = 2{,}29\,(0.58 z_1 + 0.005 z_2) + 6.0.$$

Здесь при оценке блоковой пористости кроме первого использован и второй фактор, который, как было показано выше, может контролировать изменение каверновой пористости.

Таблица II.12

Результаты компонентного анализа (скв. 833-Малгобек)

Переменная	Факторы		Среднее значение переменной	Стандартное отклонение
	Первый	Второй		
$J_{НГМ}$	−0,87	0,05	3,07	0,73
$\ln \rho_8$	−0,97	−0,32	2,55	0,77
$\ln \rho_4$	−0,96	−0,19	2,72	0,84
$\ln \rho_2$	−0,95	−0,10	2,98	0,84
$\ln \rho_1$	−0,95	0,07	3,24	0,88
$\ln \rho_{0.5}$	−0,83	0,13	2,77	0,56
$\ln \rho_{БМ}$	−0,83	0,49	2,65	0,67
Вес факторов, %	83,1	5,9	—	—

Среднее значение блоковой пористости определено по методу БЭЗ, оно оказалось равным 6%; стандартное отклонение и вектор корреляции (необходим для расчетов компонент вектора Ф) взяты по аналогии с данными по скв. 154-Карабулак-Ачалуки. Изменение пористости по разрезу скв. 833-Малгобек, восстановленное по факторам, представлено на рис. II.5.

Пример 4. Аналогичная оценка пористости проведена для пород, вскрытых скв. 38-Дурбе. В данном случае уравнение, связывающее пористость с первым фактором (второй фактор, как было показано ранее, в этом случае не связан с блоковой пористостью), имеет следующий вид:

$$K_{п.\,бл} = 3{,}96 \cdot 0.597 z_1 + 9{,}61.$$

Почти аналогичный вид принимает уравнение регрессии, полученное по данным, характеризующим разрез скв. 3-Ремте.

На рис. II.3 и II.4 приведены кривые изменения пористости, рассчитанные по соответствующим уравнениям регрессии. Сравнивая эти кривые с фактическими значениями пористости (определена по керну), можно убедиться, что они близки. Некоторое различие объясняется тем, что пористость, установленная по образцам, относится в основном к межзерновому типу, в то время как полученная по геофизическим данным является суммарной (межзерновая плюс каверновая).

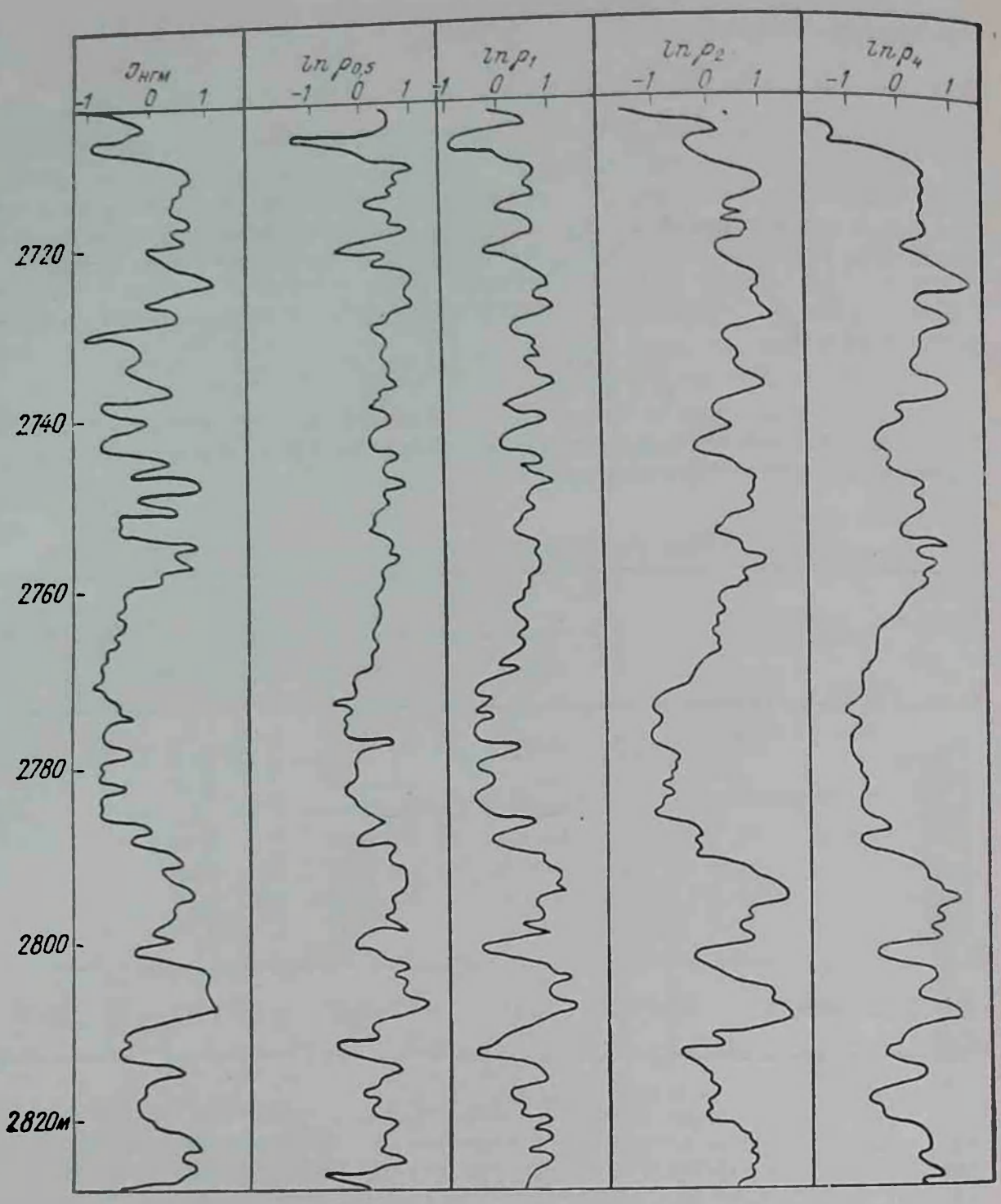
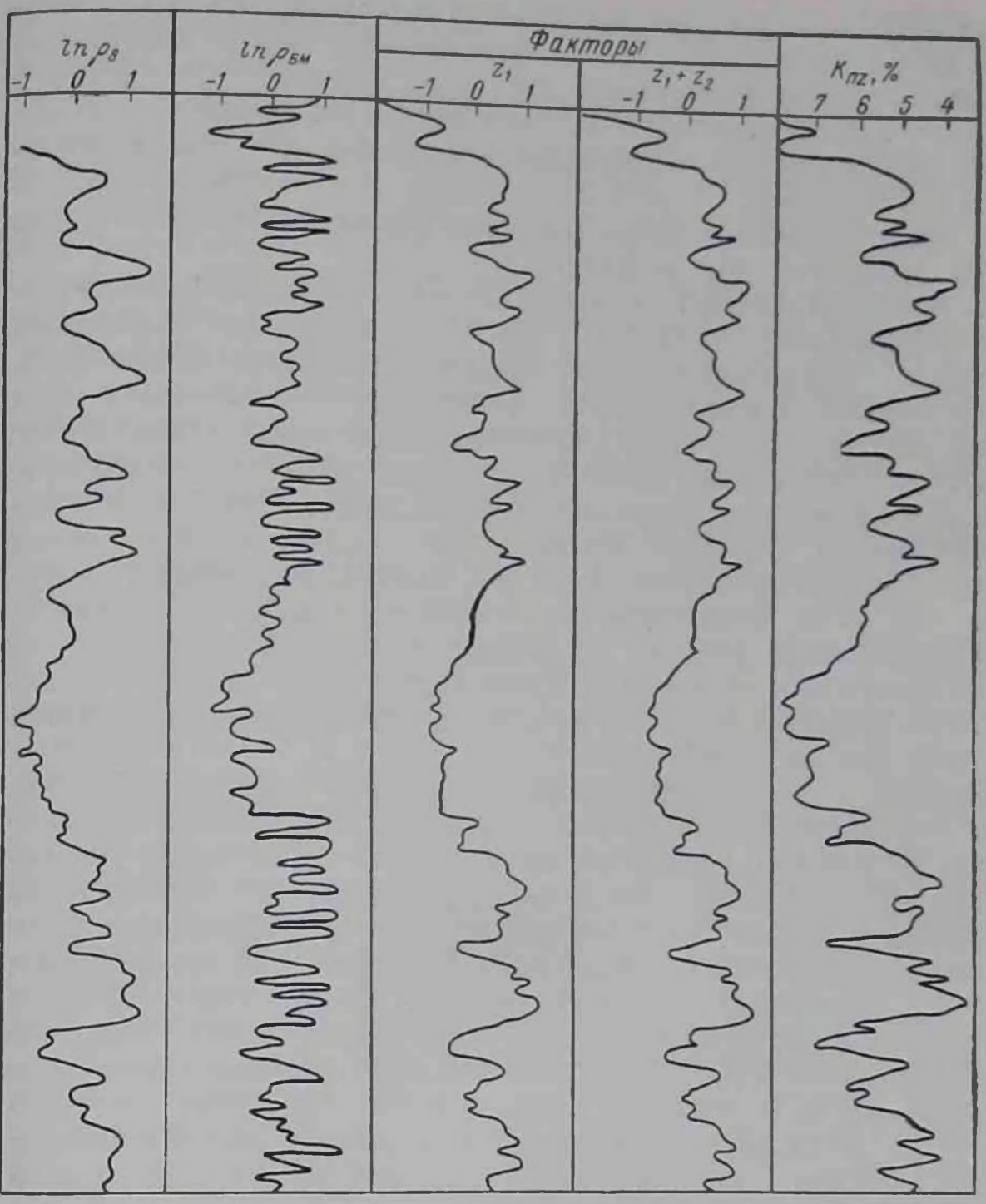


Рис. II.5. Геофизические переменные и выделенные из них

Из приведенных примеров видно, что оценка блоковой пористости может быть успешно осуществлена с помощью регрессионной модели с учетом главных компонент. Однако при этом необходимо знать среднее значение пористости для всего изучаемого комплекса отложений, его стандартное отклонение и коэффициенты корреляции между пористостью и анализируемыми геофизическими показателями. Эта информация, вообще говоря, на определенных стадиях разведки имеется. Так, среднее значение пористости достаточно удовлетворительно может быть оценено по данным либо БЭЗ, либо НГМ, либо по керну. При этом весь исследуемый интервал карбонатного разреза целесообразно сна-



ортогональные компоненты (факторы) в скв. 833-Малгобек.

чала принять за единый пласт. Что касается стандартного отклонения, то его легко установить, анализируя изменение пористости по керну соседних скважин.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Подводя итог выполненной интерпретации, необходимо отметить следующее.

1. Истолкование большого числа геологических и геофизических данных с целью выявления тех или иных природных факторов, обусловливающих взаимосвязанную изменчивость исходных наблюдений, целесообразно проводить с помощью компонентного

анализа. Этот анализ дает возможность либо утвердиться в априорных представлениях об изменении состава пород, слагающих разрез, их коллекторских свойств, типа флюида, заполняющего пустоты, либо по-новому объяснить причину изменений, основываясь на конкретной информации о взаимосвязях между исходными показателями. Иногда это приводит к разрушению привычных представлений о взаимодействии различных свойств пород и их поведении по разрезу.

2. Большая система характеризующих разрезы геолого-геофизических показателей (от 7 до 18) может быть эффективно заменена значительно меньшим числом факторов (не более трех), ответственных за изменчивость анализируемых показателей. При такой замене сохраняется основная часть информации (более 70%), достаточная для детального освещения исследуемого объекта. Кроме того, исходная информация, расщепленная на составные элементы (главные компоненты), поддается более ясной и определенной трактовке. Так, для карбонатных толщ установлено, что если комплексы геолого-геофизических показателей, характеризующие разрезы различных скважин в одном районе, почти одинаковы, то смысл каждого выделенного фактора во всех разрезах (в фиксированном стратиграфическом диапазоне) остается большей частью постоянным.

Первый фактор, выделенный из системы геолого-геофизических показателей, характеризующих верхнемеловые отложения Северо-Восточного Предкавказья и Южного Дагестана, контролирует распределение объема раствора (фильтрата промывочной жидкости) в породах-коллекторах и тем самым изменение межзерновой (и частично каверновой) пористости. В ордовикских разрезах Прибалтики первый фактор, кроме того, отражает колебание глинистости в породах. Надо заметить, что этим свойством практически однозначно определяется изменение межзерновой пористости. В подобных случаях тесная связь (генетического характера) двух параметров приводит к выводу, что в принципе нельзя исключать влияние глинистости при оценке межзерновой пористости по геофизическим данным. Совершенно ясно, что устранив влияние одной особенности, мы тем самым тотчас исключим влияние и другой.

Второй фактор, выделяемый при помощи компонентной модели, объясняет значительно меньшую долю дисперсии системы анализируемых показателей. В разрезах верхнего мела Северо-Восточного Предкавказья и Южного Дагестана второй фактор, как правило, контролирует изменение содержания нерастворимого остатка в породе, а так как оно часто связано с увеличением каверновой пористости, то можно сказать, что второй фактор контролирует изменение каверновой пористости. В разрезах ордовика Прибалтики второй фактор отражает изменение некоторой части межзерновой пористости. Но, как показал анализ, происхождение пустот, определяющих эту пористость, может быть

искусственным: они возникли в процессе сушки и экстрагирования образцов керна.

Исключив влияние первых двух факторов, мы пришли к заключению, что третий фактор в некоторых случаях контролирует изменение трещиноватости (Северо-Восточное Предкавказье). Большей же частью этот фактор настолько сложный, что однозначное его истолкование крайне затруднительно. Как правило, напрашивается вывод, что трещиноватость вызывает только незначительную изменчивость геофизических показателей, уловить которую достаточно сложно, так как многие другие особенности карбонатных пород (специфические особенности минерального скелета и структуры пород, свойств флюида и т. п.) могут быть причиной соизмеримой по интенсивности изменчивости анализируемых геофизических показателей.

3. В определенных случаях наличие факторов с достаточно однозначной геологической трактовкой позволяет, с одной стороны, проследить их изменение по разрезу исследуемых толщ, а с другой — восстановить по ним ту или иную исходную переменную. Восстановленные переменные получаются как бы отфильтрованными от многих воздействий случайного характера. Исходя из того, что радио- и электрометрические показатели (НГМ, БЭЗ, БМ) заключают в себе разную информацию о типах пористости, и сравнивая восстановленные по факторам (в нормированном виде) геофизические показатели, по расхождению между ними можно проследить, как изменяется по разрезу тип пористости карбонатного коллектора. Зоны расхождения приурочены к интервалам, которые прямыми и косвенными методами определяются как трещиноватые.

4. Регрессионный анализ по методу главных компонент имеет то преимущество, что коэффициенты регрессии оказываются более устойчивыми, поэтому уравнение регрессии, найденное на одном объекте, можно перенести на другой без опасения получить существенные искажения в результатах. В частности, это позволило восстановить значения пористости, исходя лишь из геофизических данных, в тех интервалах разреза, где прямой информации о пористости не было.

5. Интерпретация геолого-геофизических показателей на основе компонентного анализа позволяет судить о типе флюида, насыщающего поры. Так, установлено, что коллекторы верхнего мела (Восточное Предкавказье, Южный Дагестан) и ордовика (Прибалтика) в исследуемых интервалах вмещали фильтрат промывочной жидкости, который еще в процессе бурения пропитывал блоки пород. При смене минерализации промывочной жидкости новый фильтрат успевает заполнить не только трещины и каверны, развитые по трещинам, но и притрещинные пустоты (часть межзерновых пустот блоков пород). Этот факт заставляет осторожно относиться к результатам оценки трещинной пористости (или пористости, определяемой суммой трещинных пустот

и каверн, развитых по трещинам) с помощью метода «двух растворов». При насыщении притрещинной зоны непременно возникает эффект фиктивного раскрытия трещин, что обязательно отразится на изменении удельного электрического сопротивления породы, расшифровка которого покажет завышенную трещинную пористость.

В газосодержащих коллекторах (отложения карбона и перми в скважинах Вуктыл) процесс пропитки не наблюдается. Наличие газа, вероятно, сильно противодействует формированию зоны проникновения фильтрата промывочной жидкости. Это приводит к тому, что смысл первого фактора может оставаться прежним (например, контроль межзерновой пористости пород), но отражения в геофизических показателях (БЭЗ, БМ, НГМ) этот фактор не получает. Иными словами, в этом случае изменчивость геофизических показателей связывается с изменением не емкости пород, а содержания остаточной воды в порах, особенностей структуры породы, ее минерального скелета и т. п.

6. По данным компонентного анализа геологического материала, характеризующего верхнемеловые отложения скв. 30-Карабулак, формирование емкости карбонатных коллекторов зависит главным образом от их первичной структуры и вещественного состава. Причем если структура пород в основном определяет первичную пористость, то вещественный состав обусловливает как первичную, так и вторичную пористость. Основная доля емкости карбонатных коллекторов приходится на межзерновые поры седиментационного происхождения (этот вывод может быть распространен и на породы ордовика). Доля вторичных пустот (кавернового и трещинного типа) невелика и составляет в среднем не более 1%. Следовательно, при прогнозировании карбонатных коллекторов по площади или по разрезу целесообразно ориентироваться на выявление зон и участков с повышенной межзерновой пористостью первичного (седиментационного) происхождения. Это успешно можно осуществить, если комплекс геофизических показателей интерпретировать с помощью компонентной модели.

В заключение отметим, что сопоставление тех или иных одноименных факторов по разным скважинам может помочь в решении задач по расчленению и корреляции разрезов. Очевидно, что с учетом генетической информации, а факторы, несомненно, такую информацию содержат, эта задача может быть решена намного обоснованнее и эффективнее.

Глава III

ДИАГНОСТИКА ГАЗОНОСНЫХ И ВОДОНОСНЫХ ПЛАСТОВ

ПОСТАНОВКА ЗАДАЧИ И МЕТОДЫ ЕЕ РЕШЕНИЯ

Остановимся на решении одной из наиболее важных задач нефтепромысловой геологии и геофизики: по комплексу геофизических данных, характеризующих пласты, необходимо определить, относится изучаемый пласт к нефте-, газо- или водонасыщенным. Этот вопрос является частным случаем более общей задачи, которая может быть сформулирована следующим образом: по комплексу косвенных признаков установить, с каким объектом мы имеем дело. Такие задачи называют задачами классификации или распознавания образов; иногда их именуют также диагностическими.

Задача диагностики объекта требует отнести этот объект, используя определенный набор характеризующих его признаков (в дальнейшем этот набор целесообразно обозначать вектором x с компонентами $x_1, x_2, \ldots, x_m$), к одному из классов К. Классами могут служить различные объекты, например породы-коллекторы, породы-неколлекторы, нефте-, водо-, газоносные или пустые пласты, различные по литологическому составу типы пород и т. п. Признаками в интересующем нас плане должны являться данные геофизических исследований скважин, характеризующие изучаемые объекты. Очевидно, что указанное отнесение выполнить тем легче, чем сильнее различаются геофизические показатели, свойственные разным объектам.

Но в условиях, когда существует большое число мешающих факторов и геофизические показатели являются «интегральными», зависящими не от одного фактора, а от их совокупности, это различие, как правило, уловить трудно. Поэтому необходимы методы, которые, во-первых, учитывали бы специфику геофизического материала, а во-вторых, позволили сравнивать геофизические показатели в комплексе. Традиционные методы здесь в принципе оказываются бессильными: дело даже не столько в том, что они не учитывают многофакторность объекта и наличие случайных флуктуаций в геофизических показателях, сколько в том, что они не в состоянии охватить и обозреть всю ситуацию в целом, во всем ее многообразии и сложности.

В принципе задачи по диагностике объектов возникли давно, но начало их правильной формулировки, постановки и реализации следует отнести к тому времени, когда стал создаваться аппарат теории классификации и распознавания образов. Для интерпретации промыслово-геофизического материала этот метод пришелся как нельзя кстати, дав нетрадиционное решение в общем-то традиционной задачи.

С математической точки зрения распознавание заключается в следующем [14]. Вводится некоторое преобразование

$$\mathbf{Y} = \varphi(\mathbf{x}),$$

или в векторной форме

$$y_1 = \varphi_1(\mathbf{x}), \quad y_2 = \varphi_2(\mathbf{x}), \quad \ldots, \quad y_m = \varphi_m(\mathbf{x}),$$

где $\mathbf{x}$ — вектор признаков объекта.

Преобразование $\varphi(\mathbf{x})$ ставит в соответствие исходному описанию объекта некоторое новое описание $\mathbf{Y}$. Вводится условие, согласно которому объект $\mathbf{x}$ относится к первому классу, если выполняется неравенство

$$\sum_{i=1}^{m} \lambda_i \varphi_i(\mathbf{x}) > 0;$$

в противном случае объект относится ко второму классу.

Такой схеме легко дать простую геометрическую интерпретацию: в пространстве признаков $\mathbf{x}$ задана гиперповерхность

$$\sum_{i=1}^{m} \lambda_i \varphi_i(\mathbf{x}) = 0,$$

которая делит это пространство на два полупространства (выше и ниже гиперповерхности). Если объект $\mathbf{x}$ находится по одну сторону от поверхности, т. е. $\sum_{i=1}^{m} \lambda_i \varphi_i(\mathbf{x}) > 0$, то он соответствует первому классу; если же по другую от нее сторону, то второму. Такие гиперповерхности называются разделяющими.

Каждой гиперповерхности пространства $\mathbf{X}$ в пространстве $\mathbf{Y}$ с координатами $y_1 = \varphi_1(\mathbf{x})$, $y_2 = \varphi_2(\mathbf{x})$, ..., $y_m = \varphi_m(\mathbf{x})$ соответствует гиперплоскость

$$\sum_{i=1}^{m} \lambda_i y_i = 0.$$

Тем самым в пространстве $\mathbf{Y}$ разделяющая гиперповерхность заменяется разделяющей гиперплоскостью, в связи с чем пространство $\mathbf{Y}$ получило название спрямляющего. В этом пространстве каждому объекту $\mathbf{x}$ ставится в соответствие вектор $\mathbf{Y} = (y_1, y_2, \ldots, y_m)$. Этот вектор относится к тому или другому классу в зависимости от его расположения относительно разделяющей гиперплоскости $\sum_{i=1}^{m} \lambda_i y_i = 0$.

Для решения задачи классификции необходимо иметь так называемую обучающую последовательность объектов, принадлежность которых к тому пли иному классу известна. Задача распознавания заключается в том, чтобы, используя обучающую

последовательность, построить такую математическую процедуру (решающее правило), которая позволяла бы классифицировать вновь предъявляемые незнакомые объекты. Поиск решающего правила проводится при определенных требованиях: считается, что объекты, предъявляемые для классификации, выбираются случайно в соответствии с функцией распределения вероятностей $P(\mathbf{x})$ появления объектов того или иного класса. Иначе говоря, функция $P(\mathbf{x})$ является характеристикой среды, в которой будет работать классифицирующее правило, она показывает, как часто могут встречаться объекты каждого класса. Обучающая последовательность составляется из объектов, выбранных случайно и независимо из той среды, для которой будет оцениваться качество полученного решающего правила.

Обычно считается, что задано некоторое множество решающих правил и задача заключается в том, чтобы найти в этом множестве нужное правило, используя обучающую последовательность, состоящую из фиксированного числа объектов. Сформулируем критерии, оценивающие качество решающих функций (правил).

Пусть объекты $\mathbf{x}$, появляющиеся случайно и независимо в соответствии с функцией распределения вероятностей $P(\mathbf{x})$, классифицируются некоторым «идеальным» образом по классам (для простоты положим, что $K = 2$) согласно условной вероятности $P(\omega|\mathbf{x})$, где $\omega = 1$, если объект $\mathbf{x}$ отнесен к первому классу, и $\omega = 0$, если ко второму. Ни характеристика среды $P(\mathbf{x})$, ни правило классификации $P(\omega|\mathbf{x})$ нам неизвестны. Известно лишь, что обе функции существуют и, следовательно, существует совместное распределение вероятностей

$$P(\omega, \mathbf{x}) = P(\mathbf{x})\, P(\omega|\mathbf{x}).$$

Пусть известно множество Ω решающих правил $F(\mathbf{x}, \alpha)$, каждое из которых определяется заданием параметра α. Все правила $F(\mathbf{x}, \alpha)$ — характеристические функции, т. е. могут принимать только значения 0 (если вектор $\mathbf{x}$ отнесен к первому классу) и 1 (если вектор $\mathbf{x}$ отнесен ко второму классу). Для каждой функции $F(\mathbf{x}, \alpha)$ может быть определено качество $P(\alpha)$ как вероятность несовпадения классификаций объектов, полученных с помощью «идеального» правила $P(\omega|\mathbf{x})$ и принятого правила $F(\mathbf{x}, \alpha)$. Если пространство $\mathbf{X}$ дискретно и состоит из точек $\mathbf{x}_1$, $\mathbf{x}_2$, ..., $\mathbf{x}_N$, то качество $P(\alpha)$ функции $F(\mathbf{x}, \alpha)$ определяется как

$$P(\alpha) = \sum_{\omega=0}^{1} \sum_{i=1}^{N} [\omega - F(\mathbf{x}_i, \alpha)]^2\, P(\mathbf{x}_i)\, P(\omega|\mathbf{x}_i),$$

где $P(\mathbf{x}_i)$ — вероятность появления объекта $\mathbf{x}_i$.

В общем случае

$$P(\alpha) = \int_{\mathbf{x},\, \omega} [\omega - F(\mathbf{x}, \alpha)]^2\, dP(\mathbf{x}, \omega).$$

Очевидно, что чем меньше вероятность несовпадения классификаций, найденных с помощью правил $P(\omega|\mathbf{x})$ и $F(\mathbf{x}, \alpha)$, тем лучше правило $F(\mathbf{x}, \alpha)$. Следовательно, среди всех функций $F(\mathbf{x}, \alpha)$ надо найти такую $F(\mathbf{x}, \alpha_0)$, которая минимизирует по α функционал $P(\alpha)$. Задача о поиске минимума по α функционала вида

$$R(\alpha) = \int Q(\mathbf{z}, \alpha)\, dP(\mathbf{z})$$

получила название задачи минимизации среднего риска. Функцию двух переменных $Q(\mathbf{z}, \alpha)$ часто называют функцией потерь. Отсюда ясно, что задача обучения распознаванию образов есть частный случай задачи минимизации среднего риска. Функция потерь здесь задана в виде $[\omega - F(\mathbf{x}, \alpha)]^2$, а функция, минимум которой надо определить,

$$P(\alpha) = \int [\omega - F(\mathbf{x}, \alpha)]^2\, dP(\mathbf{x}, \omega).$$

Известно три пути решения задачи о поиске наилучшего правила, обеспечивающего минимизацию среднего риска.

Первый путь. По выборке восстанавливаются неизвестные многомерные функции распределения признаков объектов первого и второго классов, чтобы затем по восстановленным функциям построить оптимальную классификацию объектов. В этом случае классификация проводится с помощью так называемой дискриминантной функции. Применение дискриминантных функций для классификации получило наиболее широкое распространение во многих областях знаний [82], и в частности в геологии [57].

Следует иметь в виду, что построение дискриминантной функции — задача значительно более простая, чем восстановление двух m-мерных функций распределения вероятностей. Правда, если признаки независимы, т. е. координаты вектора $\mathbf{x} = (x_1, x_2, \ldots, x_m)$ распределены независимо, то функция $P(\mathbf{x}|\omega) = P(x_1|\omega) \times P(x_2|\omega) \ldots P(x_m|\omega)$ и дело значительно упрощается, так как в этом случае задача о восстановлении многомерных функций распределения вероятностей вырождается в задачу восстановления $2m$ одномерных функций

$$P(x_i|\omega = 0), \quad P(x_i|\omega = 1) \quad (i = 1, 2, \ldots, m).$$

Обозначим $P(\mathbf{x}|\omega = 1)$ и $P(\mathbf{x}|\omega = 0)$ — плотность распределения вероятностей векторов соответственно первого и второго класса; $P(\omega = 1)$ и $P(\omega = 0)$ — вероятности появления объектов соответственно первого и второго класса. При этих обозначениях дискриминантная функция имеет вид

$$F(\mathbf{x}) = \theta\{\ln P(\mathbf{x}|\omega = 1) - \ln P(\mathbf{x}|\omega = 0) + \ln [P(\omega = 1)/P(\omega = 0)]\}.$$

Если векторы распределены нормально, то

$$F(\mathbf{x}) = 0 \{ (\mathbf{x}-\mu_2)' \Gamma_2^{-1} (\mathbf{x}-\mu_2) - (\mathbf{x}-\mu_1)' \Gamma_1^{-1} (\mathbf{x}-\mu_1) - \\ - \ln (|\Gamma_1|/|\Gamma_2|) + \ln [P(\omega=1)/P(\omega=0)]\},$$

где μ_1, Γ_1 и μ_2, Γ_2 — математическое ожидание и ковариационная матрица векторов соответственно первого и второго класса.

Если ковариационные матрицы равны между собой (обозначим их Г), то представляет интерес величина, называемая расстоянием Махаланобиса:

$$\Delta^2 = (\mu_1 - \mu_2)' \Gamma^{-1} (\mu_1 - \mu_2).$$

Через эту функцию вычисляется вероятность правильной классификации оптимального решающего правила

$$P = \frac{1}{2\pi} \int_{-\infty}^{\Delta/2} e^{-t^2/2}\, dt = \Phi(-\Delta/2),$$

где Ф — интегральная функция нормального распределения.

Критерий Махаланобиса характеризует расстояние между классами. Он представляет собой отношение расстояния между центрами двух классов к дисперсии проекции векторов каждого класса на направляющий вектор разделяющей гиперплоскости:

$$\Delta^2 = (\mu_1 - \mu_2)^2/\sigma^2.$$

Критерий Махаланобиса позволяет определить нижнюю оценку решающего правила, построенного по выборке фиксированного объема. Это расстояние сужается с уменьшением размерности пространства, одновременно снижается вероятность правильной классификации с помощью оптимального в этом пространстве решающего правила.

Второй путь. Параметр α, доставляющий минимум функционалу $P(\alpha) = \int [\omega - F(\mathbf{x}, \alpha)]^2 dP(\mathbf{x}, \omega)$, отыскивается путем организации рекуррентной процедуры. Одним из методов, реализующих эту идею при геологических исследованиях, является метод «потенциальных» функций [14]. Суть его заключается в следующем. На пространстве исходных признаков $\mathbf{X}$ задается функция расстояния между точками, называемая потенциалом, которая, следовательно, определяет близость двух точек $\mathbf{x}_1$ и $\mathbf{x}_0$. Примерами потенциальной функции могут служить

$$K(\mathbf{x}_1, \mathbf{x}_0) = 1/(1 + r^2\alpha); \quad K(\mathbf{x}_1, \mathbf{x}_0) = e^{-r^2\alpha},$$

где

$$r = \sqrt{\sum_{i=1}^{m} \left(x_i^{(0)} - x_i^{(1)}\right)^2}$$

— расстояние от точки $\mathbf{x}_0 = (x_1^{(0)}, x_2^{(0)}, \ldots, x_m^{(0)})$ до точки $\mathbf{x}_1 = (x_1^{(1)}, x_2^{(1)}, \ldots, x_m^{(1)})$; α — константа.

Эти функции используются для построения потенциального поля $\Phi(\mathbf{x})$, что осуществляется следующим образом. Для первого члена обучающей последовательности $\mathbf{x}_1$ определяется потенциал с центром в этой точке. Ему приписывается знак в зависимости от того, к какому классу относится предъявляемый объект: положительный, если к первому, и отрицательный, если ко второму. Для второго объекта обучающей последовательности вычисляется потенциал $K(\mathbf{x}_2, \mathbf{x}_1)$. Если его знак положительный и объект относится к первому классу, то потенциальное поле на пространстве $\mathbf{X}$ не меняется; если же объект относится ко второму классу, то в точке $\mathbf{x}_2$ потенциалу приписывается отрицательный знак и потенциальное поле (суммарный потенциал) изменяется на

$$\Phi(\mathbf{x}) = K(\mathbf{x}, \mathbf{x}_1) - K(\mathbf{x}, \mathbf{x}_2).$$

Таким образом, по обучающей последовательности суммарное поле каждый раз исправляется так, чтобы оно соответствовало предъявляемым объектам. Все пространство разделяется на две части: в одной значения суммарного потенциала положительные (соответствует объектам первого класса), в другой отрицательные (соответствует объектам второго класса). Разделяющая поверхность в этом случае находится легко: на ней суммарный потенциал принимает нулевые значения.

Третий путь. Неизвестный функционал

$$P(\alpha) = \int Q(\mathbf{z}, \alpha)\, dP(\mathbf{z})$$

заменяется функцией

$$P_{\text{эмп}}(\alpha) = \frac{1}{l} \sum_{i=1}^{l} Q(\mathbf{z}_i, \alpha),$$

построенной по случайной и независимой выборке $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_l$.

Функция $P_{\text{эмп}}(\alpha)$ исчисляет эмпирический риск. Идея метода состоит в том, чтобы найти такие значения параметров $\alpha = \alpha^*$, при которых функция $P_{\text{эмп}}(\alpha)$ достигает минимума, а затем в качестве решения задачи предложить функцию с этими значениями параметров, т. е. $Q(\mathbf{z}, \alpha^*)$. Такой подход к решению задачи называется методом минимизации эмпирического риска. Преимуществом этого метода является то обстоятельство, что число объектов в обучающей последовательности можно значительно сократить, так как она используется до тех пор, пока ее объекты не будут разделены гиперплоскостью безошибочно. Этот момент означает, что выбрано решающее правило, минимизирующее эмпирический риск.

Одним из алгоритмов, реализующим идею поиска минимума эмпирического риска, является «Кора». Этот алгоритм плодотворно использован при разделении по промыслово-геофизическим характеристикам нефте- и водоносных пластов [29, 30, 58]. Суть его заключается в следующем.

На множестве объектов первого $\{x_{\omega=1}\}$ и второго $\{x_{\omega=0}\}$ классов задается множество характеристических функций $\psi(x, \tau)$, которые называются признаками. В качестве класса характеристических функций $\psi(x, \tau)$ берутся все возможные конъюнкции двух-трех переменных. Из множества признаков алгоритм выделяет так называемые достаточные. Для векторов первого класса достаточным является такой признак $\psi(x, \tau)$, который на всех векторах второго класса принимает значение 0, а на некоторых векторах первого класса 1. Аналогично определяются достаточные признаки второго класса. Алгоритм выбирает некоторое одинаковое число достаточных признаков второго и первого класса. При этом важно, чтобы для каждого объекта обучающей последовательности нашлось несколько достаточных признаков.

Опознание нового объекта проводится путем подсчета числа достаточных признаков первого и второго класса, принявших на этом объекте значение 1. Объект относится к тому классу, для которого число таких признаков больше. Характерной особенностью алгоритма «Кора» является небольшая емкость класса решающих правил и чрезвычайно простой (хотя и эвристический) метод поиска правила, минимизирующего эмпирический риск.

Мы обратимся к описанию и практическому применению двух классификационных процедур. Первая использует принципы и алгоритмы последовательного статистического анализа Вальда [13], основанного на построении отношения функций максимального правдоподобия. Вторая процедура — линейный дискриминантный анализ [2, 80, 82] — относится к классу решающих правил, где восстанавливаются плотности распределения признаков, используемых при классификации объектов исследования.

ПОСЛЕДОВАТЕЛЬНЫЙ СТАТИСТИЧЕСКИЙ АНАЛИЗ

Рассматриваемая нами задача распознавания газо- и водоносных пластов является альтернативной. К классу собственно альтернативных (дихотомических) относятся такие задачи, в которых требуется принять одно из двух (а не из многих) решений. Классификацию, отвечающую специфике альтернативных задач, можно осуществить на основе последовательной процедуры Вальда. Эта процедура позволяет использовать большое число характеристик (признаков), является компактной, легко воспринимается, и, что, пожалуй, особенно важно, техника вычислений при этом не требует применения ЭВМ. Более того, если, например, процедура дискриминантного анализа предполагает, что распределение каждого из признаков подчинено нормальному закону, то в процедуре Вальда могут быть использованы признаки, характеризующиеся различной формой распределения. В этом случае говорят, что процедура свободна от вида распределения признаков; следовательно, она не требует вычисления таких трудоемких парамет-

ров, характеризующих распределение, как математическое ожидание признака, его дисперсия и т. п.

При решении задачи мы придерживались не строгой концепции последовательного анализа Вальда, а лишь его принципов. Связано это со следующим обстоятельством. Если говорить строго, то последовательный анализ Вальда построен для случая, когда информация об объекте накапливается поэтапно. Обработав на каждом шаге поступившую «порцию» информации, можно принять то или иное решение из двух альтернативных либо, если информации оказалось недостаточно, продолжить ее накопление и затем вновь обработать. При таком подходе, как показал А. Вальд, для принятия решения, к какому классу отнести объект (с определенным уровнем надежности), информации требуется значительно меньше, чем при других подходах.

Однако при геолого-геофизических исследованиях поэтапная стратегия не всегда выполнима. Действительно, после геофизических исследований скважины интерпретатор располагает всей возможной косвенной информацией о свойствах объекта (например, пласта) и именно на основе этой информации ему необходимо принять то или иное диагностическое решение. Казалось бы, если поэтапного (последовательного) поступления информации нет, то тем самым одна из основных особенностей анализа Вальда использована быть не может и, следовательно, в такой ситуации мы не должны прибегать к процедуре последовательного анализа.

Но такой вывод, вообще говоря, слишком категоричен. Наиболее привлекательная особенность анализа Вальда заключается в том, что его процедура позволяет построить классификационное правило, разделяющее пространство наблюдений $\mathbf{X}$ на два подпространства: $\mathbf{X}_1$ и $\mathbf{X}_0$. Иными словами, если в результате измерений получаем некоторое значение x, попадающее в подпространство $\mathbf{X}_0$, то принимаем гипотезу H_0 (пласт водонасыщен); если же попадаем в подпространство $\mathbf{X}_1$, то принимаем альтернативную гипотезу H_1 (пласт газонасыщен).

Для разделения пространства $\mathbf{X}$ на два непересекающихся подпространства Дж. Нейман и Е. Пирсон сформулировали принципы, которые легли в основу их теоремы (теорема Неймана — Пирсона). Последняя утверждает, что если A — действительное неотрицательное число, то найдется некоторая критическая область $\mathbf{X}_1(A)$, состоящая из всех значений x, для которых отношение максимального правдоподобия

$$\frac{P(\mathbf{x} \mid \omega = 1)}{P(\mathbf{x} \mid \omega = 0)} = \frac{P_1(\mathbf{x})}{P_0(\mathbf{x})} = \frac{P_1(x_1) P_1(x_2) \ldots P_1(x_m)}{P_0(x_1) P_0(x_2) \ldots P_0(x_m)} > A,$$

где $\mathbf{x} = (x_1, x_2, \ldots, x_m)$ — совокупность величин, образующих выборку объемом N из генеральной совокупности, с плотностью вероятности $P(\mathbf{x})$.

Это отношение определяет наиболее мощный критерий выбора между альтернативными гипотезами H_0 и H_1. Очевидно, что для

двух совпадающих распределений $P(x)$ отношение правдоподобия равно единице. Когда же распределения отличны друг от друга, то выбор числа A, задающего границу, при переходе через которую принимается то или иное диагностическое решение, крайне затруднен. Однако А. Вальд показал, что если $A_0 > P_1(x)/P_0(x)$, то принимается гипотеза H_0; если $P_1(x)/P_0(x) > A_1$, то принимается гипотеза H_1. Числа

$$A_0 = \alpha/(1-\beta); \quad A_1 = (1-\alpha)/\beta,$$

где α и β — ошибки соответственно первого и второго рода: ошибка первого рода — отвергнуть гипотезу H_0, когда на самом деле она верна; ошибка второго рода — принять гипотезу H_0, когда в действительности верна гипотеза H_1.

А. А. Гепкин и Е. В. Гублер [32] обосновали применение анализа Вальда в том случае, когда используется отношение правдоподобия градаций распределений разных признаков, соответствующий комплекс которых характеризует два альтернативных класса объектов, подлежащих классификации. Если, например,

$$\frac{P_1(x_{11})}{P_0(x_{11})} \cdots \frac{P_1(x_{23})}{P_0(x_{23})} \cdots \frac{P_1(x_{35})}{P_0(x_{35})} \cdots \frac{P_1(x_{ij})}{P_0(x_{ij})} > A_1,$$

то принимается гипотеза H_1; если

$$\frac{P_1(x_{11})}{P_0(x_{11})} \cdots \frac{P_1(x_{23})}{P_0(x_{23})} \cdots \frac{P_1(x_{35})}{P_0(x_{35})} \cdots \frac{P_1(x_{ij})}{P_0(x_{ij})} < A_0,$$

то принимается гипотеза H_0; здесь i — номер признака; j — номер градации функции распределения признака.

Поясним смысл этих неравенств. Величина $P_1(x_{ij})/P_0(x_{ij})$ — отношение вероятностей для фиксированной градации, где $P_1(x_{ij})$ — вероятность в j-й градации распределения признака x_i, характеризующего объект первого класса; $P_0(x_{ij})$ — то же для объекта второго класса. Таким образом, имея несколько признаков (комплекс характеристик), описывающих объекты двух разных классов, подлежащих распознаванию, и, следовательно, эмпирические формы распределений этих признаков, можно вычислить отношение правдоподобия для разных признаков и разных градаций и при этом выработать классификационное правило. Достигая границы A_0 или A_1, мы тем самым принимаем то или иное решение относительно исследуемого объекта. Если при использовании всей имеющейся информации ни одно из двух неравенств не выполняется, то делается вывод, что для принятия решения этой информации недостаточно (неопределенный ответ).

Для удобства обсчета фактического материала приведенные неравенства целесообразно прологарифмировать; более того, чтобы логарифмы представляли собой целые числа, их следует умножить на коэффициент 10. Величину, которую при этом получают,

называют диагностическим коэффициентом (ДК). Например, коэффициент i-го признака x_i в его j-й градации равен:

$$\text{ДК}(x_{ij}) = 10 \lg [P_1(x_{ij})/P_0(x_{ij})].$$

Соответственно и диагностические пороги A_0 и A_1 должны быть заменены их логарифмами, умноженными на коэффициент 10. Например, при заданных ошибках первого ($\alpha = 0{,}10$) и второго ($\beta = 0{,}05$) рода величины A_0 и A_1 (в дальнейшем именуемые кратко: пороги A_0 и A_1) примут следующие значения:

$$A_0 = 10 \lg \frac{\alpha}{1-\beta} = 10 \lg \frac{0.1}{1-0{,}05} = 10 \lg \frac{0.1}{0{,}95} = 10(-0{,}978) =$$

$$= -9{,}78 \approx -10;$$

$$A_1 = 10 \lg \frac{1-\alpha}{\beta} = 10 \lg \frac{1-0.1}{0{,}05} = 10 \lg \frac{0{,}9}{0{,}05} = 10 \cdot 1{,}255 = 12{,}55 \approx 13.$$

Имея диагностические коэффициенты, распознавание сводим к суммированию этих коэффициентов. Если при этом достигается порог A_0 или A_1, то соответственно принимается гипотеза H_0 или H_1; если ни один из порогов не достигнут, то выносится неопределенный ответ, т. е. имеющейся информации недостаточно для принятия решения.

При использовании изложенной диагностической процедуры (впрочем, как и любой процедуры распознавания) большое значение имеет информативность признаков (геофизических показателей), характеризующих объект классификации. Под дифференциальной информативностью признака понимают степень различия его распределений. Очевидно, чем сильнее различаются распределения одного и того же признака, характеризующего объекты разных классов, тем больше в них заложено информации, позволяющей различать эти объекты.

Наиболее ценной является мера информативности Кульбака

$$I(x_i) = \sum_j [P_1(x_{ij}) - P_0(x_{ij})] \lg [P_1(x_{ij})/P_0(x_{ij})],$$

как раз и устанавливающая степень различия между распределениями. С учетом диагностического коэффициента эту формулу можно переписать иначе:

$$I(x_i) = \sum_j \text{ДК}(x_{ij}) [P_1(x_{ij}) - P_0(x_{ij})].$$

Если подсчитанная информативность больше единицы, то признак информативен и его можно включать в расчеты [32].

Если подходить более строго, то вопрос об отборе наиболее информативных (дискриминантных) признаков оказывается намного сложнее, чем это может показаться с первого взгляда. Увеличение числа признаков приводит к росту объема вычисли-

тельной работы; кроме того (что, пожалуй, самое главное), включение малоинформативных признаков в анализ распознавания часто вносит «шумовые эффекты», которые серьезно могут осложнить этот процесс. Кардинального решения задачи отбора информативных признаков еще нет.

Кроме критерия Кульбака рекомендуется пользоваться некоторыми непараметрическими статистиками [32, 80]. При их применении не требуется разбивать совокупность значений признака на диапазоны, что, как известно, достаточно сильно влияет на конечные результаты. Кстати, последний недостаток присущ и мере Кульбака. Кроме меры информативности признаков можно ввести меру информативности градаций (диапазонов) признака. Так, согласно формуле Кульбака информативность диапазона j признака i

$$I(x_{ij}) = \text{ДК}(x_{ij})\,[P_1(x_{ij}) - P_0(x_{ij})].$$

Поставленная задача диагностики пластов решалась по данным геофизических исследований скважин, которые характеризуют ряд водоносных (40 объектов) и продуктивных (34 объекта) пластов (результаты собраны и переданы авторам Е. Г. Бро). Пласты относятся к отложениям мелового возраста нескольких разведочных площадей Енисей-Хатангской нефтегазоносной области; глубина их залегания 810—2660 м.

Комплекс исходных признаков, по которым проводилась классификация, состоит из значений кажущегося удельного электрического сопротивления $\rho_к$, измеренного градиент-зондами разной длины (0,5; 1,05; 2,25 м); истинного сопротивления пласта $\rho_п$, определенного по палеткам БЭЗ; собственного потенциала $\Delta U_{ПС}$, отношения $\Delta U_{ПС}/\rho_п$; мощности пласта $h_п$.

Результаты вычисления диагностических коэффициентов и информативности диапазонов признака $\rho_п$ приведены в табл. III.1. Чтобы свести

Таблица III.1

Диагностические коэффициенты и информативность диапазонов (распределение признака $\rho_п$)

Диапазоны (границы изменения признака)	Число наблюдений, попадающих в диапазон	Частость (вероятность), %	Сглаженная частость, %	Диагностический коэффициент	Информативность
2,5—15	1/25	2,9/62,5	6,4/50,5	—9	3,9
16—30	6/9	17,6/22,5	14,0/29,5	—3	0,5
31—45	9/6	26,4/15,0	25,7/16,7	2	0,2
46—60	6/0	17,6/0	17,0/2,2	9	1,3
61—75	3/0	8,8/0	12,6/1,5	9	1,0
76—90	4/0	11,7/0	10,2/0,5	13	1,3
91—105	2/0	5,8/0	7,3/0	—	—
106—120	3/0	8,8/0	6,0/0	—	—
Всего	34/40	99,6/100	99,2/100,9	—	8,2

Примечание. Числитель — для газа, знаменатель — для воды.

Таблица III.2

Информативность признаков и диагностические коэффициенты (газо- и водонасыщенные пласты месторождений Енисей-Хатангского прогиба)

Признак	Диапазон изменения признака	Диагностический коэффициент	Информативность
Сопротивление пласта $\rho_п$	2,5—15	—9	8,2
	16—30	—3	
	31—45	2	
	46—60	9	
	61—75	9	
	76—90	13	
	91—105	0	
	106—120	0	
Собственный потенциал пласта $\Delta U_{ПС}$	<25	11	3,4
	26—45	4	
	46—65	—2	
	66—85	—2	
	86—105	—3	
	106—125	—3	
	126—145	—1	
	146—165	—4	
Кажущееся сопротивление пласта:			
$\rho_{0,5}$	<15	—7	2,3
	16—30	—7	
	31—45	2	
	46—60	0	
	61—75	1	
	76—90	2	
	91—105	3	
$\rho_{2,25}$	<20	—2	1,3
	21—40	0	
	41—60	—1	
	61—80	—2	
	81—100	0	
	101—120	3	
	121—140	6	
	141—160	5	
	161—180	3	
	181—200	0	
$\rho_{1,05}$	<25	—3	0,78
	26—50	—1	
	51—75	0	
	76—100	0	
	101—125	2	
	126—150	2	
	151—175	2	
	176—200	2	
	201—220	1	
	221—250	1	

Продолжение табл. III.2

Признак	Диапазон изменения признака	Диагностический коэффициент	Информативность
Отношение $\Delta U_{ПС}/\rho_п$	<2	4	2,2
	2,1—4	2	
	4,1—6	—2	
	6,1—8	—3	
	8,1—10	—6	
	10,1—12	0	
	12,1—14	0	
	14,1—16	0	
	>16		
Мощность пласта $h_п$	4,1—6	2	1,9
	6,1—8	1	
	8,1—10	2	
	10,1—12	1	
	12,1—14	0	
	14,1—16	—3	
	16,1—18	—4	
	18,1—20	—7	
	>20	—11	

к минимуму влияние выбора границ градаций на результаты распознавания, проводилось сглаживание частостей (вероятностей) методами взвешенного скользящего среднего [81]. Для этого использовались формулы, учитывающие частости анализируемого признака в четырех соседних диапазонах:

$$\bar{x}_3 = (x_1 + 2x_2 + 4x_3 + 2x_4 + x_5) : 10;$$

$$\bar{x}_4 = (x_2 + 2x_3 + 4x_4 + 2x_5 + x_6) : 10.$$

Для крайних диапазонов В. Ю. Урбах рекомендует следующую формулу:

$$\bar{x}_1 = (7x_1 + 5x_2 - x_3 - x_4) : 10;$$

$$\bar{x}_2 = (3x_1 + 5x_2 + x_3 + x_4) : 10.$$

Полученные диагностические коэффициенты для j-х диапазонов i-х признаков представлены в табл. III.2; в этой же таблице приведена информативность признаков по Кульбаку.

Как и следовало ожидать, наиболее дискриминантным признаком является истинное удельное сопротивление пласта $\rho_п$, полученное по данным БКЗ; наименее дискриминантным — кажущееся сопротивление, измеренное градиент-зондом, длина которого 1,5 м; этот признак в дискриминантный анализ следует включать с осторожностью. В процессе распознавания объектов (пластов) целесообразно сначала использовать признаки, значение информативности которых наивысшее.

Распознавание пластов осуществляется на основе сводной таблицы. Делается это так. Допустим, что поступившие данные, характеризующие пласт на глубине 830 м (скв. 106, Мессояхская площадь), имеют следующие значения признаков: $h = 7{,}2$ м, $\Delta U_{ПС} = 30$ мВ, $\rho_п = 32$ Ом·м, $\rho_{0{,}5} = 43$,

$\rho_{1,05} = 52$, $\rho_{2,25} = 42$ Ом·м, $\Delta U_{ПС}/\rho_п = 0{,}94$. Из таблицы находим, что диагностические коэффициенты для данных значений признаков соответственно равны: 1; 4; 2; 2; 0; — 1; 4; алгебраическая сумма коэффициентов составляет 12.

Рассчитаем пороговые уровни коэффициентов A_0 и A_1. Если принять, что ошибки $\beta = 0{,}1$, $\alpha = 0{,}05$, то $A_0 = -12{,}5$, $A_1 = +10$ (табл. III.3). При этих порогах среди диагнозов A_0 и A_1 должно быть не более 5% ошибочных решений. Полученная нами сумма диагностических коэффициентов превышает значение порога A_1, следовательно, пласт является газонасыщенным (что подтверждено испытанием).

Таблица III.3

Пороговые суммы диагностических коэффициентов A_1 и A_0 при разном допустимом проценте ошибок первого α и второго β рода

β, %	α, %							
	20	10	5	2	1	0,5	0,2	0,1
20	+6 / −6	+6,5 / −9	+7 / −12	+7 / −16	+7 / −19	+7 / −22	+7 / −26	+17 / −29
10	−9 / −6,5	+9,5 / −9,5	+10 / −12,5	+10 / −16,5	+10 / −19,5	+10 / −22,5	+10 / −26,5	+10 / −29,5
5	+12 / −7	+12,5 / −10	+13 / −13	+13 / −17	+13 / −20	+13 / −23	+13 / −27	+13 / −30
2	+16 / −7	+16,5 / −10	+17 / −13	+17 / −17	+17 / −20	+17 / −23	+17 / −27	+17 / −30
1	+19 / −7	+19,5 / −10	+20 / −13	+20 / −17	+20 / −20	+20 / −23	+20 / −27	+20 / −30
0,5	+22 / −7	+22,5 / −10	+23 / −13	+23 / −17	+23 / −20	+23 / −23	+23 / −27	+23 / −30
0,2	+26 / −7	+26,5 / −10	+27 / −13	+27 / −17	+27 / −20	+27 / −23	+27 / −27	+27 / −30
0,1	+29 / −7	+29,5 / −10	+30 / −13	+30 / −17	+30 / −20	+30 / −23	+30 / −27	+30 / −30

Примечание. Числитель — для коэффициента A_1, знаменатель — для A_0.

ДИСКРИМИНАНТНЫЙ АНАЛИЗ

Рассмотрим классификацию описанных выше объектов с помощью линейного дискриминантного анализа [2, 82]. Мы уже приводили выражение для линейной дискриминантной функции $F(\mathbf{x})$, в данном примере положим $\mathbf{\Gamma}_1 = \mathbf{\Gamma}_2$, тогда $F(\mathbf{x}) = \boldsymbol{\beta}'\mathbf{x}$. Эта функция максимизирует нормированное расстояние между центрами преобразованных совокупностей:

$$\Delta^2 = (\mu_1 - \mu_2)^2/\sigma^2.$$

Это расстояние максимально, если β есть решение системы уравнений $\Gamma\beta = \mu_1 - \mu_2$, где Γ — обобщенная ковариационная матрица, $\Gamma = (\Gamma_1 + \Gamma_2)/2$. Отсюда следует, что

$$F(x) = (\mu_1 - \mu_2)' \Gamma^{-1} x.$$

Квадратичная дискриминантная функция

$$F(x) = (x - \mu_2)' \Gamma_2^{-1} (x - \mu_2) - (x - \mu_1)' \Gamma_1^{-1} (x - \mu_1)$$

переходит в данную линейную функцию при $\Gamma_1 = \Gamma_2$.

Если распределения признаков в классах являются нормальными, то линейная и квадратичная дискриминантные функции дают оптимальное разделение соответственно при $\Gamma_1 = \Gamma_2$ и $\Gamma_1 \neq \Gamma_2$. Преобразование $F(x) = \beta' x$ превращает оба m-мерных распределения в одномерные, которые в той или иной степени перекрываются. Поэтому требуется выбрать некоторое значение $F_0(x)$, разделяющее оба класса. Критерием выбора служит средняя квадратическая погрешность $P(\alpha)$, которая в обоих классах должна быть минимальной. Для ее отыскания используется байесовский подход, который требует знания априорных вероятностей P_1 и P_2 обоих классов. Тогда $P(\alpha) = P_1 P(\alpha_1) + P_2 P(\alpha_2)$. Если $P_1 = P_2 = 0.5$, то можно показать, что

$$F_0(x) = (1/2)(\mu_1 - \mu_2)' \beta.$$

После выбора $F_0(x)$ отнесение объекта к тому или иному классу определяется знаком величины

$$F(x) - F_0(x) = \beta' x - F_0(x).$$

Это уравнение описывает $(m-1)$-мерную гиперплоскость в m-мерном гиперпространстве. Это и есть разделяющая гиперплоскость; знак величины $\beta' x - F_0(x)$ указывает, по какую сторону от нее лежит данный объект.

Поскольку после преобразования $F(x) = \beta' x$ полученные распределения обычно в той или иной степени перекрываются, некоторые объекты будут ошибочно отнесены не в свой класс. При наилучшем выборе решающего правила вероятность ошибочной классификации объекта

$$P(\alpha) = \Phi(-\Delta/2),$$

где Φ — интегральная функция нормального распределения; $\Delta^2 = (\mu_1 - \mu_2)' \Gamma^{-1} (\mu_1 - \mu_2)$ — расстояние Махаланобиса.

Ошибка в классификации может возникнуть по двум причинам: 1) использовано не оптимальное решающее правило; 2) дискриминация проводится по неинформативным признакам. Решающие правила уже обсуждались; что же касается признаков, то здесь надо отметить следующие моменты.

Если классификация проводится не по выборочным оценкам, то с ростом числа признаков доля ошибочных отнесений умень-

шается. Однако при этом увеличивается объем вычислительной работы, поэтому даже по этой причине встает вопрос о выборе наиболее дискриминантных признаков. Решение задачи основывается на том, что отдельные признаки вносят неодинаковый вклад в расстояние Δ^2. Поэтому желательно отбросить те признаки, которые не оказывают существенного влияния на качество дискриминации. Для их поиска предложено достаточное число методов, вскрывающих всю сложность этого вопроса.

Если параметры распределений оцениваются по выборкам (что всегда имеет место в геологических и геофизических исследованиях), то вопрос об отбрасывании малодискриминантных признаков еще более осложняется. Попытки решения этой задачи сводились, в частности, к тому, чтобы не принимать во внимание те признаки, для которых разности $\bar{x}_1 - \bar{x}_2$ незначимы. Однако невозможность оценить дискриминантность признака по значению $\bar{x}_1 - \bar{x}_2$ была неоднократно доказана [82]. Поэтому использование для отбора признаков t-критерия следует считать ошибочным. Этот вопрос еще далек от своего удовлетворительного разрешения.

Мы интересуемся различением по комплексу геофизических признаков газо- и водоносных пластов. В качестве таких признаков, как уже указывалось, выбраны: собственные потенциалы $\Delta U_{ПС}$, истинное сопротивление

Таблица III.4

Результаты дискриминантного анализа

$$F(\mathbf{x}) = 0{,}033\,\Delta U_{ПС} - 0{,}1578\,\rho_п + 0{,}2405\,(\Delta U_{ПС}/\rho_п) - 0{,}1350\rho_{0,5} + 0{,}0668\,\rho_{1,05} + 0{,}0221\,\rho_{2,25};$$

$$\Delta^2 = 2{,}7317;$$

$$F_0(\mathbf{x}) = -3{,}9791$$

Дискриминант		Результаты распознавания		Дискриминант		Результаты распознавания	
Газ	Вода	Газ	Вода	Газ	Вода	Газ	Вода
—6,123	2,1604	0	1	—9,7885	0,5556	0	1
—7,552	0,8993	0	1	—11,0762	0,1020	0	1
—6,420	—0,7235	0	1	—12,3913	--3,4637	0	1
—5,9013	—1,4316	0	1	—5,1137	--1,2591	0	1
—6.6681	1,1146	0	1	—9,9704	2,0900	0	1
—5,9430	—0,0281	0	1	—10,9104	3,9320	0	1
—7,5273	—0,1514	0	1	—12,4501	--0,4496	0	1
—5,5142	—1,1628	0	1	—8,8435	--3,1602	0	1
—6,3643	2,0692	0	1	—6,1766	--1,4891	0	1
—12.8323	2,6840	0	1	—12,6574	--1,1844	0	1
—11,6751	3,9286	0	1	—6,7161	--3,7416	0	1
—8,4134	0,7693	0	1	—7,2037	--2,0744	0	1
—3,8396	2,4067	1	1		--0,3422		1
—7,6071	2,0768	0	1		--5,4293		0
—1,5839	0,7879	1	1		--4,1612		0
—2,4241	4.9026	1	1		--4,3827		0
—8,3982	—0,4802	0	1		0,4833		1
—4,2523	1,6432	0	1		--3,1941		1

пласта $\rho_п$, оцененное по БЭЗ; отношение $\Delta U_{ПС}/\rho_п$; кажущееся сопротивление $\rho_к$ при трех разных размерах зонда. Число газонасыщенных пластов 30, а водонасыщенных 36. Число m признаков (6), выбранных для распознавания, намного меньше объема n анализируемых совокупностей (30 и 36), что позволяет с уверенностью интерпретировать результаты дискриминантного анализа. Если бы число признаков было сравнимо с объемом выборки, сказывалось бы влияние эффектов «большой размерности», которое учесть невозможно. Поэтому разумное применение данного метода возможно лишь при $m \ll n$.

Результаты дискриминантного анализа (табл. III.4) показывают, что по значениям дискриминантной функции можно вполне удовлетворительно классифицировать пласты по шести показателям, полученным на основе геофизических методов исследования скважин, и проводить распознавание этих пластов. Из 30 газонасыщенных пластов лишь три отнесены не к своему классу, т. е. эмпирическая вероятность ошибки $P(\alpha_1) = 0{,}1000$; водонасыщенные пласты распознаются примерно так же: из 36 пластов три классифицированы неправильно, $P(\alpha_2) = 0{,}0833$.

Чтобы распознать пласт, не вошедший в обучающую последовательность, надо по значениям его характеристик $\Delta U_{ПС}$, $\rho_п$, $\Delta U_{ПС}/\rho_п$, $\rho_{0.5}$, $\rho_{1.05}$ и $\rho_{2.25}$ вычислить дискриминант $F(\mathbf{x})$ и сравнить его с граничным значением $F_0(\mathbf{x}) = -3{,}9791$. Если $F(\mathbf{x}) < F_0(\mathbf{x})$, то пласт газоносный; если $F(\mathbf{x}) > F_0(\mathbf{x})$, то пласт водонасыщенный. Вычисления, проведенные для того же пласта, который рассматривался выше (скв. 106, Мессояхская площадь, глубина 830 м), показали, что и по данным дискриминантного анализа он распознается как газоносный.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Итак, мы показали, что диагностика пластов по комплексу геофизических данных, как и любая другая классификация, может быть осуществлена несколькими методами. Естественно возникает вопрос, какой из алгоритмов является наиболее рациональным. Полученные результаты позволяют утверждать: если некоторая задача может быть решена с помощью одного классификационного правила, то близкое качество решения и аналогичные ответы дает и другое правило. Вероятно, при изменении вида распределения, вида функции потерь, решающего правила имеет место устойчивость результатов классификации, т. е. малые изменения этих факторов вызывают тоже малые сдвиги в конечных результатах.

Следует подчеркнуть, что любое решающее правило, конечно, если оно выбрано на разумной статистической основе, справляется с классификацией цифровой информации гораздо лучше, чем с классификацией объектов повседневной жизни. Действительно, диагностика геометрических фигур, распознавание музыкальных произведений или рисунка танца любым человеком, даже неспециалистом в данной области, выполняется на глаз или на слух достаточно уверенно и, конечно, без привлечения каких-либо формальных процедур. Но визуально расклассифицировать большие массивы цифровой информации, кодирующей геолого-геофизические показатели газо- и водонасыщенных пластов, — задача крайне сложная даже для специалиста.

Вероятно, умение распознавать зрительные и акустические сигналы выработалось у человека в процессе эволюции и гене-

тически закрепилось. Навыки же в классификации цифровых объектов человек начал приобретать в своей практической деятельности сравнительно недавно. И не случайно поэтому опытный специалист быстрее и правильнее может решать задачи по классификации производственных объектов, чем новичок в этой области. Чтобы уменьшить срок приобретения опыта и увеличить число интерпретаторов, способных правильно и быстро решать определенный круг диагностических задач, необходимо иметь соответствующий инструмент для решения указанных задач. Таким инструментом и могут явиться алгоритмы распознавания образов, классифицирующие абстрактные цифровые массивы. Действительно, в некоторых ситуациях подобные алгоритмы наряду с ЭВМ при наличии большого числа признаков позволяют решать такие задачи, которые неподвластны специалистам даже с большим опытом и острой интуицией.

Отметим одно важное обстоятельство. Все изложенные методы являются состоятельными лишь для выборок достаточно большого объема, т. е. с ростом числа наблюдений они позволяют как угодно близко, асимптотически, подойти к оптимальному решению. Однако ясно, что на практике объем выборки всегда ограничен. Поэтому возникает задача поиска методов минимизации риска на конечных выборках, т. е. необходимы не асимптотически-оптимальные, а конечно-оптимальные алгоритмы, приводящие к наилучшему обучению распознаванию образов на выборках фиксированного объема. Определенные достижения в этом направлении уже имеются. Отличие конечно-оптимальной теории от асимптотически-оптимальной сводится к следующему [14]. Качество алгоритма в первом случае естественно определять качеством решающего правила, выбранного по обучающей последовательности. Но поскольку обучающая последовательность составляется случайно, то и качество $P(\alpha)$ будет величиной случайной. Поэтому при решении задачи обучения распознаванию на выборках конечного размера лучшим следует считать тот алгоритм, для которого математическое ожидание величины $P(\alpha)$ меньше. Поиск конечно-оптимальных алгоритмов осуществляется так называемым методом упорядоченной минимизации риска. Схема реализации этого метода такова.

Вводятся априорные сведения о тех задачах, которые предстоит решать. Пусть Q — множество всех возможных задач. Их можно представить в виде системы вложенных подмножеств

$$Q_1 \subset Q_2 \subset \ldots \subset Q_N \subset Q$$

с вероятностью появления каждой задачи $P(Q_i)$. При этом по д множества выделяются так, чтобы

$$P(Q_1) < P(Q_2) < \ldots < P(Q_N) < 1.$$

В соответствии с подмножествами задач можно ввести упорядочение в классе решающих правил S, построив соответствующую систему вложенных множеств

$$S_1 \subset S_2 \subset \ldots \subset S_N = S$$

так, чтобы алгоритмы класса S_i применялись к обучающей последовательности, рассчитанной на задачи класса Q_i.

В каждом классе S_i отыскивается правило, минимизирующее эмпирический риск, и вычисляется оценка качества

$$P(i) \leqslant P_{\text{эмп}}(i) + \varepsilon(i).$$

Слагаемое $\varepsilon(i)$ задает доверительный интервал для класса S_i. Это слагаемое монотонно растет с ростом номера i, в то время как величина $P_{\text{эмп}}(i)$ не возрастает, поскольку эмпирически оптимальное правило из предыдущего класса содержится во всех последующих.

Сравнивая полученные значения $P(i)$, выбираем тот класс К и то правило $F(x, К)$, которые соответствуют минимальному значению $P(i)$. Это правило и будет удовлетворительным. Правило, минимизирующее $P(i)$, обеспечивает минимальную гарантированную вероятность ошибочной классификации, равную $P_{\text{эмп}}(К) + \varepsilon(К)$.

Описанная процедура может быть использована для выбора оптимальной совокупности признаков. В этом случае априорно должна быть задана ранжировка не задач, а систем признаков. Как и ранее, класс S линейных решающих правил разбивается на вложенные подклассы S_i, но уже так, чтобы в подкласс S_i попали правила, использующие только первые i признаков, т. е. работающие в подпространстве E_i, а затем отыскивается правило, обеспечивающее минимум $P(i)$ и соответствующий ему класс К. Набор признаков, отвечающий классу К, и будет оптимальным.

Часть вторая

ИЗУЧЕНИЕ РАЗРЕЗОВ ОСАДОЧНЫХ ТОЛЩ НА ОСНОВЕ АНАЛИЗА СТРУКТУРЫ ГЕОЛОГО-ГЕОФИЗИЧЕСКИХ РЕГИСТРОГРАММ

Глава IV

ИЗУЧЕНИЕ ВЗАИМОСВЯЗИ МЕЖДУ ПОСЛЕДОВАТЕЛЬНЫМИ ЗНАЧЕНИЯМИ РЕГИСТРОГРАММЫ

ПРЕДВАРИТЕЛЬНЫЕ ЗАМЕЧАНИЯ И ОПРЕДЕЛЕНИЯ

На протяжении всех последующих глав изучение разрезов осадочных толщ будет проводиться с определенных позиций, обязанных специфике характеризующего их геолого-геофизического материала. Эта специфика требует рассматривать любую кривую, описывающую изменение по разрезу осадочной толщи значений геологических или геофизических переменных, как реализацию некоторой случайной функции. О том, что подобного рода функции являются случайными, свидетельствует следующее. Представим себе, что одни и те же природные условия, определившие состав и строение осадочной толщи, могли повторяться на данном участке несколько раз. Очевидно, что каждый раз образовывались бы несколько разные толщи или, как говорят иначе, реализации этих условий были бы разными. Если бы регистрировалось изменение какой-либо характеристики по разрезу таких гипотетических толщ, то полученные кривые тоже не были бы точной копией друг друга, в чем и проявляется случайный характер исходных кривых (переменных). Элемент случайности в совокупности измеренных реализаций (кривых) может быть различным. Когда случайность полностью отсутствует, мы приходим к вполне детерминированному поведению характеристики в координатах исследуемого пространства (например, разреза осадочной толщи), что в геологических исследованиях вряд ли возможно.

В широком, хотя и не очень точном понимании, случайной функцией $y(h)$ является любая функция, управляемая вероятностными законами. Более строго, случайная функция $y(h)$ — это такая функция, значение которой при любом возможном аргументе h есть случайная величина. Случайную функцию принято называть случайным (или стохастическим) процессом, если h меняется непрерывно (например, непрерывная запись диаграммы каротажа), или случайной последовательностью, если h принимает дискретное множество значений. Такую функцию в наших приложениях будем обычно именовать регистрограммой, но иногда могут быть даны и другие названия: эмпирическая переменная, сигнал, кривая.

Изучение строения разреза осадочной толщи по регистрограмме, вообще говоря, сводится к изучению ее конфигурации. При этом стараются либо обнаружить в этих переменных закономерные (эволюторные) компоненты, связанные с эволюционным развитием общего геологического процесса, либо отделить закономерные компоненты от случайных, флуктуирующих, возмущений, которые всегда сопровождают измеренные геолого-геофизические характеристики и существенно осложняют их истолкование.

При такой постановке задачи сразу возникает несколько вопросов. Каким образом можно определить, что собственно следует понимать под закономерной компонентой, которая отражает часть изменчивости геолого-геофизических данных, описывающих разрез осадочной толщи? Как найти отличительные черты, выработать определенные критерии, позволяющие одну часть информации отнести к закономерной (детерминированной), а другую — к случайной (стохастической), когда оба ее вида объединены и содержатся в исходных результатах наблюдений? В этом случае, как и ранее, целесообразно использовать определенные математические модели.

Однако разделением информации на «полезную» и «мешающую» дело не ограничивается. Хорошо известно, что существует причинная зависимость между определенным состоянием исследуемого геологического объекта, возникновение которого вызвано породоформирующим процессом, и изменчивостью разного рода геологических и, следовательно, зависящих от них геофизических характеристик, описывающих этот объект. Иначе говоря, характер поведения случайной функции, полученной в результате геологических или геофизических наблюдений, или, более строго, структура (конфигурация) этой функции обусловлена некоторым природным механизмом (геологическими причинами) или, что то же самое, работой динамической системы. Поэтому часто интерпретацию «полезной» информации мы будем проводить, чтобы вынести суждение о природном механизме, определившем строение изучаемого разреза.

Таким образом, круг очень важных вопросов, рассматриваемых во второй части книги, выходит за рамки общеизвестных задач. На эти вопросы невозможно ответить, используя только традиционные приемы обработки и интерпретации результатов геофизических исследований скважин. Поэтому привлекается ряд методов теории случайных функций, в частности авторегрессионный, спектральный анализы, разложение Карунена — Лоэва, анализ вскрытия и выделения периодичностей. Все эти методы не являются традиционными. К сожалению, при истолковании исходных наблюдений они применяются исключительно редко, хотя конечные результаты их использования трудно переоценить.

При интерпретации данных геолого-геофизических исследований в разрезе скважин, рассмотренной в предыдущих главах, пласты считались изолированными друг от друга, т. е. не учитывалась взаимосвязь их свойств и характеристик. Однако хорошо известно, что взаимозависимость свойств внутри последовательности пластов, слагающих разрез осадочной толщи, непременно существует. Эта связь косвенно или прямо учитывается исследователями. Не случайно поэтому эмпирический материал по разрезу обычно представляют в виде разного рода кривых распределения измеренных свойств с глубиной. Уже само такое представление позволяет вскрыть характерные особенности разреза и тем самым лучше его изучить.

При описании кривых часто встречаются такие выводы: «пористость пород вверх по разрезу постепенно возрастает» или «изменение кривой ПС по разрезу контролируется изменением пористости и литологического состава пород и носит ритмический характер». В этих фразах подчеркивается общая мысль: значение любого геолого-геофизического признака, характеризующее данную точку разреза, предопределено (в вероятностном смысле) значениями признака в предшествующих точках, т. е. имеется взаимосвязь последующих свойств породы с предшествующими. Очевидно, что эта взаимозависимость обусловливает определенную конфигурацию (структуру) геолого-геофизических переменных, характеризующих исследуемый разрез. В широком смысле можно утверждать: конфигурация кривой обусловлена некоторым природным механизмом, предопределяющим последействие, которое воспроизводит взаимозависимость между последовательностью ординат этой кривой. Очевидно, что было бы разумно при интерпретации разного рода геологических признаков и геофизических диаграмм (регистрограмм), при выделении из них закономерных и случайных составляющих использовать эту внутреннюю взаимозависимость между последовательными ординатами переменных.

Наибольший эффект достигается при обращении к хорошо разработанным подходам, основанным на методах теории случайных функций. Обрабатывая соответствующим образом эмпирическую переменную, характеризующую разрез, можно составить определенное суждение о природном механизме (динамической системе), функционирование которого вызвало наблюдаемую последовательность слоев именно с таким распределением геолого-геофизических свойств, причем функционирование может быть выражено в терминах стохастического дифференциального уравнения. Параметры уравнения и некоторые вспомогательные функции этих параметров дают возможность установить основные и специфические особенности работы природного механизма, вскрыть общие и частные моменты в поведении исследуемых эмпирических переменных и, следовательно, полнее и детальнее изучить разрез.

Авторегрессионный анализ

Предположим, что имеется некоторая динамическая система (процесс), вызывающая направленное изменение какой-либо контролируемой по разрезу характеристики (пористости, глинистости, размера зерен или геофизических данных, зависящих от этих параметров). Если направленного изменения не наблюдается, то характеристики все равно будут иметь некоторый «внутренний» разброс, что фиксируется их хаотическим изменением по разрезу. Это явление следует отнести на счет большого числа разнообразных случайных причин. Такой характер изменения контролируемой переменной хорошо согласуется с поведением так называемой случайной стационарной функции $e(h)$. По причинам, которые будут изложены ниже, случайную функцию $e(h)$ удобно называть «входом». Эта не известная нам функция характеризует устойчивость самой системы, ее «сбои».

В результате функционирования динамической системы возникает наблюдаемое в настоящее время поведение по разрезу какой-либо регистрируемой характеристики. Это поведение описывается функцией $y(h)$, которую удобно называть «выходом». Условно можно представить, что динамическая система преобразует случайную функцию $e(h)$ в некоторую функцию $y(h)$. Именно функция $y(h)$ регистрируется исследователем, и по ней необходимо составить определенное суждение о работе динамической системы.

Предположим, что существует линейное соотношение между «выходом» и «входом», которое можно записать в виде [8, 37]

$$y(h)-\mu=\int_0^{\infty} h(\tau)\, e(h-\tau)\, d\tau. \qquad \text{(IV.1)}$$

Известно [37], что эта формула является решением линейного дифференциального уравнения вида

$$b_m \frac{d^m y}{dh^m} + b_{m-1} \frac{d^{m-1} y}{dh^{m-1}} + \ \ldots\ + b_0 [y(h)-\mu] = e(h), \qquad \text{(IV.2)}$$

которое может достаточно подробно описывать работу динамической системы, преобразующей «входную» функцию $e(h)$.

Дискретным аналогом последнего уравнения служит следующая запись:

$$y_h-\mu=\beta_1 (y_{h-1}-\mu) + \ \ldots\ + \beta_p (y_{h-p}-\mu) + e_h. \qquad \text{(IV.3)}$$

Это уравнение носит название процесса авторегрессии p-го порядка для дискретного случая. Оно связывает последующее значение исходной регистрограммы с p предшествующими.

Так как значения y_h в дискретных точках h известны, то можно провести оценку коэффициентов β. Как и прежде (см. гл. I), в матричных обозначениях

$$\beta = (\mathbf{Y}'\mathbf{Y})^{-1}\mathbf{Y}'\mathbf{y}, \tag{IV.4}$$

где $\mathbf{Y}$ — матрица вида

$$\mathbf{Y} = \begin{pmatrix} y_{p-1}-\mu & y_{p-2}-\mu & \dots & y_1-\mu & y_0-\mu \\ y_p-\mu & y_{p-1}-\mu & \dots & y_2-\mu & y_1-\mu \\ \dots & \dots & \dots & \dots & \dots \\ y_{h-1}-\mu & y_{h-2}-\mu & \dots & y_{h-p+1}-\mu & y_{h-p}-\mu \end{pmatrix};$$

$\mathbf{y}$ — вектор наблюдений с компонентами

$$\mathbf{y}' = [(y_p-\mu),\ (y_{p+1}-\mu),\ \dots,\ (y_h-\mu)].$$

Вектор $\mathbf{y}$ и столбцы матрицы $\mathbf{Y}$ центрированы; если, кроме того, их еще и нормировать, то оценка вектора примет следующий вид:

$$\beta = \mathbf{R}^{-1}\mathbf{r}, \tag{IV.5}$$

где

$$\mathbf{R} = \frac{1}{n\sigma_{y_h}}[\mathbf{Y}'\mathbf{Y}]$$

— автокорреляционная матрица;

$$\sigma^2_{y_h} = \frac{1}{n}\sum_{h=0}^{n}(y_h-\mu)^2$$

— дисперсия эмпирической переменной (мера отклонения ординат функции y_h от ее среднего уровня); $\mathbf{r}$ — вектор, компонентами которого определяются значения автокорреляционной функции.

Оценка автокорреляционной функции дается следующим выражением:

$$r(\tau) = \frac{1}{n-\tau}\sum_{h=0}^{n-\tau}\frac{(y_h-\mu)(y_{h+\tau}-\mu)}{\sigma^2_{y_h}}. \tag{IV.6}$$

Как видим, коэффициенты β_i, определяющие дифференциальное уравнение, полностью зависят от элементов автокорреляционной матрицы и функции автокорреляции. Это же относится и к структуре исходной эмпирической переменной. Таким образом, вид функции автокорреляции может являться источником исходных идей при построении вероятностной модели механизма, порождающего структуру (конфигурацию) исходной эмпирической переменной.

Опыт показывает, что для описания многих достаточно сложных систем можно использовать приведенное уравнение не выше второго порядка [37]. При этом для оценки коэффициентов β_i по-

требуются значения автокорреляционной функции не более чем при двух сдвигах τ. Детальное изучение автокорреляционных функций при достаточно больших сдвигах приводит к иным моделям, анализ которых дается в последующих параграфах.

Рассмотрим более подробно уравнения авторегрессии первого и второго порядка.

Функция $y(h)$, генерируемая уравнением

$$b_1 \frac{dy}{dh} + b_0 [y(h) - \mu] = e(h), \qquad \text{(IV.7)}$$

называется процессом авторегрессии первого порядка или марковским процессом первого порядка. Аналогично функция $y(h)$, порождаемая уравнением

$$b_2 \frac{d^2y}{dh^2} + b_1 \frac{dy}{dh} + b_0 [y(h) - \mu] = e(h), \qquad \text{(IV.8)}$$

называется процессом авторегрессии второго порядка или марковским процессом второго порядка.

Дискретные аналоги уравнений (IV.7) и (IV.8) записываются в виде разностных уравнений:

$$y_h - \mu = \beta_1 (y_{h-1} - \mu) + e_h; \qquad \text{(IV.9)}$$

$$y_h - \mu = \beta_1 (y_{h-1} - \mu) + \beta_2 (y_{h-2} - \mu) + e_h. \qquad \text{(IV.10)}$$

Решение приведенных уравнений относительно y_h имеет вид

$$y_h - \mu = \sum_{\tau=0}^{\infty} \hbar_\tau e_{h-\tau}. \qquad \text{(IV.11)}$$

где $\hbar_\tau$ — функция, по виду которой можно составить определенное суждение о функционировании непосредственно самой динамической системы, выявить основные черты ее работы в условиях отсутствия случайных искажений (помех); иначе говоря, эта функция является характеристикой собственно динамической системы, т. е. соответствующего природного процесса.

Физический смысл функции $\hbar_\tau$ вполне очевиден. Так, если на «вход» динамической системы подать мгновенное возмущение и проследить реакцию системы на «выходе» (т. е. определить отклик системы на единичный импульс), то эта реакция и будет соответствовать функции $\hbar_\tau$.

Для динамической системы, поведение которой описывается уравнением первого порядка, функция $\hbar_\tau$ совпадает с корреляционной функцией процесса y_h и равняется

$$\hbar_\tau = \beta_1^{|\tau|} = r(\tau), \qquad \text{(IV.12)}$$

где $\tau = 0, 1, 2, \ldots$; $r(\tau)$ — автокорреляционная функция.

В моделях авторегрессии значение y_h должно удовлетворять условию стационарности. Это условие будет выполняться лишь

в том случае, когда $|\beta_1| \leqslant 1$. Отсюда видим, что функция отклика $\hbar_\tau$ динамической системы первого порядка при $\beta_1 > 0$ затухает экспоненциально, а при $\beta_1 < 0$ она является знакопеременной функцией, огибающая которой затухает также экспоненциально. Чем ближе $|\beta_1|$ к единице, тем медленнее затухание, т. е. последействие системы распространяется на большее число последующих ординат функции y_h.

В данном и рассмотренном ниже случаях наличие условия стационарности крайне важно. Физически оно означает, что эмпирическая переменная не должна иметь эволюционирующих ординат с ростом аргумента h, т. е. кривая y_h должна осциллировать около постоянного уровня μ, хотя амплитуда некоторых всплесков может быть значительной.

Для динамической системы, поведение которой описывается уравнением второго порядка, функция y_h будет удовлетворять условию стационарности лишь в случае, когда выполнены следующие ограничения:

$$(\beta_1 + \beta_2) < 1;\ (\beta_1 - \beta_2) > -1;\ -1 \leqslant \beta_2 \leqslant 1.$$

Вид функции отклика $\hbar_\tau$ зависит от того, какие корни имеет характеристическое уравнение вида

$$\varepsilon^2 - \beta_1 \varepsilon - \beta_2 = 0.$$

Для действительных корней, т. е. когда $\beta_1^2 \geqslant -4\beta_2$, функция отклика имеет вид

$$\hbar_\tau = (\varepsilon_1^{\tau+1} + \varepsilon_2^{\tau+2})/(\varepsilon_1 - \varepsilon_2), \qquad \text{(IV.13)}$$

где ε_1 и ε_2 — корни уравнения.

Когда корни комплексные, т. е. при $\beta_1^2 < -4\beta_2$, функция

$$\hbar_\tau = \{c^\tau \sin [2\pi f_0 (\tau + 1)]\}/\sin (2\pi f_0). \qquad \text{(IV.14)}$$

Корреляционная функция $r(\tau)$ также зависит от вида корней характеристического уравнения. При действительных корнях

$$r(\tau) = [\varepsilon_1 (1 - \varepsilon_2^2)\, \varepsilon_1^{|\tau|} - \varepsilon_2 (1 - \varepsilon_1^2)\, \varepsilon_2^{|\tau|}]/[(\varepsilon_1 - \varepsilon_2)(1 + \varepsilon_1 \varepsilon_2)]; \qquad \text{(IV.15)}$$

при комплексных

$$r(\tau) = [c^{|\tau|} \cos (2\pi f_0 \tau - \varphi_0)]/\cos \varphi_0. \qquad \text{(IV.16)}$$

В последних выражениях коэффициент затухания c, частота f_0 и фаза φ_0 задаются формулами

$$c^2 = -\beta_2;\ \cos (2\pi f_0) = \beta_1/(2c);$$

$$\operatorname{tg} \varphi_0 = [(1 - c^2)/(1 + c^2)] \operatorname{tg} (2\pi f_0).$$

Эти параметры имеют четкую геологическую трактовку. При этом функция отклика, соответствующая формуле (IV.14), носит колебательный характер. Коэффициент c указывает на то, как

быстро угасает колебание системы в случае вывода ее из состояния равновесия и «сбоя». Вполне понятно, что чем выше значение c, тем медленнее система стабилизируется и тем большее число последующих значений функции y_h испытывает влияние предшествующих «сбоев». В этом случае f_0 есть собственная частота колебания режима системы. Из формулы (IV.11), видно, что для любой динамической системы формирование структуры функции y_h полностью определяется функцией отклика $\hbar_\tau$. Причем чем медленнее происходит затухание $\hbar_\tau$, тем большее число последующих значений функции y_h будет испытывать последействие, т. е. влияние нестабильности динамической системы.

Для построения функции $\hbar_\tau$, решения разностных уравнений необходимо знать коэффициенты β_i и параметр μ. Их оценка может быть проведена непосредственно по имеющимся данным переменной y_h. Метод наименьших квадратов дает следующие формулы:

для уравнения первого порядка

$$\hat{\beta}_1 = r(1);\ \mu = [1/(n+1)] \sum_{h=0}^{n} y_h = \bar{y}; \qquad \text{(IV.17)}$$

для уравнения второго порядка

$$\left.\begin{aligned} &\hat{\beta}_1 = r(1)\,[1 - r(2)]/[1 - r^2(1)];\\ &\hat{\beta}_2 = [r(2) - r^2(1)]/[1 - r^2(1)];\\ &\mu = [1/(n+1)] \sum_{h=0}^{n} y_h = \bar{y}, \end{aligned}\right\} \qquad \text{(IV.18)}$$

где $r(\tau)$ — выборочные оценки первых значений автокорреляционной функции ($\tau = 1;\ 2$).

Если функция $\hbar_\tau$ характеризует влияние «сбоев» системы, определяя вид функции y_h, то коэффициенты β_i показывают, в чем собственно заключается направленное действие системы — смысл ее работы. В данном случае система функционирует по принципу: каждое последующее значение пропорционально предыдущему или двум предыдущим. Коэффициенты $\hat{\beta}_i$ и есть коэффициенты пропорциональности. В этом смысле их можно трактовать как меру проявления марковости первого или второго порядка в исходных случайных функциях.

В уравнениях (IV.9) и (IV.10) положим $y_h - \mu = \tilde{y}_h$ и умножим обе части каждого из них на $\tilde{y}_h$, получим

$$\tilde{y}_h \tilde{y}_h = \beta_1 (\tilde{y}_h \tilde{y}_{h-1}) + e_h \tilde{y}_h$$

для уравнения (IV.9) и

$$\tilde{y}_h \tilde{y}_h = \beta_1 (\tilde{y}_h \tilde{y}_{h-1}) + \beta_2 (\tilde{y}_h \tilde{y}_{h-2}) + e_h \tilde{y}_h$$

для уравнения (IV.10).

Суммируя по h от 0 до n каждое слагаемое и принимая во внимание, что

$$\frac{1}{n+1}\sum_{h=0}^{n}\tilde{y}_h\tilde{y}_h=\sigma^2_{\tilde{y}_h};\quad \frac{1}{n+1}\sum_{h=0}^{n}e_h^2=\sigma^2_{e_h};$$

$$r(1)=\frac{\sum\tilde{y}_h\tilde{y}_{h-1}}{(n+1)\,\sigma^2_{\tilde{y}_h}};\quad r(2)=\frac{\sum\tilde{y}_h\tilde{y}_{h-2}}{(n+1)\,\sigma^2_{\tilde{y}_h}},$$

а ковариации $\sum e_h\tilde{y}_{h-1}$ и $\sum e_h\tilde{y}_{h-2}$ равны нулю в силу независимости (ортогональности) e_h и y_h, получаем

$$\sigma^2_{\tilde{y}_h}=\sigma^2_{e_h}/[1-\beta_1 r(1)] \qquad \text{(IV.19)}$$

для авторегрессионного уравнения первого порядка и

$$\sigma^2_{\tilde{y}_h}=\sigma^2_{e_h}/[1-\beta_1 r(1)-\beta_2 r(2)] \qquad \text{(IV.20)}$$

для авторегрессионного уравнения второго порядка.

Таким образом, имея значения эмпирической переменной, характеризующей разрез, т. е. значения функции на «выходе» динамической системы, можно провести оценку дисперсии «входного» сигнала — случайного процесса e_h, характеризующего стабильность работы уже самой системы, т. е. соответствующего природного процесса. В получении характеристик самого процесса и заключается смысл геологических реконструкций.

Мера изменчивости функции y_h, оцениваемая дисперсией $\sigma^2_{\tilde{y}_h}$, может быть разложена на меру изменчивости функции e_h с дисперсией $\sigma^2_{e_h}$ и остаточную изменчивость:

$$\sigma^2_{\tilde{y}_h}=\sigma^2_{e_h}+\sigma^2_{\text{ост}}. \qquad \text{(IV.21)}$$

Изменчивость $\sigma^2_{\text{ост}}$ является мерой интенсивности работы собственно динамической системы, так как представляет собой «добавку», которая образуется после того, как «входной» сигнал e_h был подвергнут трансформации, связанной с действием динамической системы. Можно также утверждать, что мера $\sigma^2_{\text{ост}}$ показывает, как сильно преобразуется (искажается) после работы динамической системы «входная» функция e_h.

Величины

$$\sigma^2_{e_h}/\sigma^2_{\tilde{y}_h}=1-\beta_1 r(1); \qquad \text{(IV.22)}$$

$$\sigma^2_{e_h}/\sigma^2_{\tilde{y}_h}=1-\beta_1 r(1)-\beta_2 r(2) \qquad \text{(IV.23)}$$

представляют собой меру избыточной корреляции; они показывают, в какой степени дисперсия исходного процесса связана с дисперсией наблюдений при условии, что этот процесс относится к марковскому типу того или иного порядка.

Таким образом, мы опять приходим к разложению общей дисперсии исходной функции в ряд по ее элементарным составляющим. Но путь разложения несколько отличен от ранее использованного, где мы применяли модели дисперсионного и факторного анализов. Если раньше предполагалось, что факторы, вызывающие изменчивость (дисперсию) контролируемых переменных, действуют лишь в отдельно взятых точках и взаимодействия между соседними точками нет, то авторегрессионная модель позволяет сначала вскрыть, а затем учесть это взаимодействие.

Теоретические выкладки свидетельствуют, что по форме автокорреляционной функции, по ее первым последовательным значениям можно судить о конфигурации (структуре) исходной

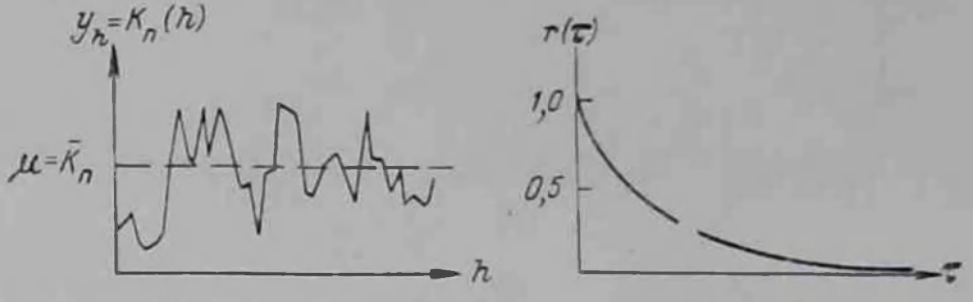


Рис. IV.1. Изменение пористости $K_п$ (h) по разрезу скв. 1 Успинской площади и соответствующая функция автокорреляции.

эмпирической переменной, характеризующей разрез, если ее формирование идет по схеме марковского процесса. Чтобы проиллюстрировать, насколько тонко реагирует форма автокорреляционной функции на изменение конфигурации исходной эмпирической переменной, приведем некоторые примеры процессов авторегрессии первого и второго порядка.

Процесс авторегрессии первого порядка. Кривая изменения пористости, характеризующая фрагмент разреза карбонатных пород каменноугольного возраста (Тимано-Печорская нефтегазоносная область), может быть аппроксимирована процессом авторегрессии первого порядка (рис. IV.1), уравнение которого имеет вид

$$\tilde{y}_h = 0{,}9\tilde{y}_{h-1} + e_h.$$

Соответствующая этому процессу теоретическая функция автокорреляции задается формулой

$$r(\tau) = 0{,}9^{|\tau|}.$$

Надо отметить, что выборочная оценка автокорреляционной функции, рассчитанная непосредственно по наблюденным значениям пористости с использованием формулы (IV.6), не всегда затухает так же быстро, как математическое ожидание оценки, т. е. автокорреляционная функция, рассчитанная по теоретической формуле. Несовпадение есть следствие того, что выборочная

оценка рассчитывается, как правило, по коротким рядам наблюдений (n мало). Отсюда напрашивается вывод, что придавать слишком важное значение видимой конфигурации автокорреляционной функции, особенно при больших τ, не следует. Еще большую осторожность надо проявлять, когда функция автокорреляции рассчитана по коротким рядам наблюдений, здесь необходимо проверять значимость даже первых результатов автокорреляции.

Кривая изменения плотности пород в разрезе меловых отложений (флишевая формация, Западное Предкавказье) может быть удачно аппроксимирована процессом авторегрессии первого

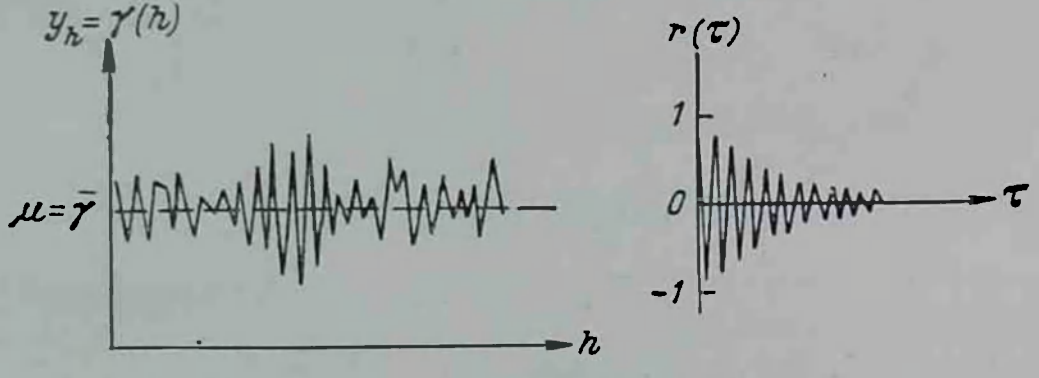


Рис. IV.2. Изменение плотности γ (h) по разрезу меловых отложений флишевой формации Западного Предкавказья и соответствующая функция автокорреляции.

порядка (рис. IV.2), где коэффициент $\hat{\rho}_1$ является отрицательным числом, а именно

$$\tilde{y}_h = -0{,}9\tilde{y}_{h-1} + e_h.$$

Теоретическая функция автокорреляции, рассчитанная по формуле (IV.12), имеет вид

$$r(\tau) = -0{,}9^{|\tau|}.$$

Как видим, функция r (τ) является знакопеременной, свидетельствуя, что исходная кривая имеет сильную осцилляцию определенного характера: пласту с большой плотностью пород предшествует пласт, плотность которого значительно меньше, что в общем-то типично для флишевых образований.

Процесс авторегрессии второго порядка. Ранее приведенные теоретические выкладки свидетельствуют, что при рассмотрении процессов авторегрессии второго порядка следует различать четыре случая, которые определяются знаками коэффициентов разностного уравнения, а именно: 1) β_1 и β_2 положительные, 2) β_1 отрицательный, а β_2 положительный, 3) β_1 положительный, а β_2 отрицательный, 4) β_1 и β_2 отрицательные. Очевидно, что два последних случая соответствуют ситуации, когда корни характеристического уравнения $\varepsilon^2 - \beta_1\varepsilon - \beta_2 = 0$ комплексные.

Чтобы нагляднее представить, каким может быть характер изменения (конфигурация) исходных кривых, которые аппрокси-

мируются процессом авторегрессии второго порядка, приведем сначала пример кривой авторегрессии, полученной искусственно [8]; последующие примеры взяты непосредственно из практики. Представленная на рис. IV.3 кривая порождена процессом авторегрессии второго порядка, уравнение которого имеет вид

$$\tilde{y}_h = 0{,}75\tilde{y}_{h-1} - 0{,}50\tilde{y}_{h-2} + e_h.$$

Следовательно, коэффициенты β_1 и β_2 равны соответственно 0,75 и —0,50. Очевидно, что корни характеристического уравнения $\varepsilon^2 - 0{,}75\,\varepsilon + 0{,}50 = 0$ комплексные и поэтому автокорреляционная функция должна иметь вид затухающей косинусоиды.

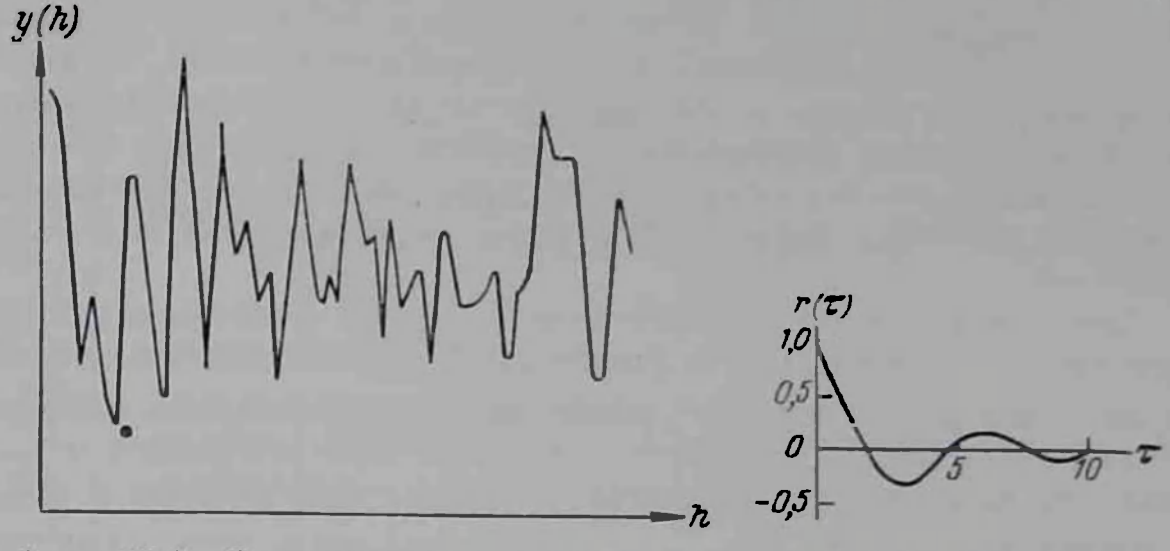


Рис. IV.3. Регистрограмма, порожденная процессом авторегрессии второго порядка с коэффициентами $\beta_1 = 0{,}75$ и $\beta_2 = -0{,}50$, и соответствующая функция автокорреляции [37].

Согласно формуле (IV.16) параметр затухания $c = \sqrt{0{,}50} = 0{,}71$. Присутствие периодической компоненты в функции автокорреляции свидетельствует о том, что и в исходной кривой должна улавливаться периодическая компонента с частотой, рассчитываемой по формуле

$$f_0 = \frac{1}{2\pi}\arccos\frac{\beta_1}{2c} = \frac{1}{2\pi}\arccos 0{,}5303 = \frac{1}{6{,}2}.$$

Таким образом, основной период функции автокорреляции и периодической компоненты равен 6,2.

Обратимся к специальным приемам анализа пористости карбонатных пород (определенной по керну) на основе авторегрессионной модели второго порядка. Такой анализ позволяет не только проиллюстрировать возможности модели авторегрессии, но и показать, как с помощью указанной модели можно осветить некоторые вопросы формирования общей пористости карбонатных пород и как общая пористость может быть расщеплена на составные элементы (допустим, первичную и вторичную пористость). Если исследователь располагает геофизическими показателями, которые уверенно контролируют общую пористость карбонатных

пород, то подобный анализ может быть проведен с использованием этих данных.

В нашем распоряжении имелись значения пористости (определенные по керну), характеризующие разрезы карбонатных пород осинского горизонта нижнего кембрия (Марковское месторождение нефти). Всего было выбрано 20 скважин (рис. IV.4). Равномерно по мощности горизонта в каждой из скважин было получено 20 определений пористости (с учетом интерполяции). Параметры уравнений авторегрессии второго порядка представлены в табл. IV.1 и на рис. IV.4.

По распределению показателя $\sigma^2_{\varepsilon_n}/\sigma^2_{u_n}$ все исследованное поле распадается на две части: в западной избыточная корреляция велика, а в восточной она существенно меньше. Область, охватывающая залежь нефти, выделяется как аномальная. Иными словами, процесс формирования пористости протекал так, что в восточной части и в области залежи корреляция каждого последующего значения пористости с двумя предыдущими выражена наиболее ярко.

Информация о специфическом воздействии функционирующей динамической системы на пористость, об интенсивности ее работы заключена в показателе $\sigma^2_{ост}$. Значения этого показателя в общем возрастают с запада на восток. На фоне этой тенденции в центральной части площади выделяется аномальный участок с пониженными значениями $\sigma^2_{ост}$. В зоне, переходной к этому участку, расположена залежь нефти; она приурочена к локальному максимуму.

Показатель $\sigma^2_{\varepsilon_n}$ отражает изменчивость (дисперсию) «входного» сигнала, обязанную нестабильности режима формирования пористости, «сбоям» динамической системы. Область наиболее стабильного режима имеет вид изогнутой полосы, протягивающейся с запада на восток и охватывающей центральную часть площади.

Динамическая система «работала» на рассматриваемой площади в разных режимах. Так, ее реакция на импульс (функция $\hbar_\tau$) в одних случаях (скважинах) имеет вид затухающей экспоненты, а в других — затухающей косинусоиды (см. табл. IV.1). Таким образом, в одних скважинах изменение пористости носит колебательный характер, а в других характеризуется неуклонным приведением отклонений пористости к постоянному уровню (среднему значению пористости).

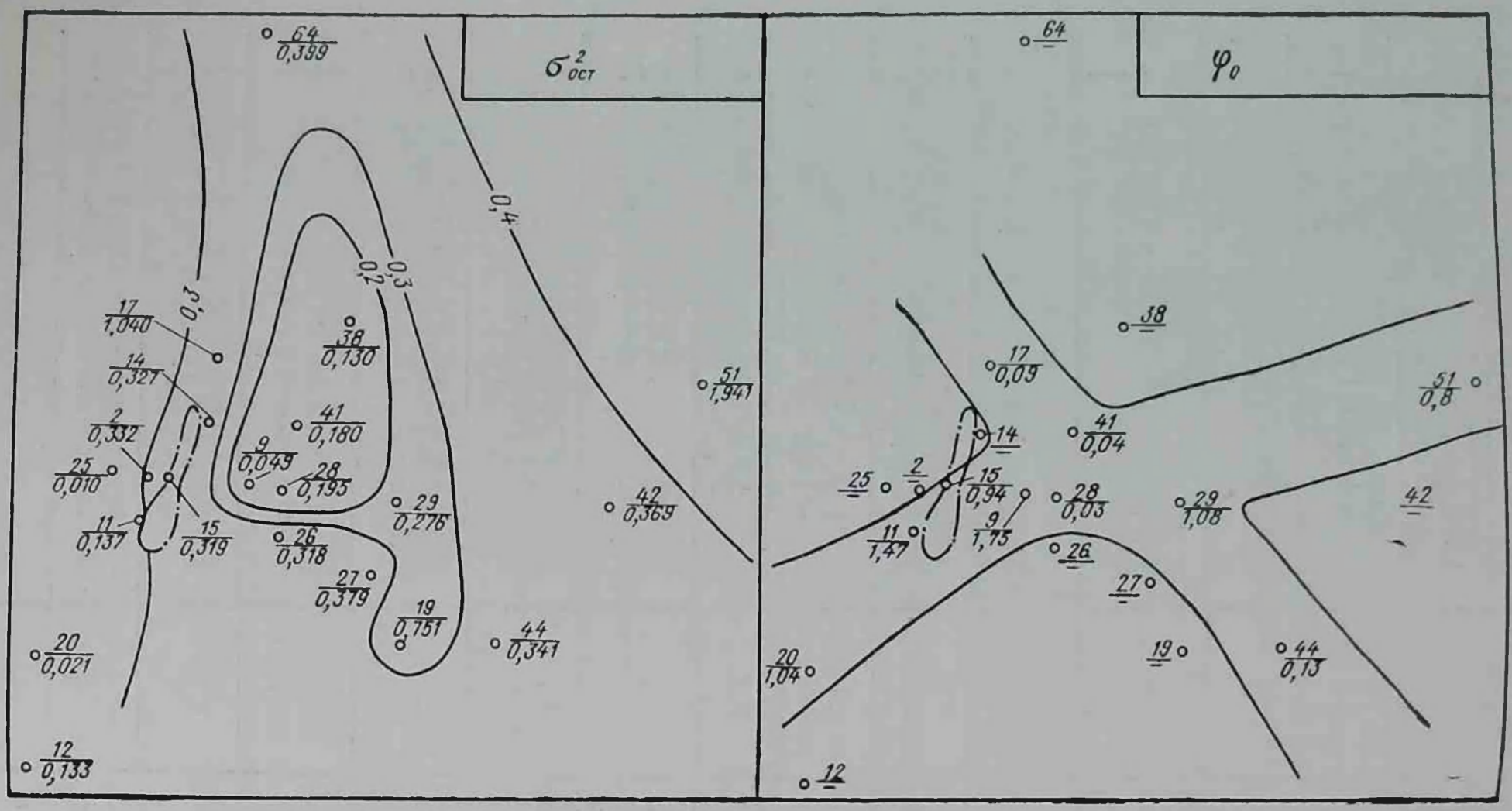
Показатель φ_0 распределен по площади также с явной закономерностью (при этом, естественно, выбраны только те скважины, где функция отклика $\hbar_\tau$ имеет вид затухающей косинусоиды, т. е. корни характеристического уравнения комплексные). Область, охватывающая отмеченные скважины, имеет вид пересекающихся полос, протягивающихся по падению и простиранию моноклинали, причем по падению пород (на северо-запад) почти не

Таблица IV.1

Элементы статистической обработки определений пористости при помощи модели авторегрессии второго порядка (Марковское месторождение)

Скважина	$\bar{y}$	$\sigma^2_{y_h}$	$r(1)$	$r(2)$	β_1	β_2	$\sigma^2_{e_h}$	$\frac{\sigma^2_{e_h}}{\sigma^2_{y_h}}$	$\sigma^2_{ост}$	Форма функции h	ε_1	ε_2	f_0	φ_0	c
2	2,64	1,886	0,417	0,215	0,396	0,049	1,554	0,823	0,332	э *	—0,099	0,495	—	—	—
9	1,73	1,066	0,215	0,030	0,218	—0,017	1,017	0,954	0,049	п **	—	—	0,0917	1,750	0,130
11	1,71	0,986	0,371	0,100	0,387	—0,043	0,849	0,861	0,137	п	—	—	0,0603	1,466	0,207
12	2,01	1,191	0,323	0,136	0,293	0,092	1,058	0,888	0,133	э	—0,19	0,484	—	—	—
14	1,31	0,615	0,725	0,530	0,642	0,116	0,288	0,468	0,327	э	—0,147	0,789	—	—	—
15	2,08	0,558	0,740	0,445	0,907	—0,225	0,239	0,428	0,319	п	—	—	0,472	0,941	0,474
17	3,95	3,233	0,472	—0,063	0,646	—0,368	2,173	0,672	1,040	п	—	—	0,160	0,092	0,606
19	1,93	1,632	0,387	0,131	—0,373	0,037	1,411	0,848	0,251	э	0,102	0,271	—	—	—
20	0,97	0,276	—0,053	—0,270	—0,067	—0,268	0,255	0,923	0,021	п	—	—	0,0897	1,039	0,518
25	1,25	1,402	0,071	0,046	0,068	0,041	1,392	0,992	0,010	э	—0,031	0,099	—	—	—
26	1,75	0,730	0,641	0,532	0,509	0,205	0,412	0,564	0,318	э	—0,0265	0,744	—	—	—
27	2,44	1,413	0,517	0,283	0,506	0,022	1,034	0,922	0,379	э	0,160	0,345	—	—	—
28	0,97	0,390	0,603	0,070	0,882	—0,462	0,195	0,500	0,195	п	—	—	0,137	0,025	0,679
29	4,25	2,992	0,643	0,327	0,737	—0,146	1,716	0,573	1,276	п	—	—	0,043	1,075	0,382
38	1,88	0,314	0,640	0,366	0,687	—0,073	0,184	0,587	0,130	э	—0,093	0,781	—	—	—
41	3,06	0,423	0,566	0,055	0,786	—0,389	0,243	0,577	0,180	п	—	—	0,142	0,040	0,624
42	1,95	0,528	0,835	0,722	0,765	0,088	0,159	0,301	0,369	э	—0,096	0,861	—	—	—
44	2,48	1,587	0,863	0,642	1,211	—0,404	1,246	0,214	0,341	п	—	—	0,049	0,128	0,635
51	5,07	6,284	0,549	0,232	0,604	—0,098	4,343	0,691	1,941	п	—	—	0,043	0,182	0,313
64	3,49	1,664	0,486	0,271	0,463	0,046	1,265	0,762	0,399	э	—0,084	0,547	—	—	—

* Экспоненциальная.
** Периодическая.



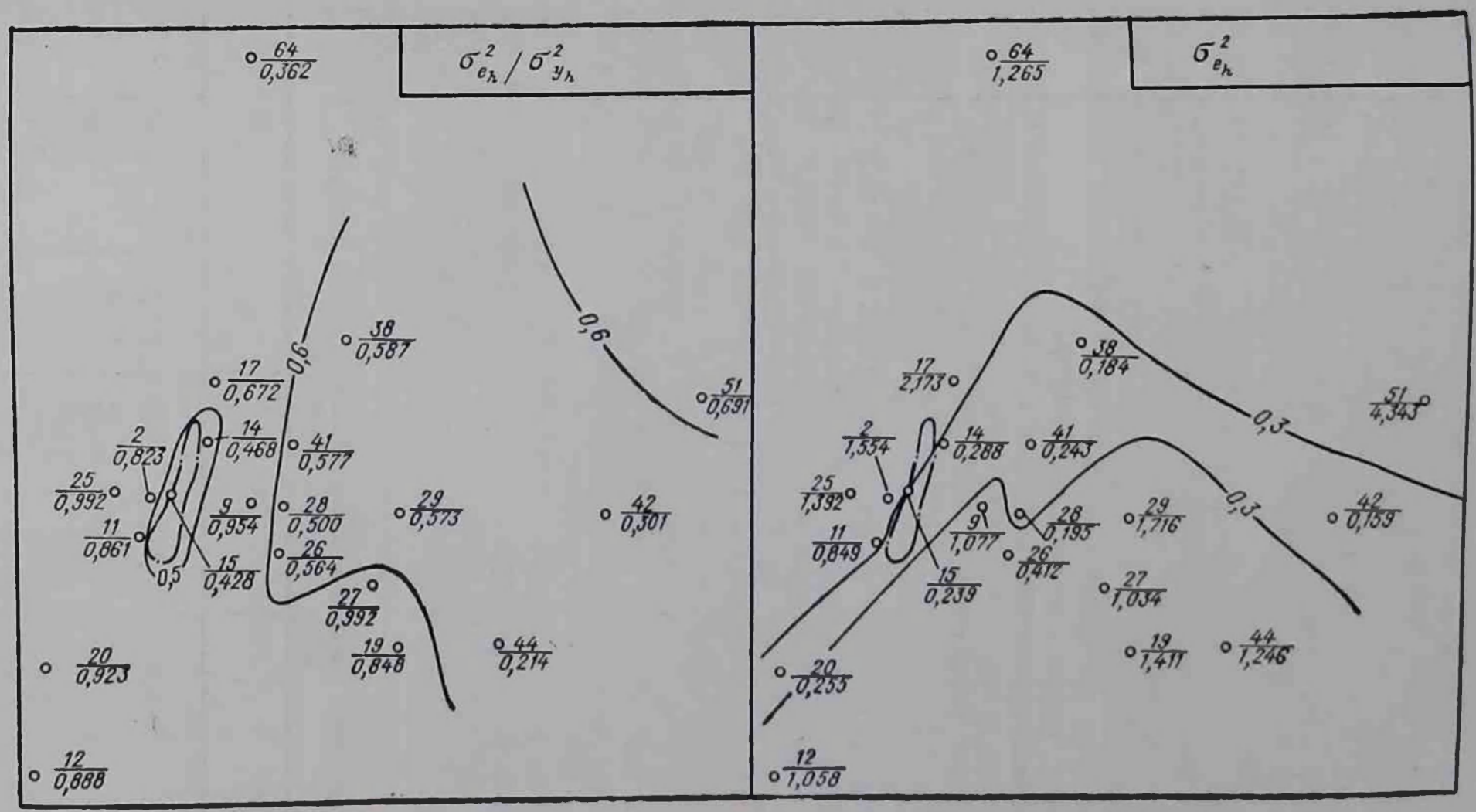


Рис. IV.4. Карты статистических характеристик пористости карбонатных пород осинского горизонта (Марковская площадь).

1 — скважина (в числителе — ее номер, в знаменателе — значение соответствующего показателя); *2* — изолинии показателя; *3* — залежь нефти.

наблюдается запаздывания системы на входную реакцию (значения φ_0 малы), а по простиранию запаздывание достигает максимальных значений. Возможно, что это запаздывание связано с разницей во времени между моментом наступления условий, при которых может идти тот или иной процесс, и моментом реализации этих условий, т. е. началом самого природного процесса.

Таким образом, анализ эмпирических переменных, характеризующих разрезы осадочных толщ, с помощью модели авторегрессии позволяет:

1) судить о наличии взаимосвязи между последовательными значениями пористости и понимать природу этого явления;

2) оценивать различие условий формирования пористости, когда ее распределение по разрезу является относительно стабильным, и в то же время судить о преобладающем механизме, управляющем распределением пористости;

3) оценивать эффект запаздывания в реакции системы и картировать его значение, что важно для понимания связи между распределением пористости и геологическими условиями.

В приведенном примере анализировались значения пористости, определенные по керну. Однако такие определения не всегда легко получить. Поэтому логично обратиться к данным геофизических исследований скважин — непрерывным записям характеристик. Но чтобы использовать геофизические характеристики для подобного анализа, необходимо удостовериться, насколько полно они отражают поведение по разрезу того или иного параметра (например, общей пористости). При исследовании вопросов, связанных с функционированием природного механизма, диктующего наблюдаемую структуру эмпирической переменной, нас интересует не конкретное ее значение в точке разреза, а относительное изменение по разрезу.

Авторегрессионные модели позволяют установить сходство или различие эмпирических переменных разной физической природы на основе критериев, не связанных с обычным коэффициентом корреляции. Такое исследование необходимо, когда нет возмож-

Таблица IV.2

Параметры уравнения авторегрессии второго порядка (скв. 38-Дурбе, Прибалтика)

Эмпирическая переменная	$r(1)$	$r(2)$	β_1	β_2	$\sigma^2_{e_h}/\sigma^2_{y_h}$
Пористость	0,593	0,402	0,546	0,077	0,645
$\rho_{\text{ПЗ}}$	0,608	0,348	0,628	0,132	0,629
ρ_1	0,551	0,399	0,475	0,137	0,683
ρ_2	0,664	0,489	0,607	0,086	0,555
$J_{\text{ГМ}}$	0,649	0,462	0,580	0,106	0,572
$J_{\text{НГМ}}$	0,674	0,450	0,667	0,087	0,554

ности точно привязать один ряд наблюдений к другому ряду, но при этом известно, что оба ряда способны контролировать анализируемую характеристику. Так, по скв. 38-Дурбе (Прибалтика) карбонатный разрез ордовика кроме прямых определений межзерновой пористости был охарактеризован комплексом данных каротажа (КС, НГМ, ГМ). Ставился вопрос, насколько хорошо геофизические данные отражают характер изменения пористости по разрезу. По результатам расчета параметры уравнения авторегрессии второго порядка оказались практически идентичными (табл. IV.2). Это показывает, что исходные регистрограммы имеют одинаковую конфигурацию и, следовательно, для изучения характера распределения пористости по разрезу в данном случае может быть выбрана диаграмма любого из методов.

Разложение Карунена-Лоэва

Взаимосвязь между последовательными значениями регистрограммы можно изучать и другим путем. Суть его заключается в том, что сопоставляются и сравниваются между собой кривые, полученные из регистрограммы последовательным сдвигом ее относительно самой себя на один шаг. В конечном итоге процесс сводится к разложению исходной регистрограммы на отдельные составляющие, содержащиеся в ней с разным весом. При этом никаких предположений о форме составляющих не делается; выдвигается только следующее условие: выделенная составляющая характеризует отдельные черты общего геологического процесса и не зависит от последующей составляющей, которая характеризует другие черты процесса.

А. М. Обуховым [64] предложен способ выделения естественных составляющих, оптимальный в том смысле, что он обеспечивает наибольшую скорость сходимости разложения. При анализе совокупности исходных регистрограмм, которые рассматриваются как различные реализации некоторого процесса, составляющие извлекаются путем выделения собственных векторов ковариационной матрицы, описывающей связь между значениями исходных регистрограмм, т. е. искомые составляющие совпадают с собственными векторами.

Модель Карунена — Лоэва позволяет провести разложение исходной регистрограммы в ряд по естественным ортогональным (независимым) компонентам произвольной формы. Как и при компонентном анализе, выделение ортогональных составляющих осуществляется путем линейного преобразования

$$\mathbf{Z} = \mathbf{Y}\mathbf{U}\Lambda^{1/2},$$

где $\mathbf{U}$ — матрица, составленная из собственных векторов корреляционной матрицы вида $\mathbf{R} = \overline{\mathbf{Y}'\mathbf{Y}}$; $\mathbf{Y}$ — матрица, представляющая собой систему переменных, получаемых из исходной

регистрограммы последовательным сдвигом τ относительно самой себя на один шаг.

При этом предполагается, что чем больше сдвигов проведено, тем большее количество предшествующих ординат переменной оказывается способным влиять на последующую ординату. Иными словами, какое-либо значение эмпирической переменной испытывает последействие ровно стольких значений, сколько выполнено последовательных сдвигов на один шаг.

Полное решение проблемы собственных чисел и собственных векторов корреляционной матрицы **R** позволяет выделить ортогональные составляющие в порядке убывания доли, вносимой

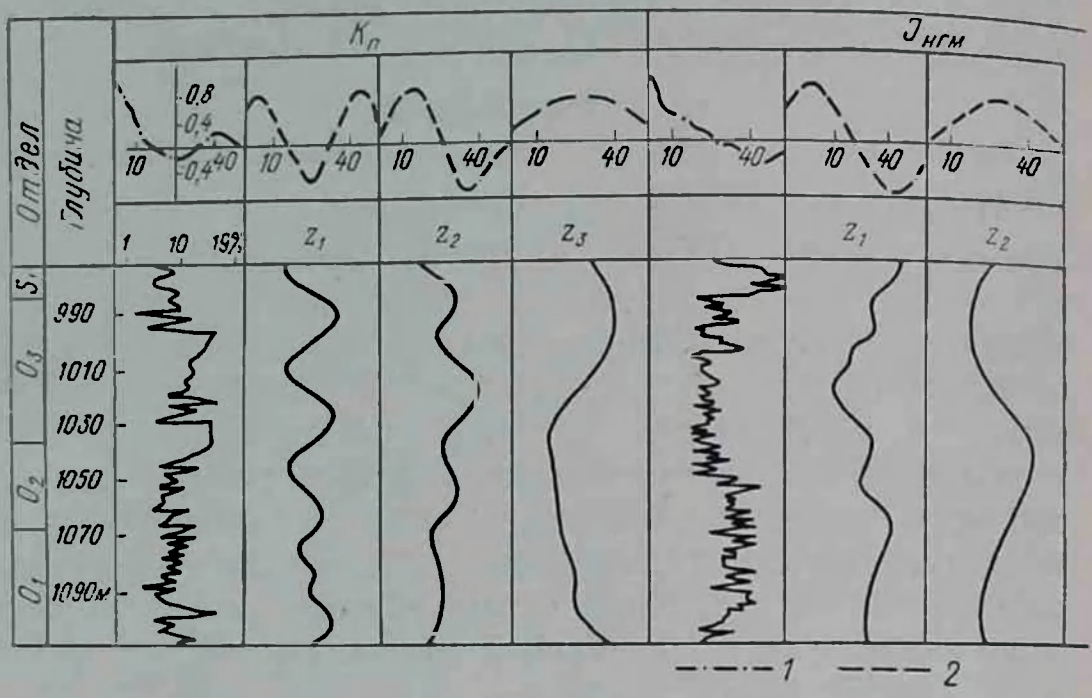


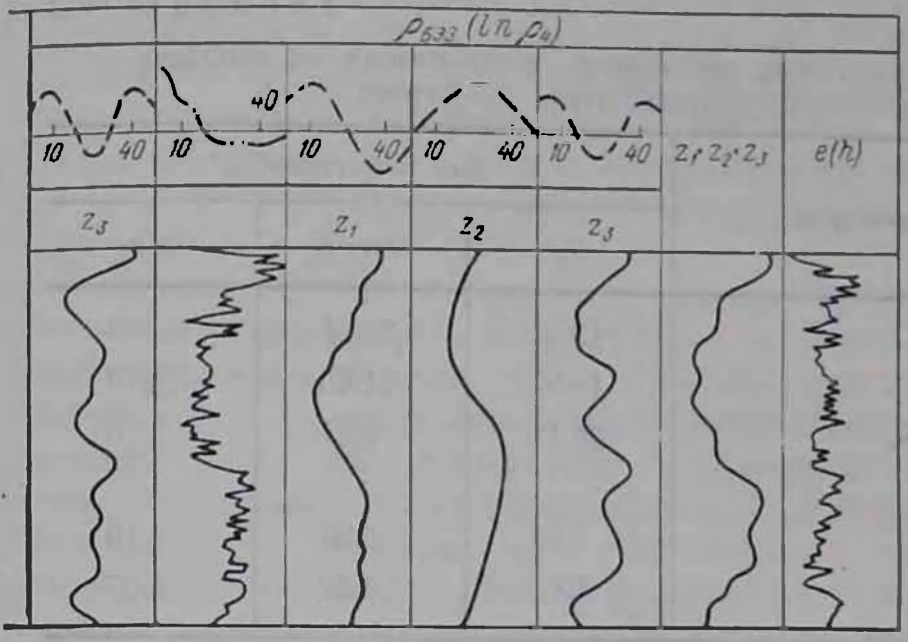
Рис. IV.5. Результаты разложения геолого-геофизиче
нена—
1 — корреляционная функция;

каждой из них в общую дисперсию исходной регистрограммы. Дисперсия каждой составляющей оценивается соответствующим собственным числом матрицы **R**. Когда собственные числа, относящиеся к ортогональной компоненте, становятся достаточно малыми, последующий ряд компонент может быть исключен из рассмотрения, так как они почти не влияют на поведение исходной регистрограммы.

Собственные векторы корреляционной матрицы отражают структуру взаимосвязей между последовательно сдвигаемыми частями исходной регистрограммы. При этом первый вектор описывает «форму» наиболее тесных связей, второй — тех связей, которые выявляются, когда тесные связи устранены, и т. д. При этом заметим, что собственный вектор является своеобразным фильтром. «Пропуская» через него исходную переменную, мы получаем компоненту (естественную составляющую) с внутренними (детерминированными) связями, которые предопределены функционированием природного механизма (динамической сис-

темы). Анализ собственных векторов дает не менее важную информацию о детерминированной изменчивости исходных переменных, чем поведение самих ортогональных составляющих.

Для иллюстрации применения модели Карунена — Лоэва обратимся к геологическим и геофизическим данным, характеризующим карбонатный разрез ордовика по скв. 38-Дурбе (Прибалтика). Были получены корреляционные функции, собственные векторы и соответствующие им естественные ортогональные компоненты z_1—z_3 (рис. IV.5). Обработке (разложению) подвергались: регистрограмма межзерновой пористости $K_п$, диаграммы методов БЭЗ и НГМ. Вместе с разложением исходных кривых



ских переменных по скв. 38-Дурбе с помощью модели Кару-Лоэва.

г — собственные векторы.

осуществлено исключение из них суммы ортогональных компонент $z_1 + z_2 + z_3$, после чего получен «остаток» e_h.

Как видно из сравнения результатов разложения, формы первых трех собственных векторов, относящихся к разным исходным переменным, идентичны (порядок вскрываемых из матрицы R собственных векторов изменен). Поведение естественных ортогональных компонент по разрезу, соответствующих идентичным по форме собственным векторам, для различных переменных также аналогично. Различны лишь веса, с которыми эти компоненты входят в исходные переменные (табл. IV.3). Следовательно, геофизические диаграммы состоят в основном из тех же элементарных составляющих, что и переменная пористости. Однако геофизические переменные по-разному реагируют на составные элементы регистрограммы пористости. Именно это и вызывает разный порядок расположения однотипных естественных ортогональных компонент (компонент, относящихся к тождественным по форме собственным векторам). Так, первая компонента, выделенная из

регистрограммы пористости (вес 20,11%), хорошо сопоставляется с третьей (но находится с ней в противофазе), выделенной из переменной $\ln \rho_4$ (вес 9,50%). Аналогичный вывод можно сделать, сравнивая естественные составляющие пористости и НГМ. Таким образом, разные составляющие пористости обусловливают лишь разную интенсивность составляющих геофизического сигнала, но не разный характер поведения их по разрезу. Этот вывод свидетельствует о том, что исходная изменчивость геофизических характеристик определяется не только собственно значением пористости, но и тем взаимодействием ее последовательных ординат, которое обусловливают неслучайное поведение пористости по разрезу.

Таблица IV.3

Форма и вес ортогональных компонент, выделенных из геолого-геофизических данных (скв. 38-Дурбе)

Форма собственных векторов	Вес компонент, %		
	$K_п$	$J_{НГМ}$	$\ln \rho_4$
	19,78	23,61	28,38
	10,35	21,43	25,73
	20,11	6,87	9,50
—	9,97	5,26	6,19
—	5,94	4,48	4,07

Какие же практические выводы следуют из результатов использования модели Карунена — Лоэва при анализе разного рода регистрограмм?

Во-первых, можно определить, насколько хорошо те или иные геофизические показатели контролируют геологические характеристики. Вопросы оценки информативности геофизических данных часто возникают при внедрении новых методов каротажа. В приведенном примере информация о пористости, заключенная в геофизических данных, составила не более 60—70%. Аналогичный результат был получен при использовании факторной модели.

Во-вторых, по конфигурации естественных ортогональных компонент им можно дать геологическую трактовку. Почти периодический характер их поведения по разрезу свидетельствует о колебательном характере процессов, управляющих распределением пористости. Забегая вперед, отметим, что это вполне согласуется с результатами, которые дает интерпретация данных геолого-геофизических исследований скважин с помощью методов выделения скрытых периодичностей. При этом необходимо подчеркнуть, что в случае использования модели Карунена — Лоэва мы не вкладываем никаких априорных сведений о периодическом поведении

исследуемых регистрограмм, их периодический характер вскрылся как бы естественно (спонтанно). При обращении же к методам выделения скрытых периодичностей (см. гл. V) такие сведения обязательны.

Можно также заметить, что модель Карунена — Лоэва вскрывает тонкие детали в конфигурации разного рода переменных. Так, например, отчетливо обнаруживается, что ортогональные компоненты носят почти периодический, но далеко не строго периодический, характер. Произвольный, но вместе с тем все же колебательный характер их поведения по разрезу, вероятно, лучше согласуется с природным явлением. Поэтому кривая, соответствующая сумме ортогональных компонент, должна точнее описывать неслучайное (эволюторное) распределение пористости по разрезу, чем кривая, соответствующая сумме строго периодических компонент. Следовательно, когда расхождение полициклической модели с фактическими данными оказывается существенным, целесообразно использовать модель Карунена — Лоэва.

Таким образом, с помощью модели Карунена — Лоэва каждая исходная переменная разложена в ряд по естественным ортогональным компонентам. При сопоставлении компонент контрастно выявляется своеобразие каждой геофизической переменной и выясняется, как она контролирует изменение пористости. Периодический (точнее, почти периодический) характер изменения любой выделенной ортогональной компоненты позволяет считать, что определенная доля исходной изменчивости пористости (и геофизических кривых, зависящих от пористости) связана с колебательными явлениями. Но так как это относится только к первичной пористости, то, исключив ее из исходной кривой (т. е. исключив составляющую, которая равна сумме естественных ортогональных компонент), получим «остаток» — флуктуационную (иррегулярную) составляющую, которая по своему смыслу должна лучше всего описывать вторичные преобразования пород, в частности изменение вторичной пористости. Однако если кривая, соответствующая сумме ортогональных компонент, контрастно вскрывает зоны повышенных или пониженных значений пористости межзернового типа, то на иррегулярной составляющей обнаружить какие-либо аномальные зоны невозможно. Видимо, множество других особенностей горных пород, воздействуя на геофизические показатели, проявляются с той же интенсивностью, что и собственно вторичная пористость.

Приведенный пример показывает, что модель Карунена — Лоэва эффективно может быть использована при анализе информативности геофизических показателей и исследовании по ним коллекторских свойств отложений. Определяемые при этом собственные векторы в силу их устойчивости могут служить критериями при корреляции разрезов осадочных толщ. С этой же целью могут быть использованы и ортогональные компоненты,

так как каждая из них отражает поступательное (эволюционное) развитие отдельных сторон сложного геологического процесса. Очевидно, что корреляция разрезов скважин на генетической основе позволит лучше и полнее вскрыть особенности геологического строения.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Мы показали, какую важную и полезную информацию о строении разреза можно извлечь из регистрограммы, если учитывать, что ее ординаты, характеризуя изменение свойств в разрезе, наделены определенными взаимосвязями. Именно эта особенность регистрограмм создает их своеобразную структуру (конфигурацию). Авторегрессионный анализ позволил вскрыть особенности этой структуры и выразить их в терминах определенных параметров. На основании параметрического описания структуры исходных регистрограмм и сравнения параметров между собой удалось ответить на вопрос, насколько схожи (идентичны) конфигурации кривых пористости и различных геофизических методов. Дальнейший анализ регистрограмм с использованием разложения Карунена — Лоэва показал, что идентичность различным образом полученных регистрограмм объясняется наличием в них одинаковых составляющих, которые описывают детерминированное распределение по разрезу одного и того же свойства (в частности, пористости).

Мы не ограничивались сопоставлением регистрограмм и выявлением их общих и специфических черт, а попытались дать генетическую трактовку полученным результатам. С помощью авторегрессионного анализа было показано, что распределением пористости управляет определенный механизм, который при некоторых особых чертах на разных участках бассейна осадконакопления описывается в общем разностным уравнением второго порядка. Исследование с помощью разложения Карунена — Лоэва позволило прийти к выводу, что специфика этого механизма предопределена колебательными движениями дна бассейна седиментации.

Глава V

ВЫЯВЛЕНИЕ И АНАЛИЗ ПЕРИОДИЧНОСТЕЙ В РЕГИСТРОГРАММАХ

ПОСТАНОВКА И СОДЕРЖАНИЕ ЗАДАЧИ

Характерной особенностью разрезов осадочных толщ, и в частности карбонатных, является их циклическое (периодическое) строение. И не случайно поэтому периодическим зависимостям, играющим основную роль в процессе осадконакопления, посвящено большое число исследований [7, 67, 69, 78]. Правда, до

недавнего времени математический аппарат, с помощью которого можно исчерпывающим образом анализировать явления периодичности [71, 74, 80], был мало знаком широким кругам геологов и геофизиков. Толчком к использованию этого аппарата при геологических исследованиях послужили фундаментальные работы А. Б. Вистелиуса [18, 19].

Небезынтересно заметить, что периодичность разных порядков часто улавливается при рассмотрении широкого круга процессов: биологических, физических, социальных и, конечно, геологических. Причины ее разнообразны, но, как показано исследованиями в области кибернетики [17, 93], системы, генерирующие процессы периодического характера, обязательно должны функционировать по принципу обратной связи. Именно обратная связь обеспечивает устойчивость физико-химического состояния системы — ее гомеостат. Этим, вероятно, и объясняются характерные изменения состава и свойств пород по вертикали разреза: не абсолютно равновесные или постоянные, а лишь подобные таковым, поскольку они имеют черты строгой или относительно строгой периодичности. Именно такой характер поведения измеренных свойств и показателей (регистрограмм) мы наблюдаем при изучении осадочных толщ.

Не вдаваясь в подробности карбонатонакопления (эти вопросы полно освещены в трудах Н. М. Страхова [78]), отметим, что основным условием образования карбонатных толщ служит сокращение поступления в осадочный бассейн терригенного материала. Это происходит лишь при относительно однотипном тектоническом режиме, который в целом можно определить как пассивный. По-видимому, не случайно именно платформы вследствие малой тектонической активности максимально приспособлены для развития карбонатных толщ. В геосинклинальных областях карбонатные породы не имеют столь широкого распространения, накопление их приурочено к ограниченным участкам морского дна, а стратиграфический диапазон измеряется обычно горизонтами. Это показывает, что пассивный тектонический режим существовал здесь лишь на локальных площадях и, как правило, был кратковременным.

Таким образом, мощные отложения примерно одного и того же литологического состава свидетельствуют, что в период их образования существовал относительно стабильный геологический режим или, что то же самое, имел место гомеостат. В такой ситуации становится объяснимой наблюдаемая повсеместно осцилляция регистрируемых свойств пород по разрезу, что вызвано лишь некоторой вариацией геологического режима. Это указывает на стремление природы удержать определенное постоянство или, иначе, на действие природного механизма (динамической системы, точнее, геогенератора), наделенного обратной связью. Не случайно поэтому вопрос о колебательных движениях решается по смене типов пород, образующихся при разных

режимах и разных скоростях прогибания области накопления [7, 69].

Ритмичность карбонатных отложений обусловливается следующими причинами [19]. Размер накапливающихся на дне частиц, как и содержание карбонатов в осадках, зависит от глубины моря и рельефа дна. К положительным элементам рельефа тяготеют биоморфные и рифовые накопления, к отрицательным — тонкозернистые карбонаты. Поскольку между рельефом дна и планом тектонических движений существует однозначное соответствие, совершенно естественно, что структура осадков сменяется в определенной исторической последовательности.

Целесообразно вплотную подойти к изучению явлений периодичности с тем, чтобы обнаруживать их не только по геологическим данным, характеризующим разрез, но и по диаграммам геофизических методов исследования скважин. Это может существенно облегчить решение широкого круга практических и теоретических задач нефтяной геологии и геофизики.

ОСНОВЫ МЕТОДА ВЫДЕЛЕНИЯ СКРЫТЫХ ПЕРИОДИЧНОСТЕЙ

Закономерная последовательность карбонатных слоев различной структуры и состава отражает изменения направления и темпов вертикальных движений. Наша задача — по всей совокупности свойств этих слоев вскрыть периодичности, найти их число, что позволит восстановить картину колебательных движений, происходивших в области седиментации. Вместе с тем анализ собственно периодических компонент, как мы убедимся позже, не обязательно должен быть направлен на восстановление колебательных движений, он может иметь непосредственный прагматический интерес. Так, выделенные закономерные периодические компоненты позволят решить задачу по разделению общей пористости карбонатных пород на ее составные части (например, первичную и вторичную) или количественно оценить каждый тип пористости.

При разработке метода следует учесть два важных обстоятельства.

1. Необходимо получить количественные характеристики выделенных периодичностей: значения периода, амплитуды, фазового сдвига (эти характеристики, очевидно, должны отражать темп, интенсивность и направление вертикальных движений).

2. Одна и та же совокупность процессов в области седиментации из-за действия случайных (неконтролируемых) причин может привести, вообще говоря, к разной последовательности свойств слоев, слагающих разрез. Поэтому мы исходили из того, что каждая из последовательностей является одной из нескольких возможных реализаций таких процессов, т. е. мы хотим подчеркнуть, что каждая реализация управляется вероятностными законами и последовательность свойств слоев является случайной.

Указанной специфике задачи отвечает метод выделения скрытых периодичностей. Ранее при решении аналогичных вопросов в качестве свойств, характеризующих разрезы осадочных толщ, служили: пористость [19], выраженные числом составы отложений [18], соотношение сланцевых и карбонатных пород в единице мощности [96], средняя грубость и нерастворимый остаток [88]. Мы кроме различных геологических данных проанализировали результаты геофизических исследований скважин: предварительно квантованные диаграммы каротажа БЭЗ, НГМ, ГМ, АМ и др. Обработка сводилась к анализу каждой регистрограммы (геологической или геофизической), их последующему сопоставлению и увязке.

Итак, если процесс осадконакопления действительно управлялся различными периодически действующими причинами, то во всех перечисленных геолого-геофизических регистрограммах должны усматриваться ритмы (периодические компоненты) различного порядка. Реализация такого процесса, оцениваемая поведением выбранной характеристики по разрезу, должна представляться в виде суперпозиции небольшого числа периодических составляющих (гармоник), осложненных помехами. Математическая модель процесса запишется в виде

$$y_h = \beta_0 + \sum_{j=1}^{q} \beta_j \cos(\omega_j h - \varphi_j) + e_h, \qquad (V.1)$$

где y_h — значение соответствующей характеристики в точке h разреза.

Временна́я шкала заменена шкалой глубины h. Эта замена возможна при условии пассивного тектонического режима осадконакопления, когда за равные периоды времени откладывается примерно одинаковая по мощности толща осадков. Величина $\omega_j = 2\pi/T_j$, где T_j — период j-й периодической компоненты, показывает, какая часть колебания уложится в 1 м мощности.

Первые два члена правой части выражения (V.1) описывают закономерное распределение характеристики по разрезу. Гармонические компоненты, присутствующие в этом распределении, отражают периодичность в изменении самой характеристики, что может быть обусловлено разным порядком колебательных движений земной коры в процессе седиментации осадка. Размах этих движений, последовательность захоронения осадков получают отражение в значениях соответствующих параметров модели: амплитуде колебаний β_j (измеряется в единицах соответствующих характеристик) и сдвиге по фазе φ_j (рад/м).

Третий член e_h определяет индивидуальные черты данной реализации, выделяющие ее из множества других возможных при тех же условиях седиментации. Часто этот член называют шумовой или мешающей компонентой. Однако надо заметить, что это название условно: шум не является безынформативным сигналом,

он может быть столь же информативным, как и закономерная гармоническая компонента.

Решение поставленной задачи сводится к выделению гармонических составляющих из исходной регистрограммы и оценке параметров модели (V.1). Однако прежде необходимо убедиться, что распределение эмпирических данных, характеризующих разрез, не противоречит этой модели. Если амплитуды периодических компонент значительны, их наличие совершенно очевидно. При малых же амплитудах периодические компоненты не проявляются столь отчетливо. Поэтому на практике проводят операцию тестирования на периодичность. Это достигается построением автокорреляционных функций по формуле (IV.6) с последующим их анализом. Если коррелограмма (график автокорреляционной функции) обнаруживает периодическое поведение с ростом сдвига τ и при некоторых сдвигах величина $r(\tau)$ значительна, то исходная эмпирическая переменная содержит периодическую (гармоническую) компоненту. В противном случае можно утверждать, что периодические компоненты отсутствуют и исходная последовательность эмпирических данных согласуется с шумовой компонентой.

Известно, что для любой эмпирической переменной (регистрограммы), и особенно для отражающей колебательные движения, наиболее ярким представлением является частотное (спектральное). Оно обычно легче поддается физической интерпретации, так как становятся очевидными механизм, породивший конфигурацию исходной эмпирической переменной, и характерные черты его функционирования.

Используя теорему Хинчина — Винера [60, 71], можно доказать, что автокорреляционная функция представима в ее частотном аналоге:

$$r(\tau) = \sum_{j=1}^{p} \sigma_j^2 \cos(\omega_j \tau), \qquad \text{(V.2)}$$

где

$$\sigma_j^2 = \frac{1}{2\pi}\left[r(0) + 2\sum_{\tau=1}^{n-1} r(\tau)\cos(\omega_j \tau)\right] \qquad \text{(V.3)}$$

— спектральное представление как автокорреляционной функции, так и исходной эмпирической переменной.

Таким образом, согласно теореме Хинчина — Винера спектр исходной функции (V.3) и автокорреляционная функция (V.2) являются Фурье-сопряженными функциями. Образно спектральное представление показывает, что дисперсия $\sigma^2_{\text{общ}}$, характеризующая рассеяние эмпирической переменной вокруг ее математического ожидания, может быть разложена на частные дисперсии σ_1^2, σ_2^2, ..., σ_p^2 для последовательности частот ω_1, ω_2, ..., ω_p (или

периодов $T_1, T_2, \ldots, T_p$) гармоник, совокупность которых обусловливает и формирует исходную эмпирическую переменную, т. е.

$$\sigma^2_{\text{общ}} = \sum_j^p \sigma^2_{\omega_j} = \sigma^2_{\omega_1} + \sigma^2_{\omega_2} + \ldots + \sigma^2_{\omega_p}.$$

Если среди последовательности дисперсий σ^2_j, относящихся к частотам ω_j, имеются высокие значения, то в исходной эмпирической переменной содержатся гармонические компоненты соответствующих частот.

Таким образом, в тех ситуациях, когда требуется подтвердить или отвергнуть гипотезу о том, что наблюдаемая переменная отражает циклические явления, необходимо обратиться к анализу вида и формы автокорреляционной функции. Спектральное представление обычно позволяет установить численные значения периодов (частот) тех скрытых периодичностей, которые содержатся в исходной эмпирической переменной и обусловливают ее конфигурацию. Кроме того, функция спектральной плотности (спектр) непременно укажет, какие именно периодические компоненты проявляются в эмпирической переменной с наибольшей интенсивностью.

Очевидно, что, определив число периодических компонент и их частоты ω_j, можно переходить к оценке других параметров модели (V.1) — коэффициентов β_j и φ_j. Однако, учитывая специфику поставленной геологической задачи, мы проводили оценку частот (периодов) и числа периодических компонент не методом построения спектральной плотности, а несколько иным путем. Дело в том, что все существующие методы выделения скрытых периодичностей из исходных регистрограмм основаны на предположении, что их длина может быть сколь угодно большой. Реальные же исходные эмпирические кривые всегда характеризуют ограниченный интервал разреза — ограниченный стратиграфический диапазон. При этом в усеченной регистрограмме искомые периоды гармонических компонент, особенно длиннопериодных (низкочастотных), могут быть соизмеримы с интервалом наблюдения или даже превышать его. В такой ситуации поиски нулевого члена (параметра β_0) и периода самой низкочастотной компоненты вызывают большие трудности. Поэтому нами был выработан специальный алгоритм, наиболее приспособленный к специфике геологической задачи. Он позволяет обойти отмеченные трудности и при минимуме информации получить достоверные оценки параметров, характеризующих модель исследуемого геологического явления.

При решении задачи были использованы как элементы статистической теории классификации с обучением, так и теория выделения полезных сигналов на фоне сильно флуктуирующих помех [54]. Сочетание двух теорий приводит к следующему алгоритму, реализованному на ЭВМ, который позволяет обнаружить,

выделить и определить число и все параметры периодических компонент по исходной регистрограмме.

1. На вход ЭВМ поступает дискретный ряд данных (ординаты регистрограммы), отражающих изменение выбранной характеристики по разрезу. Элементы ряда упорядочены и привязаны к соответствующим глубинам скважины или к координатным точкам исследуемого разреза (обнажения). Если эти элементы определены не через одинаковые интервалы, то исходная информация проходит блок линейной интерполяции, после чего дискретные данные представляются в эквидистантных точках разреза.

2. Интерполированные данные поступают в блок вычисления автокорреляционной функции, которая рассчитывается по формуле (IV.6). На основе анализа этой функции принимается одна из альтернативных гипотез: а) исследуемый упорядоченный ряд наблюдений является стационарным случайным процессом с быстро убывающей автокорреляционной функцией; б) в исследуемом упорядоченном ряду кроме стационарного шума присутствуют гармонические составляющие, в этом случае автокорреляционная функция содержит периодические компоненты с большим временем корреляции и ее изменения имеют четко выраженный периодический характер. Итак, исследуемая регистрограмма при первой гипотезе не содержит детерминированных компонент, при второй — согласуется со стохастической моделью вида (V.1).

3. При выполнении последнего условия исходная информация подвергается преобразованию с целью определения всех параметров стохастической модели (V.1). Это достигается применением принципов «обучения с учителем», в результате чего в памяти ЭВМ формируются последовательности эталонных гармонических составляющих с различными (пробными) периодами и нулевыми (пробными нулевыми) членами, задаваемыми через определенный шаг. Заметим, что шаг пробного периода не является кратным значению $2\pi/n$ (где n — число наблюдений на интервале исследования), как это делается при вычислении спектральной плотности. При известных пробном периоде и пробном нулевом члене остальные параметры модели и эталонных функций $\delta_h\ (\beta_{0i},\ T_j)$ вычисляются по эмпирическим данным методом наименьших квадратов. Затем проводится сравнение исходной регистрограммы y_h и эталонной функции $\delta_h(\beta_{0i},\ T_j)$. При этом критерием их близости служит квадратичная мера

$$S(\beta_{0i},\ T_j) = \frac{1}{n-1}\sum_{h=1}^{n}[y_h - \delta(\beta_{0i},\ T_j)]^2,$$

являющаяся функционалом двух параметров: среднего члена и периода.

В последовательности величин $S(\beta_{0i},\ T_j)$, отвечающих вариационному ряду значений нулевого члена и периода, отыскивается наименьшая величина $S(\beta_{0k},\ T_m) = \min$. Значениями β_{0k} и T_m,

при которых $S(\beta_{0k}, T_m)$ достигает общего минимума, и определяются искомые значения как нулевого члена β_0, так и периода T_m той гармонической составляющей, которая доминирует в исследуемой регистрограмме. При знании периодов гармоник (или их частот) отыскание параметров β_j и φ_j не вызывает особых трудностей и осуществляется с помощью регрессионного анализа (см. главу I). Действительно, если стохастическую модель (V.1) представить в виде

$$y_h - \beta_0 = \sum_{j=1}^{q} \beta'_j \cos(\omega_j h) + \sum_{j=1}^{q} \beta''_j \sin(\omega_j h) + e_h,$$

то соответствующая ей запись в матричных обозначениях примет вид

$$\mathbf{Y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \mathbf{e},$$

где $\mathbf{Y}$ — вектор наблюдений; $\boldsymbol{\beta}_1$ и $\boldsymbol{\beta}_2$ — искомые векторы;

$$\mathbf{X}_1 = \begin{pmatrix} \cos(\omega_1 \Delta\tau h_1) & \dots & \cos(\omega_q \Delta\tau h_1) \\ \cos(\omega_1 \Delta\tau h_2) & \dots & \cos(\omega_q \Delta\tau h_2) \\ \dots & \dots & \dots \\ \cos(\omega_1 \Delta\tau h_n) & \dots & \cos(\omega_q \Delta\tau h_n) \end{pmatrix};$$

$$\mathbf{X}_2 = \begin{pmatrix} \sin(\omega_1 \Delta\tau h_1) & \dots & \sin(\omega_q \Delta\tau h_1) \\ \sin(\omega_1 \Delta\tau h_2) & \dots & \sin(\omega_q \Delta\tau h_2) \\ \dots & \dots & \dots \\ \sin(\omega_1 \Delta\tau h_n) & \dots & \sin(\omega_q \Delta\tau h_n) \end{pmatrix};$$

$\Delta\tau$ — шаг квантования; $\omega_1, \dots, \omega_q$ — найденные частоты.

Оценка линейных коэффициентов β'_j и β''_j по методу наименьших квадратов дается формулами [43]

$$\boldsymbol{\beta}_1 = (\mathbf{X}'_1\mathbf{X}_1)^{-1}\mathbf{X}'_1(\mathbf{Y} - \mathbf{X}_2\boldsymbol{\beta}_2);$$

$$\boldsymbol{\beta}_2 = (\mathbf{X}'_2\mathbf{D}\mathbf{X}_2)^{-1}\mathbf{X}'_2\mathbf{D}\mathbf{Y},$$

где $\mathbf{D} = \mathbf{I} - \mathbf{X}_1(\mathbf{X}'_1\mathbf{X}_1)^{-1}\mathbf{X}'_1$ (произведения типа $\mathbf{X}'\mathbf{X}$ и $\mathbf{X}'\mathbf{Y}$ определенным образом нормируются); $\mathbf{I}$ — единичная матрица.

Очевидно, что график функции $S(\beta, T)$ есть не что иное, как модифицированная периодограмма, где минимальным значениям $S(\beta, T)$ отвечают искомые периоды составляющих T (при найденном и зафиксированном нулевом члене), сумма которых определяет исследуемую регистрограмму (без компоненты e). Эта функция является определенным аналогом функции спектральной плотности, с помощью которой также получают сведения о структуре изучаемого процесса (при этом значения спектральной плотности отыскиваются в точках, не обязательно кратных величине $2\pi/n$, что лучше согласуется со спецификой решаемой геологической задачи).

На выходе ЭВМ печатаются найденные периодические составляющие, их параметры и остаток, полученный после вычитания доминирующих гармонических компонент из исходной регистрограммы. Моментом окончания анализа является получение остатка е, близкого к реализации случайного стационарного процесса, что распознается в результате анализа функции автокорреляции этого остатка.

Предлагаемый алгоритм, как показали исследования, позволяет отыскать периодические компоненты, период которых в 1,5—2 раза превосходит интервал наблюдений. Следовательно, указанный подход весьма благоприятен для длиннопериодных (низкочастотных) составляющих, что присуще ограниченному числу алгоритмов и весьма ценно для решения геологических задач.

Следует подчеркнуть, что задача выделения скрытых периодических компонент из эмпирической переменной принципиально отлична от задачи разложения в ряд Фурье. В первом случае период искомой компоненты не предопределяется заранее, а устанавливается в ходе самого исследования, поэтому каждая выявленная периодическая компонента с соответствующими ей параметрами может быть отнесена к какому-либо колебательному процессу определенной длительности и интенсивности.

И еще одно замечание. Как уже отмечалось, часто возникает ситуация, когда само наличие регулярной компоненты в регистрограмме является сомнительным. Вместе с тем априорные физические соображения настолько бедны, что не дают веских оснований принять гипотезу о наличии периодичностей. И тем не менее может оказаться, что функция автокорреляции имеет длинный почти периодический «хвост», а периодограмма указывает на присутствие гармоники определенной частоты (периода). Возникает вопрос: является ли такая ситуация игрой случая или регулярный фактор действительно присутствует? В геологических исследованиях на этот вопрос иногда можно ответить однозначно, ибо, как правило, обработке подвергаются данные по ряду близко расположенных скважин, секущих осадочную толщу, формирование которой обязано какому-либо единому геологическому процессу.

Е. Е. Слуцким [74, 90] строго доказано, что случайные последовательности, обусловленные внутренней корреляцией, могут давать реализации, сколько угодно близкие к гармоническим, и, следовательно, казалось бы, никакие статистические методы анализа не позволяют с полной уверенностью доказать наличие периодических компонент (в конечной реализации случайного процесса). Тем не менее тестирование на периодичность целесообразно проводить (в сочетании с априорными данными и физическими соображениями), так как в этом случае причина возникновения внутренней корреляции (между ординатами регистрограммы) никем не установлена.

Причина же, вероятно, кроется в особенности функционирования природного механизма, воспроизводящего исследуемую оса-

дочную толщу. Если параметры механизма для разных частей площади идентичны, то внутренняя корреляция может быть одинаковой силы, что в свою очередь может обусловить периодические компоненты равных или близких периодов. Анализ, основанный на рассмотрении периодо- и коррелограмм, как раз и позволяет ответить на вопрос, образован ли анализируемый процесс последовательностью независимых случайных величин (так называемым «белым шумом») или последовательностью коррелированных величин с известным спектром.

ИЗУЧЕНИЕ РАЗРЕЗОВ КАРБОНАТНЫХ ТОЛЩ

Периодичность в изменении состава и физических свойств

При выборе объектов для исследования учитывались основные и специфические особенности накопления карбонатов в условиях аридного и гумидного климата, а также различия в тектоническом режиме платформенных и геосинклинальных областей. Очевидно, что для полноты результатов желательно было изучить разрезы карбонатных отложений различного возраста, находящиеся в различных геотектонических областях, прошедшие различную историю развития и значительно удаленные друг от друга. Однако, учитывая, что платформы по своему режиму наиболее благоприятны для карбонатонакопления и на их территории особенно четко проявляется режим колебательных тектонических движений, большинство разрезов необходимо было взять именно из этих областей. С другой стороны, для сравнительного анализа следовало изучать разрезы одного возраста, не только территориально удаленные, но и расположенные в пределах одной локальной структуры и приуроченные к разным ее элементам (свод, крыло). Специфика анализа требует, чтобы в разрезах не было перерывов в осадконакоплении. В соответствии с этими соображениями мы проанализировали 19 разрезов карбонатных толщ, вскрытых глубокими скважинами в различных районах СССР.

Четыре разреза (скв. 35, 36, 38, Дурбинское поднятие: скв. 3-Ремте, Салдусское поднятие) приходится на Прибалтийскую синеклизу, которая представляет собой область сравнительно глубокого погружения северо-западной окраины Русской платформы. Формирование синеклизы связано с каледонским тектогенезом. Впадина осложнена структурными элементами, обусловленными блоковыми подъемами кристаллического архейского фундамента. Палеозойские отложения образуют здесь платформенные структуры. Исследовались отложения ордовика, представленные мергелями, доломитами и известняками. Дурбинское и Салдусское поднятия тяготеют к приподнятым участкам фундамента в северной части синеклизы.

Три разреза (скв. 1, 3, 11, Успинская площадь) расположены на северо-восточной окраине Русской платформы. Все разрезы приурочены к Усино-Колвинскому мегавалу. Объектом изучения служили отложения карбона: доломиты и известняки, доломитизированные, органогенные и глинистые, иногда с прослоями глин и алевролитов.

Девять разрезов (скв. 2, 9—12, 15, 20, 51, 56, Марковское поднятие) расположены в пределах Усть-Киренского свода — крупной структуры

Ангаро-Ленской синеклизы Сибирской платформы. Свод осложнен несколькими валами, одним из них является Марковский, в его пределах расположено одноименное поднятие. Свод сформировался в среднем рифее. Разрезы нижнего кембрия относятся здесь к платформенному типу. Объектом исследования служил карбонатный осинский горизонт нижнего кембрия, приуроченный к мощной галогенной формации.

Остальные разрезы расположены в пределах Терско-Каспийского передового прогиба, входящего в систему Предкавказских передовых прогибов. В альпийскую складчатость они были заложены в теле палеозойского складчатого основания на полосе, примыкающей к Кавказской геосинклинали. Для прогиба характерна асимметрия: геосинклинальный борт очень узкий и крутой, платформенный склон широкий и пологий.

Терско-Сунженская нефтегазоносная область, к которой приурочены два разреза (скв. 826, 833, Малгобек-Вознесенская антиклиналь), связана с западной частью Терско-Каспийского передового прогиба; она включает в себя зону передовых складок Северо-Восточного Кавказа и геосинклинальный борт прогиба. Разрезы расположены в пределах Терского антиклинория, находящегося в южной части прогиба. Объект изучения — верхнемеловые отложения, представленные пелитоморфными и песчанистыми известняками с прослоями мергелей и глин; эти образования относятся к геосинклинальному типу.

Западная антиклинальная зона Южного Дагестана, в которой расположен один разрез (скв. 2, Сергокала), относится к Дагестанской нефтегазоносной области, охватывающей полосу передовых складок Северо-Восточного Кавказа и являющейся восточным продолжением Терско-Сунженской зоны поднятий. В тектоническом отношении зона приурочена к складчатому борту Терско-Каспийского передового прогиба. Объектом исследования служили отложения верхнего мела, представленные пелитоморфными и форaминиферовыми известняками, мергелями и глинистыми известняками с прослоями глин, которые также являются образованиями геосинклинального типа.

Таким образом, исследованию подверглись карбонатные толщи широкого возрастного диапазона (от кембрия до позднего мела), формировавшиеся в условиях как платформенного, так и геосинклинального режима. Изученные толщи входят в состав и типичных морских карбонатных, и галогенных формаций, т. е. они накапливались в самых разнообразных условиях, что позволяет придать результатам, которые мы получим ниже, самый широкий смысл.

Геофизический и геологический материал, характеризующий разрезы указанных карбонатных отложений, обрабатывался и анализировался по единой схеме. Сначала на основе исходных геофизических и геологических регистрограмм были построены автокорреляционные функции и периодограммы. Их интерпретация позволила убедиться, что в исходных регистрограммах содержатся периодические компоненты. Затем из регистрограмм были выделены доминирующие гармонические компоненты. Их увязка и сравнение показали, что в подавляющем большинстве случаев по данным геофизики мы получаем ту же информацию, что и по данным прямых геологических наблюдений.

Рассмотрим решение различных геологических задач. Следуя рецептам методики выделения скрытых периодических компонент, сначала обратимся к анализу коррелограмм.

Коррелограмм-анализ. Рассмотрение коррелограмм, рассчитанных как по кривым пористости, карбонатности, плотности и других свойств пород, определенных по керну, так и по диаграммам различных видов каротажа (НГМ, ГМ, БЭЗ и др.), показывает, что исходные регистрограммы, характеризующие разрезы карбонатных пород, содержат периодические компоненты (рис. V.1). Коррелограммы имеют вид затухающих косинусоид: то ослабевающая, то вновь усиливающаяся корреляция с ростом параметра τ (сдвига) свидетельствует о наличии периодических компонент в исходных регистрограммах.

Примечательно, что после исключения из исходных регистрограмм периодических компонент вид и форма коррелограмм сильно меняются: уже при небольших значениях τ функция корреляции резко затухает, а затем хаотически осциллирует около оси абсцисс. Такой вид коррелограммы свидетельствует, что «в остаточной» регистрограмме ярко выраженных закономерных компонент периодического характера уже нет. «Остаток» в большей мере согласуется либо со случайной стационарной последовательностью, ординаты которой не коррелированы, либо с авторегрессионной последовательностью порядка не выше первого.

Из сопоставления и анализа коррелограмм, построенных по геофизическим диаграммам и по геологическим данным, явствует, что их конфигурация практически идентична. Это свидетельствует о статистической идентичности структуры исходных регистрограмм.

Периодограмм-анализ. Более детально характер изменчивости геологических и геофизических регистрограмм, их схожесть и различие могут быть установлены при помощи периодограмм-анализа. Кроме того, по периодограммам можно точно определить, сколько периодических компонент суммируются в исходных регистрограммах и чему равен период каждой выделенной гармоники. Рассмотрим периодограммы (рис. V.2), относящиеся к разного рода геолого-геофизическим регистрограммам. Периодограммы для одного и того же разреза, но для разных геолого-геофизических показателей совпадают. Это свидетельствует, что исходные регистрограммы разной физической природы (кривые пористости, диаграммы различных геофизических методов) состоят из одних и тех же периодических компонент. Периодограммы также показывают, что число компонент небольшое. Очевидно, что структура разных исходных регистрограмм, образованных равным числом гармоник одинакового периода, должна совпадать. Следовательно, в тех случаях, когда периодограммы, относящиеся к геологическим и геофизическим показателям, совпадают, можно сделать заключение, что, обрабатывая только геофизические показатели, мы получим те же результаты, что и при обработке прямых геологических наблюдений.

Для некоторых разрезов отмечалось небольшое, реже существенное, различие в периодограммах. Оно может быть объяснено

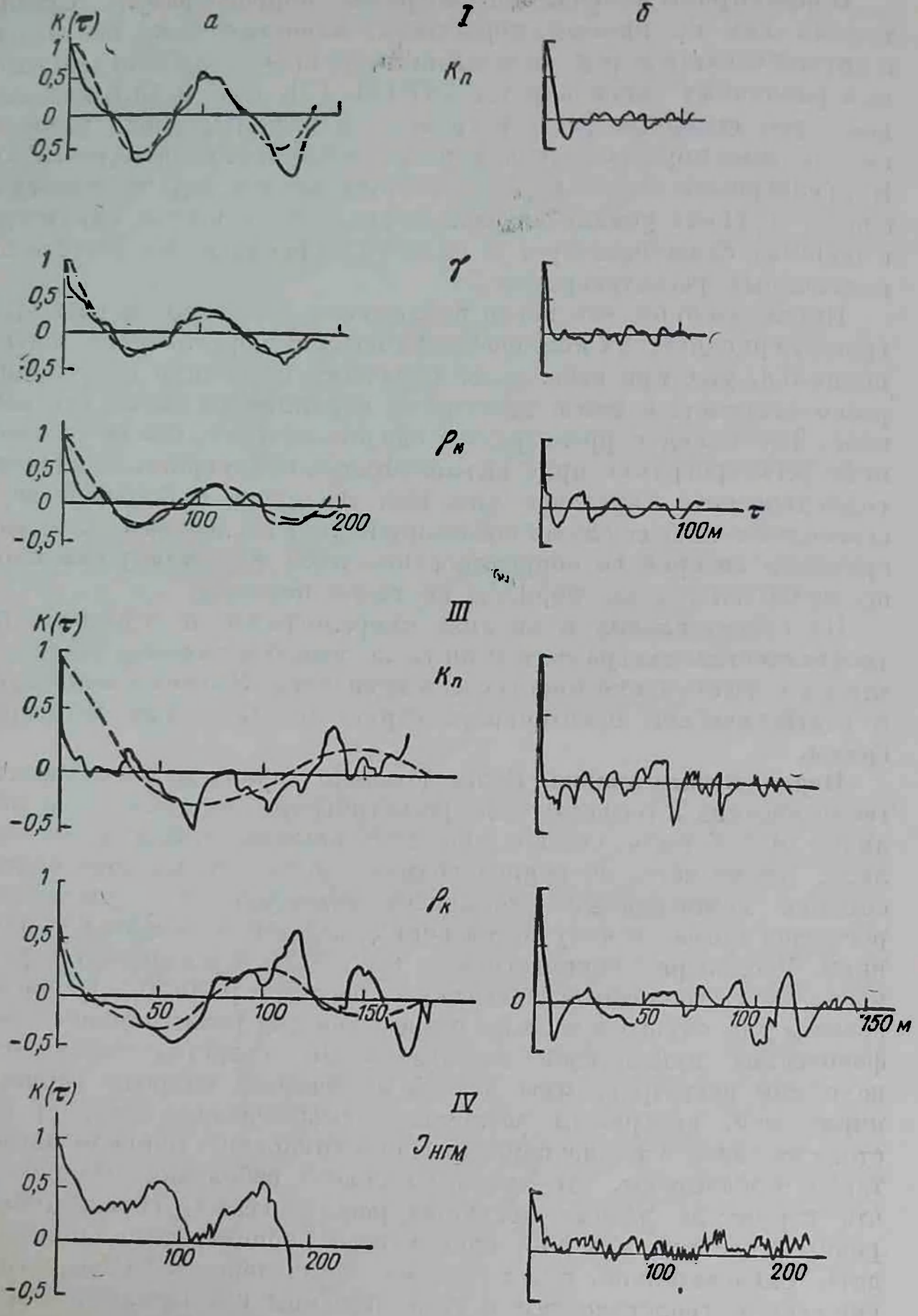
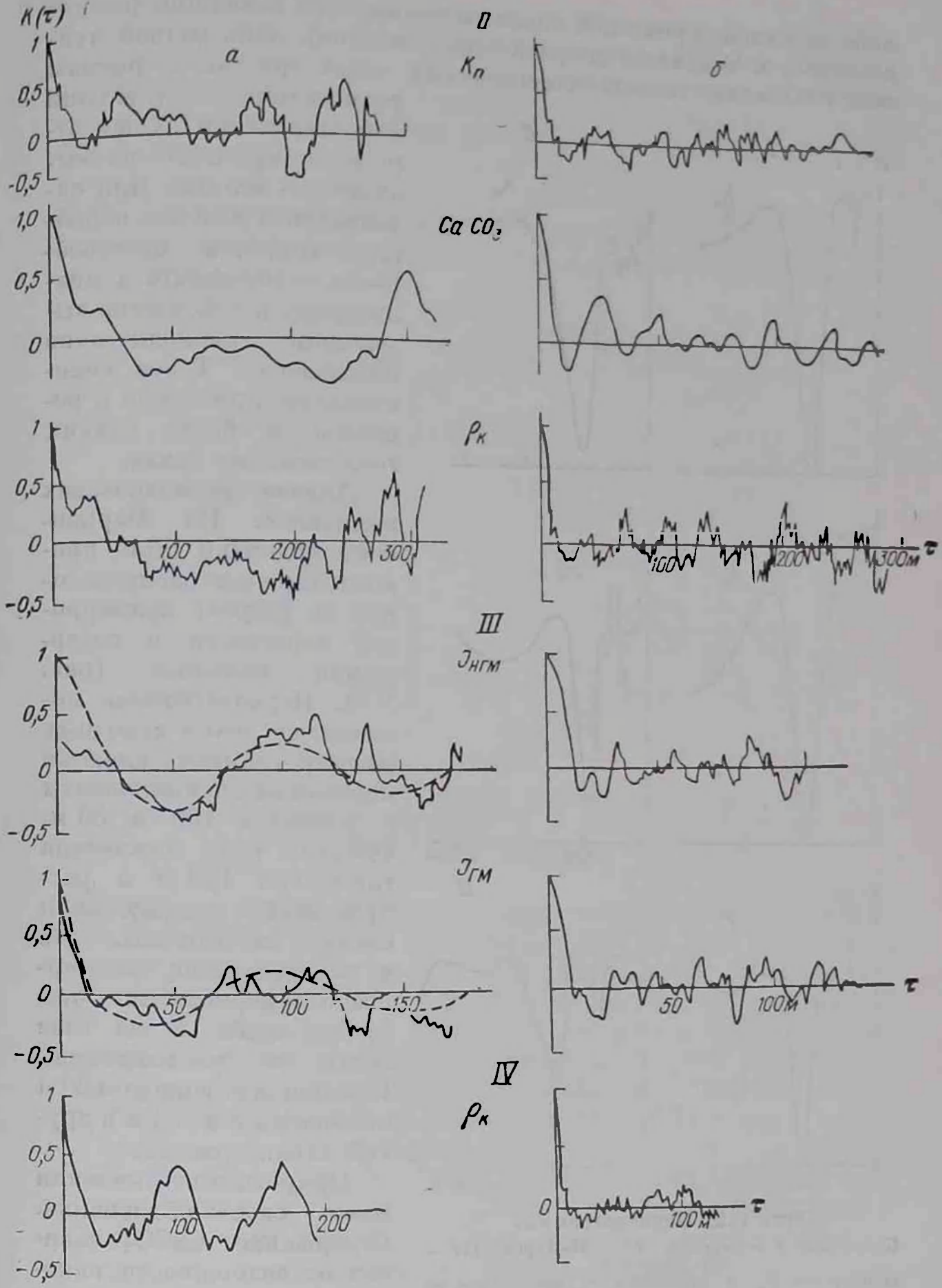


Рис. V.1. Коррелограммы (авто

Скважины: *I* — 3-Ремте, *II* — *2*-Сергокала, *а* — по исходным значениям; *б* — после Пористость $K_п$, плотность γ и карбо



корреляционные функции).

III — 38-Дурбе, *IV* — 833-Малгобек.

исключения периодических компонент.

натность $CaCO_3$ определены по керну.

либо неточной привязкой образцов керна к оси скважины (следовательно, и к данным геофизических методов), либо разной чувствительностью геолого-геофизических характеристик к ритмам разного порядка, т. е. одни характеристики лучше отражают крупные ритмы, другие — мелкие. Для однозначного решения подобных вопросов целесообразно сопоставлять и анализировать собственно выделенные периодические компоненты. Такое сопоставление приводило к решению и более тонких геологических задач.

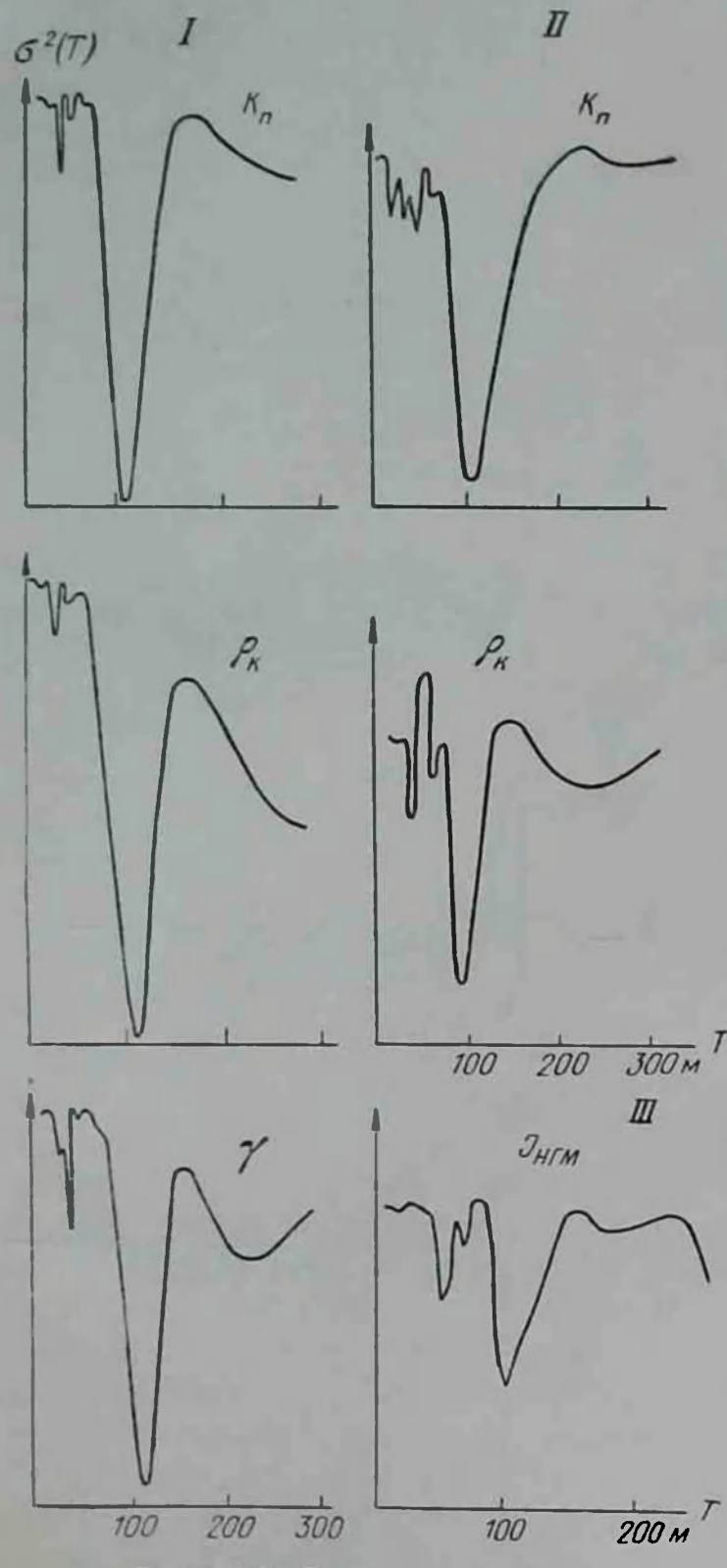


Рис. V.2. Периодограммы.
Скважины: *I* — 3-Ремте, *II* — 38-Дурбе. *III* — 826-Малгобек.
Пористость $K_п$ и плотность γ определены по керну.

Анализ периодических компонент. На Марковском поднятии было проанализировано распределение по разрезу межзерновой пористости и содержания кальцита (рис. V.3). Периодограммы показывают, что в исходных регистрограммах имеются периодические компоненты с периодом 150 и 30 м. Однако, если амплитуда гармоники 150 м в распределении содержания $CaCO_3$ значительна, то в распределении межзерновой пористости она крайне слаба и на глаз почти не улавливается. Гармоника с периодом 30 м отмечается и в той и в другой характеристике.

Природа этого явления имеет свое объяснение. Содержание $CaCO_3$ зависит от возможности накопления в породе остальных компонентов. Видимо, такая возможность определяется движениями более высокого порядка, поскольку она зависит от солености вод бассейна. В то же время на пористость влияет способ укладки накапливающихся на дне частиц, что связано в основном

с движениями более низкого порядка. Таким образом, движения разного порядка обусловливают определенные соотношения между характеристиками пород. На основании этих соотношений и их изменений в тех или иных условиях можно получить более обширную информацию о характере колебательных движений и связанных с ними особенностях осадконакопления. Очевидно, что такие выводы невозможно сделать, если использовать корреляционные методы, ибо значения коэффициента корреляции между рассматриваемыми характеристиками могли бы привести к противоположным результатам.

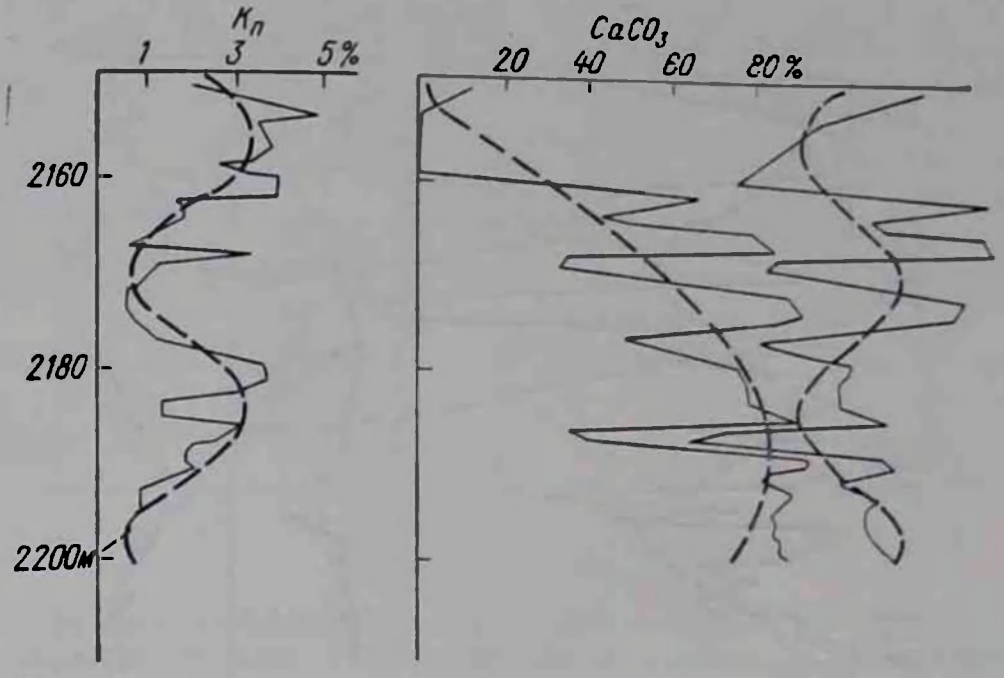


Рис. V.3. Сопоставление ритмов, выявляемых по распределению пористости и содержания $CaCO_3$ (скв. 2, Марковская площадь, осинский горизонт).

Движения одного порядка также приводят к определенному взаимоотношению тех или иных частей карбонатного осадка, что довольно четко выявляется при описанном подходе исследований.

Осинский карбонатный горизонт нижнего кембрия юга Сибирской платформы заключен в мощной соленосной формации. Породы здесь сменяют друг друга в последовательности, практически соответствующей схеме сульфатно-карбонатного осадконакопления. Предопределенность ее ритмичностью колебательных движений может быть проиллюстрирована распределением глинистого материала, засоленности и карбонатности разреза (рис. V.4). Распределение всех трех компонентов носит отчетливо выраженный ритмический характер, о чем свидетельствовали коррелограмм- и периодограмм-анализы. Значение одного из выделенных периодов у всех трех характеристик совпадает и составляет 30 м. Все это указывает на четкую связь между содержанием того или иного материала и колебательными движениями одного порядка.

В общих чертах повышению глинистости отвечает уменьшение засоленности и увеличение карбонатности. Поступление пресных

вод, выносящих терригенный материал, снижает минерализацию бассейна, препятствуя осаждению более легко растворимого галита, что в свою очередь ведет к увеличению содержания карбонатного компонента. Обращает на себя внимание некоторое запаздывание (сдвиг по фазе) в изменении концентрации разных компонентов. Так, наименьшее содержание галита в разрезе несколько сдвинуто по отношению к максимуму глинистого материала, а максимальное содержание карбоната в свою очередь также смещено по отношению к минимуму галита, причем сдвиг во всех

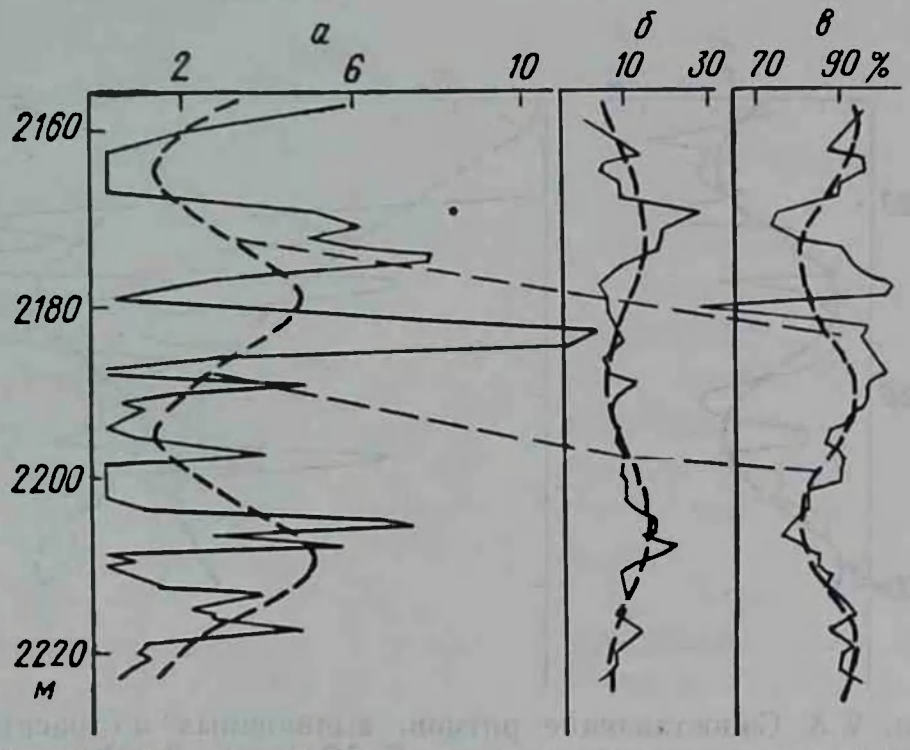


Рис. V.4. Сдвиг по фазе в распределении глинистости (*а*), засоленности (*б*) и карбонатности (*в*) по разрезу осинского горизонта (скв. 12, Марковская площадь).

случаях происходит в одну сторону. Иначе говоря, при химической седиментации (осаждение галита) момент наступления специфичных условий для садки той или иной хемогенной фазы и момент ее отложения не всегда совпадают. В каждый момент осаждаются унаследованные твердые фазы, отражающие химическую седиментацию предшествующей ступени осолонения. Наличие в хемогенных породах унаследованных твердых фаз отмечает Н. М. Страхов [78]. Благодаря используемому методу эту унаследованность можно измерить количественно — сдвигом по фазе одного компонента относительно другого. Картируя эти сдвиги, можно выявлять особенности осадконакопления, а следовательно, и колебательных движений.

Приведем еще один пример подобного рода (рис. V.5). Интервал разреза, относящийся к среднему карбону, охарактеризован комплексом геолого-геофизических данных. Периодичность их изменения не вызывает сомнения. Периодограмм-анализ позволяет выявить во всех распределениях доминирующую гармонику с периодом около 130 м, что указывает на проявление колебательного движения аналогичного порядка. Наиболее отчетливо эта.

гармоника прослеживается в регистрограммах $CaCO_3$ и $CaMg(CO_3)_2$. В геофизических диаграммах и регистрограмме пористости гармоника 130 м проявляется менее отчетливо и по отношению к характеристикам вещественного состава она сдвинута на определенный шаг.

Эти явления могут быть истолкованы следующим образом. Та или иная концентрация карбонатов является отражением колебательных движений. Следовательно, пористость в данном случае формируется под одновременным влиянием как карбонатности, так и доломитизации пород. Соотношение этих двух породообразующих компонентов, вероятно, определяет структуру пород, а следовательно, и их пористость. Надо заметить, что максимум (или минимум) содержания доломита приходится на те точки разреза, где

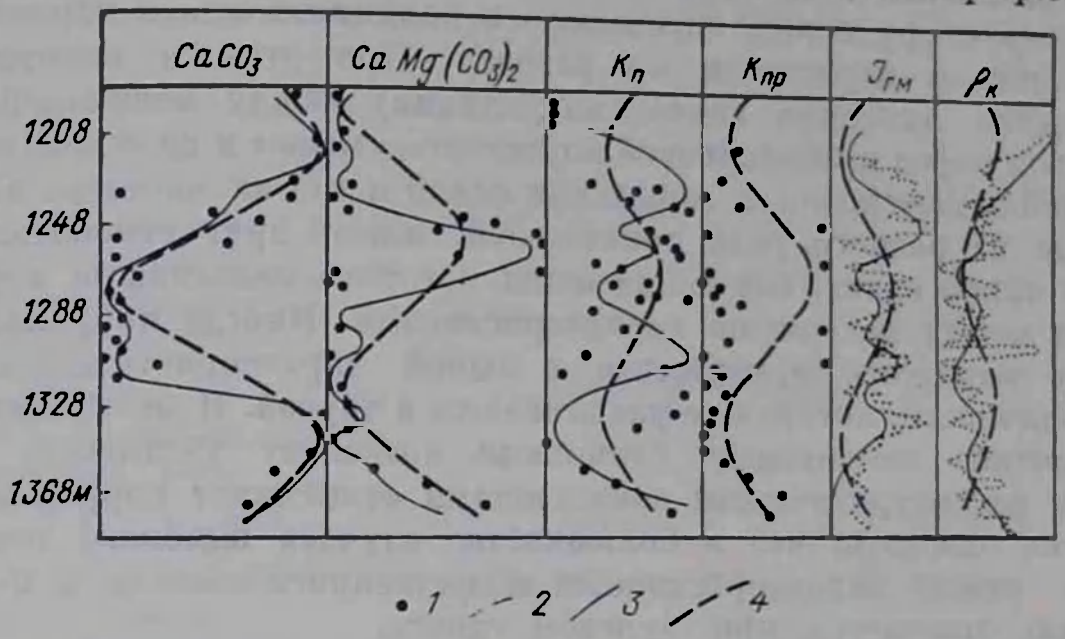


Рис. V.5. Геолого-геофизические переменные и выделенные из них периодические компоненты (скв. 11-Уса, Тимано-Печорская нефтегазоносная область).

1 — фактические значения переменных; *2* — исходные диаграммы геофизических методов; *3* — сумма периодических компонент (полициклическая составляющая); *4* — гармоническая компонента с периодом 130 м.

изменение карбонатности имеет наибольший градиент. Это приводит к тому, что на разных участках разреза карбонатность может иметь либо прямую, либо обратную связь с содержанием в породе доломита. Одновременное влияние двух породообразующих факторов вызывает сдвиг по фазе в гармонике, выделенной из регистрограммы пористости, и она занимает как бы промежуточное положение между гармониками, выделенными из распределений $CaCO_3$ и $CaMg(CO_3)_2$.

Соответствующая гармоника геофизических регистрограмм контролирует изменение либо глинистости (карбонатности) (кривая ГМ), либо пористости (кривая $\rho_к$). И хотя в общих чертах повышению пористости соответствует уменьшение сопротивления и увеличение данных ГМ, все же в данном случае в геофизических показателях присутствуют такие гармонические компоненты, которые не связаны ни с вещественным составом пород, ни с их пористостью. Эти компоненты зависят, вероятно, от таких специфических свойств пород, контроль над которыми не осуществлялся.

Итак, анализ коррело- и периодограмм приводит нас к заключению, что все геологические и геофизические регистрограммы имеют сходную структуру и согласуются с моделью (V.1). Характеризующие один и тот же разрез различные исходные геологические и геофизические регистрограммы содержат гармонические компоненты одинакового периода. Существование строгих периодич-

ностей в распределении вещественного состава пород по разрезу показывает, что изменение геологической обстановки во времени носило колебательный характер. При этом в бассейне седиментации карбонатов отмечается разный порядок колебательных движений, что предопределило появление в отмеченных распределениях гармоник соответствующей частоты (периода).

Сопоставление гармоник одного периода, выделенных из разного рода показателей, свидетельствует о том, что полициклическое изменение вещественного состава, обусловив соответствующую структуру пород, определило и полициклическое изменение межзерновой пористости по разрезу. При этом в некоторых ситуациях линейная связь (корреляция) между вещественным составом пород и межзерновой пористостью может и не отмечаться. Это наблюдается, когда гармоники одной и той же частоты, выделенные из разного рода показателей, имеют друг относительно друга такой сдвиг, без исключения которого невозможна корреляция между исходными характеристиками. Иногда корреляцию может нарушить присутствие в одной характеристике такой периодичности, которая не улавливается в другой. Поэтому только исключение «мешающей» гармоники позволяет убедиться, что между рассматриваемыми показателями существует корреляция. Следует отметить, что в большинстве случаев наиболее тесная связь между характеристиками вещественного состава и пористостью отмечается при нулевом сдвиге.

Таким образом, строгие периодичности определенной частоты, усматривающиеся в распределениях межзерновой пористости, отражают колебательные движения разного периода. Эти движения, по-видимому, проявлялись во время седиментации осадков. Поэтому вполне правдоподобно, что та составная часть пористости, поведение которой по разрезу носит строго полициклический характер, вероятно, имеет первичное (седиментационное) происхождение. Примечательно, что постседиментационные процессы непременно преобразовывали значения пористости, но обращает на себя внимание факт, что эти процессы не исказили полициклического поведения пористости по разрезу. Вероятно, менялись лишь абсолютные размеры пор (например, в процессе уплотнения пород), но относительное их распределение сохранилось.

Доминирующие периоды осадконакопления, выделенные из геологических и геофизических регистрограмм, очевидно, отражают частоту колебательных движений, воздействующих на бассейн седиментации во время карбонатонакопления (табл. V.1). Значения периодов одного и того же ритма для разных районов варьируют в незначительных пределах, несмотря на разницу в возрасте отложений и на приуроченность их к платформенным или геосинклинальным бассейнам. Поскольку периоды всюду выражены в метрах, можно считать, что за одно и то же время в геосинклиналях и на платформах откладывались примерно одинаковые мощности карбонатных осадков. Это можно объяснить только тем,

что скорости погружения платформ и геосинклиналей во время карбонатонакопления были примерно одинаковыми, т. е. в жизни геосинклиналей бывали такие кратковременные эпизоды (карбонатонакопление здесь происходило локально), когда характер их колебательных движений не отличался от платформенного.

Таблица V.1

Периоды циклов осадконакопления разного порядка

Структура	Периоды, м				
	I	II	III	IV	V
Русская платформа					
Балтийская синеклиза	220—300	104—140	65—80	35—45	15—25
Успно-Колвинский вал	220—300	115—135	65—95	35—50	15—25
Сибирская платформа					
Ангаро-Ленская синеклиза	—	120—150	55—80	30—45	18—25
Терско-Каспийский передовой прогиб					
Терский антиклинорий	175—220	100—125	75—80	35—50	20
Западная антиклинальная зона Южного Дагестана	270	100—140	75—80	35—70	20

Учитывая, что карбонатонакопление приурочено в основном к срединным стадиям геотектонических циклов, можно говорить о соизмеримости скоростей движений в геосинклиналях и на платформах именно в эти отрезки времени. Это подтверждается данными А. Б. Ронова [67], хотя он и допускает разницу в 2—3 раза. Однако, как показывают наши данные, непосредственно в моменты карбонатонакопления эта разница могла полностью сгладиться.

Характерно, что совпадает не только частота движений, но и число ритмов разного порядка. Это указывает на выдержанность колебательных движений определенного порядка в период карбонатонакопления в геосинклиналях и на платформах. Более того, можно считать, что карбонатонакопление обусловливается как бы единственным набором колебательных движений, т. е. режим гомеостата является жестким и выдерживается в достаточно узких рамках. Некоторое исключение, казалось бы, составляет карбонатный осинский горизонт юга Сибирской платформы (Марковское поднятие), ибо здесь отсутствуют ритмы высокого порядка. Однако это связано скорее с кратковременностью формирования осинского горизонта, а не с изменением характера движений.

Более крупным ритмам здесь отвечают определенные части соленосной толщи, составным элементом которой является осинский горизонт.

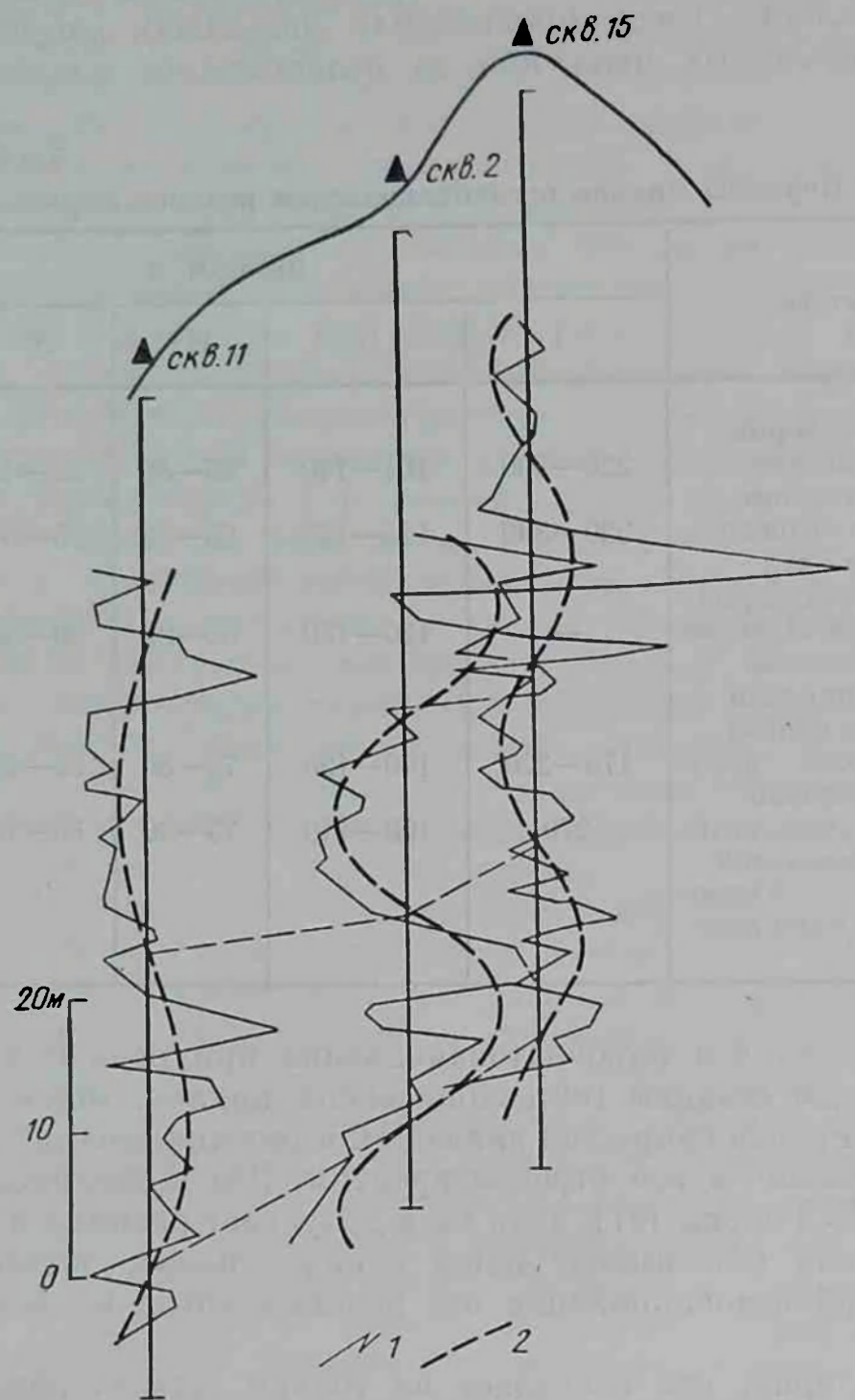


Рис. V.6. Зависимость периода ритмов от структурного плана отложений (Марковская площадь, осинский горизонт).

1 — фактические кривые пористости; *2* — доминирующие периодические компоненты.

При близких в целом периодах соответствующих ритмов их различия надо считать не случайными, а обусловленными определенными причинами. Эти причины становятся ясными, если рассматривать зависимость периода от положения разреза на тектонической структуре. Проиллюстрировать эту зависимость можно на примере Марковского поднятия (рис. V.6). По карбо-

натному осинскому горизонту Марковская структура представляет собой моноклиналь, осложненную целым рядом небольших по размерам и амплитуде поднятий. На рисунке показан разрез такого поднятия, одна из скважин расположена на его вершине, две — на крыле. Наибольший период (46 м) одного из ритмов наблюдается в скв. 11, самой удаленной от вершины; по мере приближения к вершине период уменьшается (30 м в скв. 2) и достигает минимума на вершине (28 м в скв. 15). Таким образом, период ритмов в пределах осложняющих Марковское поднятие элементов меняется в зависимости от структурного плана отложений.

Это обстоятельство позволяет картировать распределение колебательных движений разного порядка. На существование связи между периодом ритмов и тектоническим строением площади указывал А. Б. Вистелиус [19]. Он отметил, что на поднятиях значения периодов, как правило, меньше, чем в областях погружений, при этом каждому порядку колебательных движений отвечает соответствующий ритм. Картируя параметры каждого ритма, можно получить раздельные схемы распространения колебательных движений по региональным и локальным структурным элементам.

Определение типа карбонатных коллекторов и оценка их емкости

В оценке коллекторов, и в первую очередь карбонатных, самым сложным является определение их емкости. Относительно достоверные данные по этому вопросу могут быть получены с помощью геологического изучения разрезов естественных обнажений, исследования керна прямыми и лабораторными методами. Однако такие исследования, особенно в скважинах, ограничены малым выносом керна и его размерами.

Результаты существующих методов интерпретации диаграмм, записанных при геофизических исследованиях скважин, часто бывают противоречивыми и не согласуются с данными прямых геологических наблюдений. При интерпретации геофизических данных, как правило, привлекаются результаты эмпирических и аналитических зависимостей между различными типами пористости и геофизическими показателями (БМ, БЭЗ, НГМ и др.), полученные при условии аппроксимации реальной геологической среды ее идеализированным аналогом [27, 34, 66]. Разумеется, что при таком подходе конечные результаты интерпретации существенно зависят от степени соответствия выбранной физической модели реальному пласту-коллектору. Однако даже в тех случаях, когда тип коллектора достоверно известен и литологический состав пород строго однороден, по данным геофизических измерений не всегда удается однозначно определить емкостные характеристики породы.

Это объясняется тем, что реальная карбонатная порода характеризуется случайным (вероятностным) распределением объема пустот различного типа. Идеальные же среды, на основе которых выработаны аналитические зависимости, в принципе не предназначены учитывать эту случайную изменчивость емкостных параметров пород. Кроме того, множество других неконтролируемых причин (эффектов) могут настолько исказить измеряемые геофизические характеристики, что извлечь из них информацию о пористости становится невозможно, если не попытаться каким-либо способом влияние этих эффектов подавить или устранить.

Для преодоления отмеченных трудностей, вызванных случайной изменчивостью различных геологических особенностей карбонатных пород, при интерпретации данных геофизических исследований скважин целесообразно применить ту стохастическую модель и те методические разработки, которые были изложены в этой главе.

Использование стохастической модели (V.1) при интерпретации геофизических данных с целью установления типа карбонатного коллектора и оценки его емкости диктуется вполне естественными соображениями. В процессе литогенеза формирование пористости происходит под влиянием факторов двоякой природы: закономерных (детерминированных), как, например, колебание дна бассейна седиментации, и случайных (стохастических). К последним можно отнести резкое изменение трещиноватости, образование каверн в результате вторичных процессов и т. п. Развитие во времени закономерных факторов в конечном итоге приводит к тому, что распределение межзерновой пористости по разрезу карбонатных пород будет полициклическим (суперпозиция гармоник разного периода, амплитуды и фазы). Влияние случайных факторов обусловливает флуктуационную картину изменения пористости, что хорошо согласуется с поведением функции e_h. Действие постседиментационных факторов приводит к образованию дополнительных пустот вторичного происхождения (трещины, пустоты выщелачивания, перекристаллизация и т. п.).

Следовательно, если с помощью стохастической модели (V.1) провести интерпретацию геофизических регистрограмм (диаграмм методов НГМ, БЭЗ, БМ, АМ и др.), которые отражают изменение общей пористости по разрезу (сумму межзерновых, трещинных и кавернозных пустот), то можно разделить эти регистрограммы на отдельные составляющие. Закономерная (полициклическая) часть регистрограммы, по-видимому, лучшим образом будет согласовываться с изменением межзерновой (первичной) пористости, а осциллирующий остаток — с распределением суммы трещинных и кавернозных пустот.

Специфические особенности результатов геофизических методов, характеризующих карбонатные коллекторы, позволяют наметить и другой путь выделения в разрезе участков с повышенной трещинной пористостью. Так, из ряда работ по проблеме трещин-

ных коллекторов известно, что трещинная пористость карбонатных пород крайне мала и не превосходит 1% [15, 76]. Такая пористость не может существенно влиять на данные НГМ из-за слабой чувствительности этого метода. В то же время данные УЭС, полученные методом БМ или БЭЗ при соответствующем сопротивлении заполняющего трещины флюида, могут быть крайне чувствительны к изменению подобного значения трещинной пористости [27, 28, 59, 61]. Поэтому совместный анализ закономерных составляющих, выделенных при помощи стохастической модели из диаграмм указанных геофизических методов, позволяет проследить изменение суммарной пористости по разрезу, а в некоторых случаях и оценить трещинную пористость.

Распознавание и оценка проводятся на основе предположения о том, что закономерная составляющая (сумма периодических компонент), присутствующая в регистрограмме сопротивления, кроме межзерновой пористости отражает и трещинную. Закономерная составляющая в регистрограмме НГМ в основном зависит от межзерновой пористости. Сравнивая эти две закономерные составляющие, выраженные в единицах сопротивления (данные НГМ переводятся через соответствующее уравнение регрессии), можно определить тип коллектора в карбонатном разрезе по следующим критериям.

1. Если закономерные составляющие, построенные по данным НГМ и УЭС, совпадают, то пористость в интервале наблюдения относится к межзерновому типу.

2. Если значения закономерной составляющей по данным НГМ превосходят аналогичные значения по данным УЭС, то пористость в этом интервале обусловлена и трещиноватостью. Если УЭС не кажущееся, а истинное, то можно оценить реальное значение трещинной пористости общепринятыми методами [59, 62, 79].

3. Если значения закономерной составляющей по данным НГМ меньше, чем по данным УЭС, то либо поровое пространство заполнено флюидом высокого сопротивления, либо эффективная пористость меньше суммарной.

Рассмотрим примеры интерпретации геофизических данных по ряду скважин, вскрывших карбонатные разрезы ордовика в Прибалтике и верхнего мела в Восточном Предкавказье (рис. V.7). Разрезы этих скважин характеризуются сплошным выносом керна и детально изучены литологическими и петрографическими методами (исследования проведены М. Х. Булач, Л. П. Гмид, Т. В. Дорофеевой и В. Н. Калачевой). По керну и шлифам определена плотность трещин, подсчитаны параметры трещинной пористости, вторичной пористости и отдельно кавернозности. Стандартными лабораторными методами определены межзерновая пористость и карбонатность по керну. По этим данным в разрезах выделены интервалы (зоны) повышенной пористости и трещиноватости, что может служить контролем при интерпретации геофизических показаний.

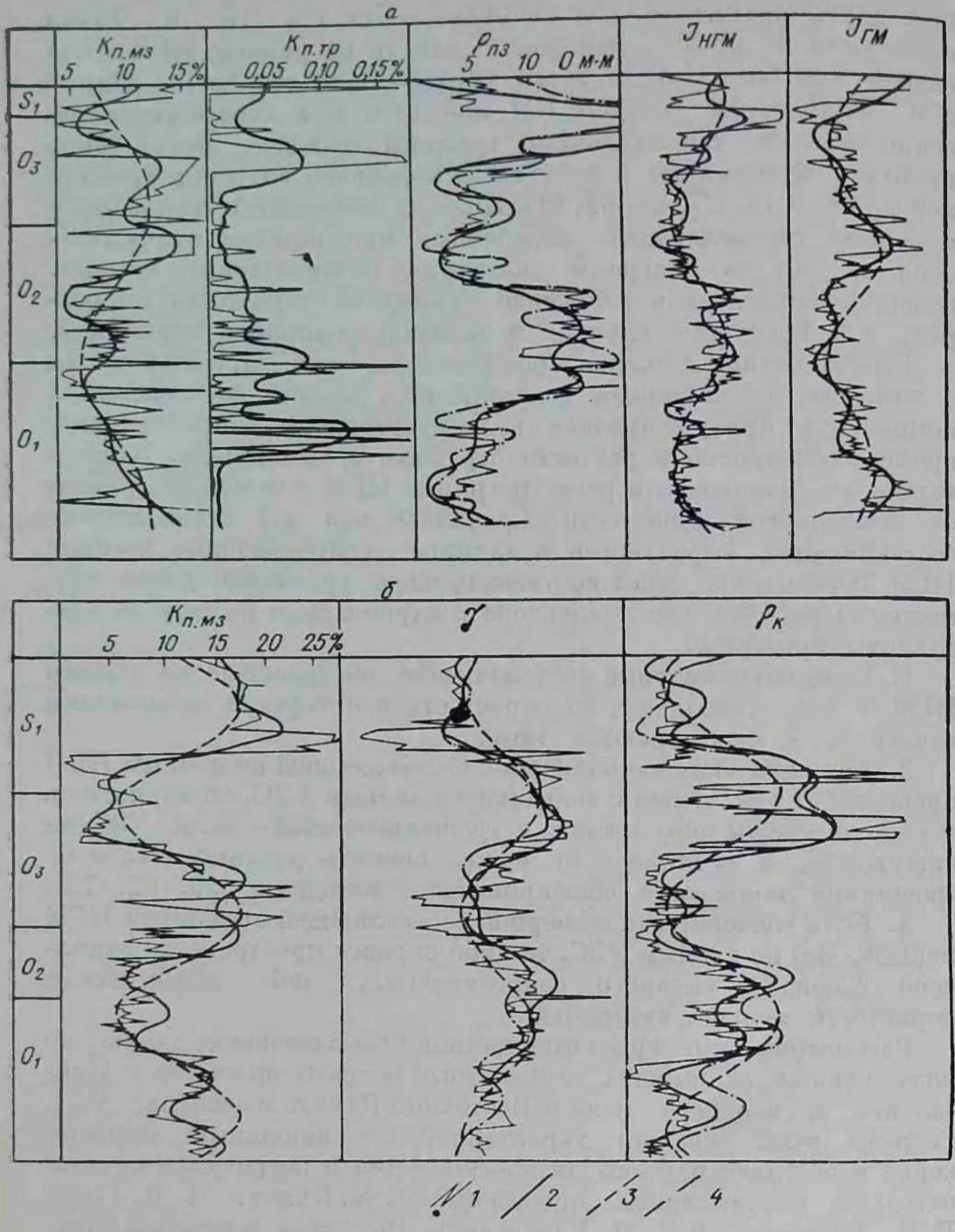
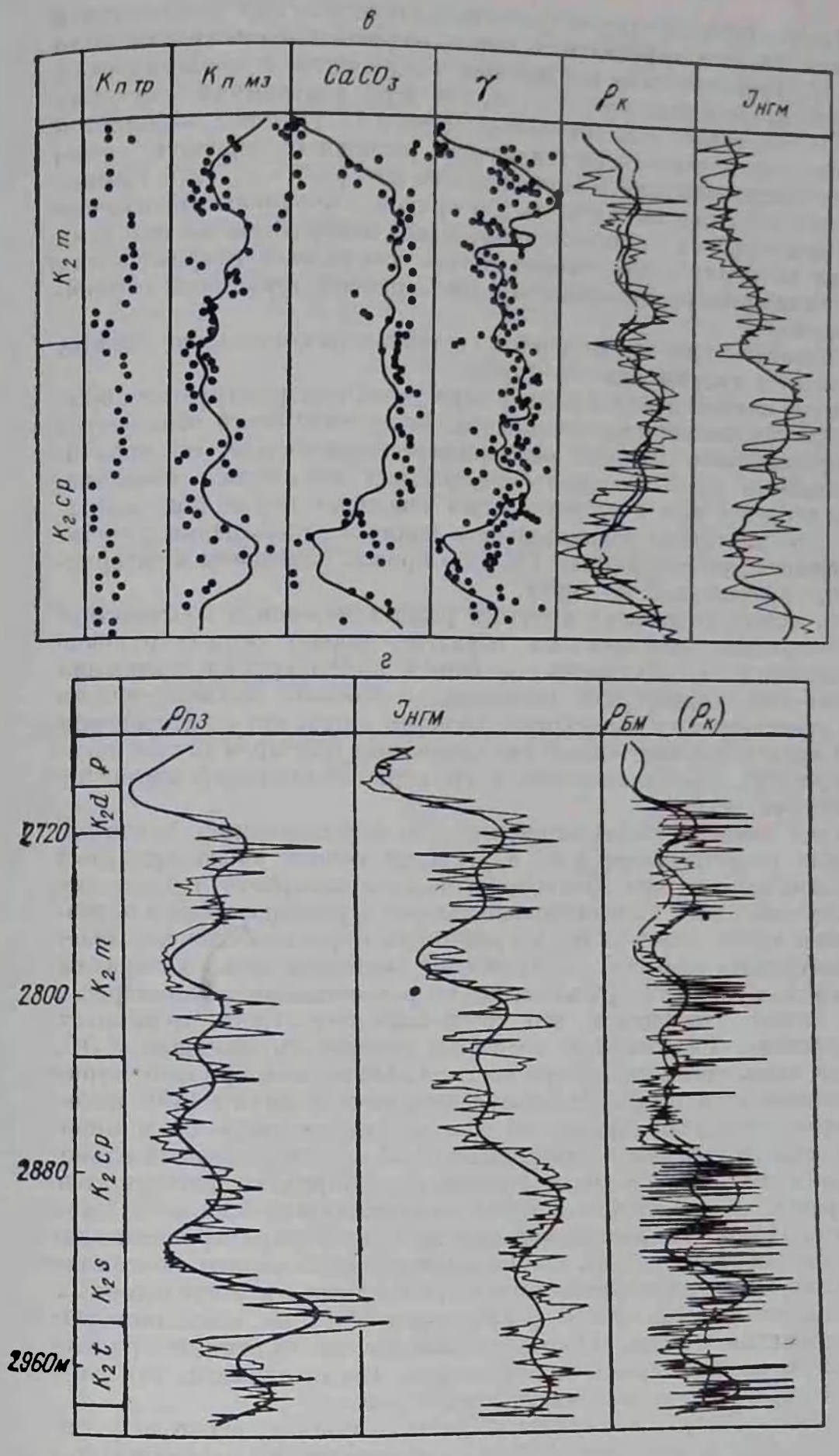


Рис. V.7. Геолого-геофизические переменные и выделенные

Скважины: *а* — 38-Дурбе, *б* — 3-Ремте.
1 — исходные геологические и геофизические переменные; *2* — полициклические составля одного периода, отмеченные во всех исходных переменных; *4* — полициклическая составля вления (рис. *а*, *в*, *г*), или полициклическая составляющая пористости, выраженная



из них периодические компоненты.

в — 2-Сергокала, *г* — 833-Малгобек.
ющие (сумма периодических компонент); *з* — периодические компоненты
ющая, выделенная из кривой НГМ и выраженная в единицах сопроти-
в единицах плотности и сопротивления (рис. *б*).

Скв. 38-Дурбе. Краткая характеристика литологических особенностей и коллекторских свойств карбонатных пород ордовикского возраста была дана в главе II. Геофизические наблюдения согласуются с геологической характеристикой разреза (см. рис. V.7, *а*). Во всех диаграммах каротажа (БЭЗ, НГМ, ГМ) на основе стохастической модели (V.1) четко выделяется небольшое число периодических составляющих различных периодов, о чем свидетельствует анализ коррело- и периодограмм (см. рис. V.1, V.2). Компоненты идентичных периодов выявляются и в кривой, показывающей изменение по разрезу межзерновой пористости (определена лабораторными методами непосредственно по керну). Это говорит о том, что данные геофизических измерений в основном несут информацию о межзерновой пористости первичного происхождения.

Более детальный анализ закономерных составляющих позволяет сделать следующие выводы о разрезе скв. 38-Дурбе.

1. Увеличение значений межзерновой пористости контролируется повышенным содержанием глинистого компонента. Вследствие этого отмечаются интервалы с относительно высокой межзерновой пористостью, но практически непроницаемые из-за большого содержания глинистого вещества. Этот вывод согласуется как с геологическим анализом керна под микроскопом, так и с результатами сравнительного анализа закономерных составляющих, выделенных по диаграммам ГМ и по кривой изменения межзерновой пористости, определенной по керну.

2. Разность между исходными кривыми разных методов и соответствующими закономерными компонентами образует сильно осциллирующие «остатки» (компоненту e_h). Выделить по общей конфигурации поведения «остатков» какие-либо характерные интервалы с особыми коллекторскими свойствами не представляется возможным. Отметим лишь, что на физические свойства пород влияет большое множество случайных факторов (в том числе и вторичные пустоты), проявляющихся с примерно одинаковой интенсивностью по всему разрезу.

3. Совместная интерпретация периодических составляющих, извлеченных из исходных регистрограмм УЭС и НГМ на основе ранее принятых критериев, свидетельствует, что практически во всем интервале наблюдения эти составляющие совпадают. Следовательно, разрез характеризуется в основном межзерновым типом пористости, а трещинная пористость не вызывает сколь-нибудь ощутимого эффекта. Заслуживают внимания лишь интервалы в середине нижнего и верхнего ордовика, где по расхождению закономерных составляющих можно утверждать, что трещинная пористость превышает свой средний уровень. Расхождение позволяет установить значение УЭС, соответствующее лишь трещинной пористости (влияние межзерновой пористости и глинистости устранено). Используя известные формулы [59], которые соответствуют идеальной среде, мы рассчитали значение трещинной пористости для этих интервалов (соответственно 0,33 и 0,79%). Относительно повышенная трещинная пористость в указанных интервалах фиксируется и по данным прямых наблюдений в шлифах (соответственно 0,26 и 0,22%).

Скв. 3-Ремте. Геолого-геофизические данные проинтерпретированы аналогичным образом (см рис. V.7, *б*). Сопоставление закономерных составляющих, выделенных из регистрограмм пористости и плотности (определенных по керну), а также электрометрии и радиометрии (ГМ), не позволяет выявить зоны интенсивной трещиноватости, так как эти составляющие практически совпадают во всем интервале исследования. Таким образом, для разреза характерна пористость лишь межзернового типа.

Скв. 2-Сергокала. Вскрыт карбонатный разрез позднемелового возраста общей мощностью 415 м (см. рис. V.7, *в*), представленный однообразной толщей известняков, в целом характеризующихся низкой глинистостью. По данным лабораторных методов и микроскопии известняки имеют невысокую межзерновую пористость (3—7%), участками интенсивно трещиноваты. Вторичная пористость выщелачивания развита участками и в целом составляет не более 0,5%. В отдельных интервалах разреза отмечаются стилолиты и трещины (открытые или выполненные кальцитом). Трещинная пористость

по данным метода шлифов составляет сотые доли процента (0,01—0,04%), трещинная проницаемость 1—3 мД. По совокупности результатов геологических методов [49] в разрезе скважины выделяются три интервала, трещиноватость пород в которых превышает фоновую. Для этих интервалов характерна интенсивная раздробленность керна открытыми минеральными трещинами и стилолитами.

Применение стохастической модели при раздельной интерпретации геофизических и геологических регистрограмм позволило в каждой переменной выявить ряд периодических компонент и остаточную флуктуационную кривую. Анализ составляющих дает право считать, что изменение межзерновой пористости контролируется в рассматриваемом разрезе изменением литологического состава. Так, доминирующие периодические компоненты равных периодов отмечаются в кривых плотности (пористости) и карбонатности (табл. V.2). В целом для разреза чем выше карбонатность, тем ниже межзерновая пористость. Сравнение периодов компонент, полученных из регистрограмм пористости (плотности), сопротивления, карбонатности и НГМ, показывает, что данные геофизических методов в основном характеризуют изменение межзерновой пористости.

Таблица V.2

Периоды гармоник разного порядка, выделенных из геологических и геофизических регистрограмм (скв. 2-Сергокала)

Регистрограмма	Периоды, м				
	I	II	III	IV	V
Плотность	142	100	76	46	28
$\rho_{\text{УЭС}}$	136	94	72	44	28
$J_{\text{НГМ}}$	120	95	75	43	28
	125 *	100 *	80 *	44 *	36 *
Карбонатность	140	95	75	44	—

* Для скв. 833-Малгобек.

Относительно большие случайные отклонения от закономерной составляющей кажущегося сопротивления, по-видимому, обусловлены фоновой трещиноватостью, которая практически одинаково развита по всему разрезу. Однако сопоставление закономерных составляющих, выделенных из регистрограмм сопротивления и НГМ, позволяет установить интервалы с относительно повышенной трещинной пористостью. Так, внизу разреза отмечается наибольшее расхождение между этими составляющими; результаты геологических наблюдений подтверждают, что эта зона характеризуется наиболее интенсивной трещиноватостью горных пород.

Скв. 833-Малгобек. Прямые геологические наблюдения по этой скважине не проводились (из-за отсутствия керна), но мы располагали довольно обширными сведениями по естественным разрезам верхнего мела Восточного Предкавказья, а также по ряду скважин месторождений Карабулак, Заманкул и Малгобек. По литологическому составу указанные разрезы крайне однообразны и сложены 300—400-метровой толщей слоистых, трещиноватых слабоглинистых известняков. Зоны повышенной трещиноватости выделяются в верхней части маастрихта и в низах кампана. Для известняков верхней части разреза свойственны вторичная пористость и кавернозность, они являются наиболее продуктивными.

В диаграммах различных геофизических методов на основе анализа периодограмм выделены периодические компоненты (см. рис. V.7, *г*). Сравнение периодов гармоник в регистрограммах, полученных по этой скважине

и скв. 2-Сергокала (см. табл. V.2), дает основание предполагать, что обнаруженные периодичности отражают циклическое изменение межзерновой пористости по разрезу.

Сравнительный анализ закономерных (полициклических) составляющих из регистрограмм НГМ и сопротивления по скв. 833-Малгобек дает наглядную картину изменения типа пористости. Так, в тех интервалах разреза, где обе составляющие почти сливаются, можно предполагать пористость межзернового типа. В интервалах, где составляющие расходятся, наряду с межзерновой пористостью отмечается и трещинная.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Нами показана возможность расщепления геологических и геофизических регистрограмм, характеризующих некоторые свойства карбонатного разреза, на составные элементы. Продемонстрировано, что стохастическая модель открывает более широкие возможности непосредственно геологического истолкования геофизических данных, т. е. она позволяет вскрыть в геофизических сведениях геологическую информацию генетического типа. Так, одним из составных элементов интегрального геофизического сигнала является его часть, приходящаяся на полициклическое поведение пористости по вертикали разреза, что связано с геологическими причинами. Предлагаемый метод позволяет устранять влияние на геофизические регистрограммы различных геологических эффектов (в частности, межзерновой пористости и глинистости) и рассматривать поведение как оставшейся части сигнала, так и устраняемой. Очевидно, что интерпретация двух составных частей общего сигнала помогает вскрыть в строении разреза общие и специфические черты. Более того, можно количественно оценить тот эффект, который вносится определенной геологической особенностью в общую изменчивость геофизической диаграммы.

Подводя итоги, надо отметить следующее.

1. Предложенный математический аппарат исследования позволяет детально изучать колебательные явления, получающие отражение в геологических и геофизических регистрограммах, характеризующих разрезы карбонатных толщ. Важно, что геофизические данные содержат в себе необходимую геологическую информацию генетического типа. Сопоставление коррело- и периодограмм для ряда характеристик обнаруживает почти полную их идентичность, свидетельствуя, что ритмы геологических и геофизических свойств тождественны. Тем самым показано, что при анализе разрезов карбонатных толщ можно использовать не только дорогостоящие сведения, получаемые по керну, но и данные различных методов геофизических исследований скважин.

2. Движения разного порядка устанавливают определенные соотношения между составом карбонатной породы и ее свойствами. Надо отметить, что существует разница во времени между моментом наступления специфических условий для осаждения данной хемогенной фазы и моментом самого осаждения. Такое запазды-

вание, видимо, связано со спецификой проявления колебательных движений в тех или иных условиях.

3. Анализ периодов гармоник в различных геолого-геофизических регистрограммах показывает, что эпохи карбонатонакопления характеризуются почти полной идентичностью колебательных движений в платформенных и геосинклинальных областях. Всюду в карбонатных толщах выделяется одинаковое число ритмов. Периоды ритмов одного порядка соизмеримы и не зависят от возраста отложений.

4. Периоды ритмов определенного порядка меняются в соответствии со структурным планом отложений. Чем выше порядок структур, тем крупнее ритм. В общем областям поднятий соответствует некоторое сокращение периода. Следовательно, классифицируя периодичности различного порядка и сопоставляя одноименные, можно не только осуществлять корреляцию разрезов, но и проводить палеотектоническую реконструкцию изучаемых осадочных толщ.

5. Периодические компоненты описывают ту часть изменчивости породы, которая связана непосредственно с процессами седиментации. Исключение этой детерминированной части из исходной регистрограммы дает «остаток», который отражает последующие вторичные процессы. Такой подход при интерпретации разного рода регистрограмм позволяет осуществить ряд конструктивных построений. Так, применительно к анализу емкости коллекторов такие построения дают возможность выделить ее составные части, связанные с межзерновой пористостью и с трещиноватостью.

Глава VI

РАСЧЛЕНЕНИЕ И КОРРЕЛЯЦИЯ РАЗРЕЗОВ СКВАЖИН НА ОСНОВЕ АНАЛИЗА КОНФИГУРАЦИИ РЕГИСТРОГРАММ

ПОСТАНОВКА И СОДЕРЖАНИЕ ЗАДАЧИ

Расчленение и сопоставление (корреляция) разрезов осадочных толщ, вскрываемых скважинами, является одной из основных задач нефтегазовой геологии и геофизики. При ее решении используется информация двух видов: сведения, получаемые при обработке керна, и данные визуального анализа диаграмм геофизических исследований скважин. Керн содержит существенную геологическую информацию, но из-за низкого его выноса и трудоемкости обработки не всегда можно составить достаточно полное представление о литологии и стратиграфии изучаемой толщи. Поэтому обычно обращаются к визуальному анализу диаграмм каротажа. Однако результаты такого анализа, даже в самых благоприятных ситуациях, часто бывают противоречивыми.

Процесс расчленения и сопоставления разрезов на основе визуального анализа диаграмм сводится к выделению на них характерных импульсов (реперов). Предполагается, что каждый репер отражает пласт определенного литологического состава. При помощи ряда геологических методов (литологических, палеонтологических) устанавливается, к какому стратиграфическому диапазону разреза приурочены реперы. Затем осуществляется поиск подобных реперов (опорных пластов) в диаграммах других скважин и одноименные реперы стратифицируются.

Иногда выделяют и зрительно запоминают какую-либо специфическую конфигурацию отрезка диаграммы. Такой конфигурации, очевидно, может соответствовать определенная последовательность чередования пластов. Затем подобную конфигурацию отыскивают в кривых каротажа соседних скважин. Если конфигурация диаграмм имеет сложную форму, приемы визуального анализа оказываются малоэффективными. Найти в таких кривых общие устойчивые реперы или их характерные специфические последовательности почти невозможно.

В целом результативность визуального анализа диаграмм геофизических методов зависит от опыта интерпретатора, степени знания им геологии исследуемого региона. Существенное значение при этом имеют опыт и интуиция специалиста. Поэтому схемы расчленения и корреляции разрезов, составленные разными интерпретаторами на одном и том же фактическом материале, часто отличаются друг от друга, и иногда существенно. Это и понятно, ибо формирование опыта и интуиции зависит от целого ряда причин, среди которых есть и крайне индивидуальные.

У интерпретатора в конечном итоге вырабатывается определенный подход — последовательность решающих правил, которые он использует при визуальном анализе диаграмм. Многие из этих правил плохо поддаются формализации, что затрудняет выработку единого алгоритма, реализация которого в виде специальной программы для ЭВМ позволила бы автоматизировать процесс расчленения и сопоставления разрезов осадочных толщ. Однако такие попытки уже предпринимаются [95].

Надо заметить, что наиболее общим моментом при визуальном анализе разного рода диаграмм являются поиски общих черт в их структурах (конфигурациях). Эта особенность может быть положена в основу алгоритма для ЭВМ, позволяющего проводить поиск идентичных «кусков» в исходных диаграммах. Ранее мы могли убедиться, что методы теории случайных функций (процессов), в частности авторегрессионный, автокорреляционный и спектральный анализы, дают возможность достаточно полно исследовать структуру любой исходной регистрограммы, на базе количественных мер вскрывать и описывать общие и специфические черты в их структуре. Наиболее эффективные результаты при этом получаются в том случае, когда исходные регистрограммы содержат периодические или квазипериодические компоненты, так

как ритмическая смена литологического состава пород или циклическое поведение пористости по вертикали разреза может свидетельствовать об аналогичном характере функционирования природного механизма (геогенератора) при седиментогенезе. Для нашей цели наиболее разумно описание регистрограмм в терминах частотного (спектрального) анализа. Ряд трансформант указанного метода, допустим, те же автокорреляционные функции или периодограммы (позже мы введем в рассмотрение ряд других трансформант), позволяют компактно описать структуру исходной регистрограммы, четко вскрыть особенности ее строения (поведения), выявить сходство и различия в структуре, например, двух регистрограмм.

МЕТОД НЕПАРАМЕТРИЧЕСКОГО ОПИСАНИЯ СТРУКТУРЫ РЕГИСТРОГРАММ (СПЕКТРАЛЬНЫЙ АНАЛИЗ)

Изложенные выше соображения позволяют предложить для расчленения и корреляции разрезов совершенно определенный алгоритм, который сравнительно легко реализуется на ЭВМ и который целесообразно использовать при анализе структуры разного рода регистрограмм. В общих чертах он выглядит следующим образом. Сначала от исходных регистрограмм, структура которых на глаз идентифицируется плохо, необходимо перейти к их аналогам (трансформантам), структура которых более простая. Затем используется либо то же визуальное сравнение и идентификация этих более простых по структуре функций, либо их математическая классификация. Если аналогами исходных регистрограмм являлись трансформанты метода теории случайных функций, то идентифицирующим выбирался метод главных компонент.

В алгоритме анализа структуры исходных регистрограмм наряду с уже известными преобразованиями использован ряд новых, смысл которых раскрывается по мере описания общего алгоритма. Предлагаемый алгоритм существенно автоматизирует процесс решения частных и общих задач расчленения и корреляции разрезов осадочных толщ на основе анализа структуры диаграмм каротажа. Единый алгоритм интерпретации разбит на ряд этапов, совокупность которых в конечном итоге представляет собой блок-схему программы, реализованной авторами с помощью алгола на БЭСМ-6.

Первый этап. Подготовка разрезов и диаграмм каротажа для обработки. Разрезы, вскрытые скважинами, разбиваются на ряд равных отрезков. Диаграммы каротажа (или иные кривые) квантуются не менее чем через 1 м нормальной мощности разреза и представляются в виде числовых (дискретных) последовательностей y_h, $h = 1, 2, 3, \ldots, n$. Число членов последовательности должно быть равно $n = 2^k$ (k — порядок массива; максимальное значение k может быть равно 12). Посредством циклического

сдвига каждый отрезок разреза может быть частично перекрыт (рис. VI.1). Каждой последовательности наблюдений, относящейся к первоначальному и перекрывающему отрезкам, приписывается порядковый номер (всего M номеров). Таким образом, общий числовой массив, введенный в ЭВМ, представляет собой матрицу наблюдений X размером $(n \times M)$.

После того как матрица X сформирована, ЭВМ выбирает из нее два вектора (размером n) по схеме y_1y_2, y_1y_3, y_1y_4, ..., y_1y_M; y_2y_3, ..., y_2y_M и т. д. до $y_{M-1}y_M$. Предварительно (как ранее для каждого вектора матрицы X) вычисляются среднее значение

$$\bar{y} = \frac{1}{n} \sum_{h=1}^{n} y_h$$

и дисперсия

$$\sigma^2 = \frac{1}{n-1} \sum_{h=1}^{n} (y_h - \bar{y})^2.$$

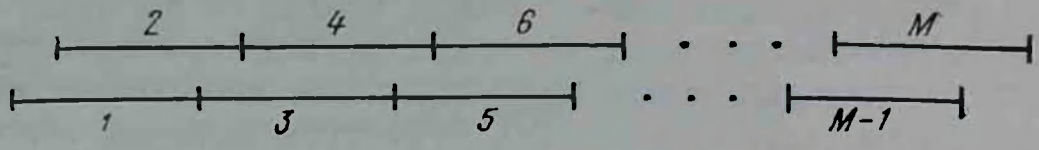


Рис. VI.1. Схема циклического сдвига.

Затем проводится центрирование и нормирование вектора Y. Иначе говоря, вместо исходной последовательности рассматривается последовательность, полученная из нее вычитанием среднего и делением на корень квадратный из дисперсии. Заметим, что такая стандартизация исходного ряда наблюдений (исходного вектора Y) является полезной процедурой, так как после нее вариация компонент векторов происходит в одном и том же диапазоне.

Второй этап. Анализ пары нормированных векторов (регистрограмм) и каждого вектора с помощью методов теории случайных функций. Сначала проводится свертка исходной информации, в результате которой от нормированных векторов переходим к их аналогам — функциям: корреляционным, взаимной корреляции, спектральной плотности, когерентности и др. Эти функции позволяют компактно описать конфигурацию любой регистрограммы, вскрыть тонкую взаимосвязь между парой регистрограмм. Ряд из этих функций намного проще по своей форме и конфигурации, чем исходные регистрограммы, и поэтому, очевидно, он может быть полнее и лучше проинтерпретирован. Математически определим функции, получаемые в процессе свертки исходной информации, и по возможности дадим каждой из них геологическую трактовку. Мы не раз убеждались в том, что анализ автокорреляционной

функции крайне полезен для вскрытия особенностей в структуре исходной регистрограммы. При $\tau = 0$ имеем $r(0) = 1$. При других значениях τ функция $r(\tau)$ варьирует в пределах от $+1$ до -1. Автокорреляционная функция характеризует изменение силы статистической линейной связи в последовательности ординат исходной регистрограммы с этими же ординатами, но сдвинутыми на шаг τ.

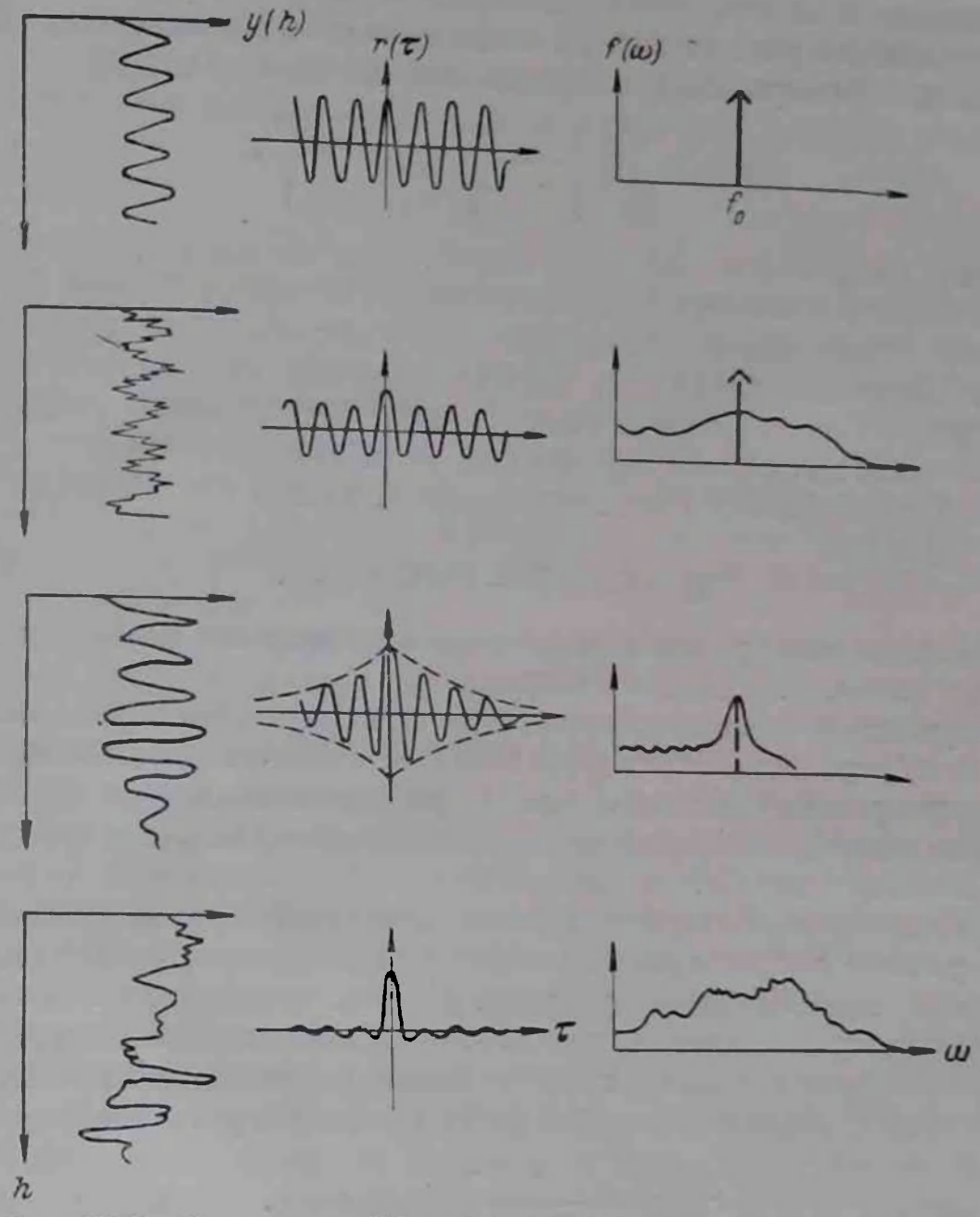


Рис. VI.2. Диаграммы ПС и соответствующие им автокорреляционные функции и спектр.

Выше мы приводили исходные регистрограммы определенной структуры и соответствующие им автокорреляционные функции. Целесообразно дать еще ряд наглядных примеров (рис. VI.2). Нетрудно убедиться, что представленные гипотетические кривые характерны для определенных режимов осадконакопления. Форма корреляционной функции существенно зависит от исходной конфигурации (структуры) диаграммы каротажа.

Обратимся к построению функции спектральной плотности, которая также заключает в себе наглядную информацию о структуре исходной регистрограммы, характеризующей разрез. Мы отмечали, что для построения этой функции обычно используют Фурье-преобразование корреляционной функции. Однако на практике иногда возникают определенные трудности математического и вычислительного характера [8, 53, 71]. Не случайно поэтому при поиске и оценке низкочастотных компонент в исходной регистрограмме мы рассматривали несколько специфичный вид периодограммы. Теперь целесообразно ввести оценку вида

$$f_{jj}(\omega) = \frac{2\pi}{n}\left|\frac{1}{2\pi}\sum_{h} y_h q_h e^{-i\omega h}\right|^2, \qquad \text{(VI.1)}$$

которая называется спектрограммой. При ее вычислении можно использовать процедуру быстрого преобразования Фурье [8, 53], которое существенно сокращает время определения спектра на ЭВМ. Весовая функция q_h вводится, чтобы приблизить оценку спектрограммы к состоятельной, ибо бесконечное увеличение длины записи исходной переменной, если она является случайной, не улучшает оценку спектрограммы. В нашем случае была принята функция

$$q_h = (1/2)\,[1 - \cos(2\pi h/n)]. \qquad \text{(VI.2)}$$

Функция спектра тонко реагирует на изменение свойств и параметров слоев, слагающих разрез.

Корреляционные и спектральные функции позволяют анализировать внутреннюю структуру каждого отдельно взятого вектора (регистрограммы). Однако можно рассмотреть и ряд функций, которые одновременно заключают в себе информацию о структуре и взаимосвязи двух случайных векторов. Пусть векторы $\mathbf{Y}_j$ и $\mathbf{Y}_k$ — центрированные и нормированные диаграммы, характеризующие либо разные разрезы одного стратиграфического диапазона, либо разные интервалы одного разреза. По аналогии с корреляционной функцией определим функцию взаимной корреляции, которая для разных значений сдвига τ дает оценки статистической линейной связи между исходными векторами (регистрограммами):

$$r_{jk}(\tau) = \frac{1}{n-\tau}\sum_{h} y_{hj} y_{(h-\tau)\,k}. \qquad \text{(VI.3)}$$

Зная эту функцию, можно также с помощью Фурье-преобразования перейти к оценке функции взаимного спектра [53]. При наличии быстрого преобразования Фурье оценку предпочтительнее получать по формуле

$$f_{jk}(\omega) = \sqrt{f_{jj}^{*}(\omega)\, f_{kk}(\omega)}; \qquad \text{(VI.4)}$$

функции $f_{jj}(\omega)$ и $f_{kk}(\omega)$ определяются из выражения (VI.1); f^* — функция комплексно сопряженная.

Анализ функции взаимного спектра показывает, как связаны между собой высоко- и низкочастотные компоненты двух разных регистрограмм.

Функция взаимного спектра является комплексной функцией частоты:

$$f_{jk}(\omega) = A(\omega) + iQ(\omega). \qquad \text{(VI.5)}$$

Вещественная часть $A(\omega)$ называется коспектром, мнимая $Q(\omega)$ — квадратурным спектром. Надо заметить, что мнимая часть взаимного спектра трудно поддается интерпретации. Поэтому вводят и рассматривают так называемую функцию когерентности

$$\mathrm{coh}_{jk}(\omega) = \frac{|f_{jk}(\omega)|^2}{f_{jj}(\omega) f_{kk}(\omega)}. \qquad \text{(VI.6)}$$

Эта функция, являясь разверткой квадрата коэффициента корреляции по частоте, отражает степень коррелированности двух исходных регистрограмм на каждой из фиксированных частот. Очевидно, что в определенном диапазоне частот значения функции могут достигать единицы и, следовательно, анализируемые регистрограммы совпадают между собой. В другом диапазоне частот значения корреляции спадают до нуля, свидетельствуя о том, что конфигурация исходных переменных в этом диапазоне частот резко различается.

Дополнительно полезно определять и рассматривать функцию разности фаз на каждой из частот:

$$\Delta\varphi(\omega) = \mathrm{arctg}\,[Q(\omega)/A(\omega)].$$

Такая функция позволяет проследить изменение фазового сдвига между компонентами одного периода, присутствующими в двух исходных переменных.

Как видим, функции, получаемые в процессе свертки исходной информации, имеют достаточно четкую геологическую трактовку. Более того, некоторые из них сходны с результатами отдельных этапов известного в геологии метода анализа мощностей. Так, если исходная диаграмма геофизического метода подробно отражает изменение литологического состава и мощности слоев в песчано-глинистом разрезе, то полученная из нее функция спектральной плотности (спектрограмма) аналогична функции распределения числа слоев определенного литологического состава по мощности. Построение последней функции при проведении анализа мощностей является весьма кропотливой работой. При отработке процедуры метода теории случайных функций на ЭВМ спектр может быть получен попутно.

Функция автокорреляции, построенная по диаграмме, характеризующей разрез песчано-глинистой толщи, может содержать два типа информации. Во-первых, она отражает характер переслаивания песчаных и глинистых слоев. Резкое падение этой функция в область отрицательных значений и знакопеременная осцилляция

кривой около оси абсцисс свидетельствуют о тонкой структуре разреза: пласты песчаников и глин маломощны (конфигурация кривой каротажа на небольшом участке разреза сильно изменчива). При достаточной мощности слоев корреляционная функция с ростом сдвига τ спадает медленнее, интервал корреляции увеличивается. Во-вторых, характер изменения функции автокорреляции позволяет судить о наличии в разрезе повторяющихся по литологическому составу и мощности пачек, в которых переслаивание слоев почти одинаковое. В этом случае автокорреляционная функция содержит косинусоидальные медленно затухающие компоненты, т. е. функционирование природного механизма, генерирующего осадочную толщу, было колебательным. Очевидно, что гармонические компоненты с таким же периодом есть и в исходных диаграммах каротажа, причем анализ спектрограммы может точно указать число и частоту этих гармоник.

Трактовка функции взаимной корреляции также не сложна. Максимумы функции при том или ином сдвиге τ показывают, что в определенных интервалах разреза структуры двух рассматриваемых диаграмм каротажа совпадают (в статистическом смысле), т. е. в сравниваемых разрезах имеются пачки с одинаковым по литологическому составу и мощности чередованием слоев. Взаимный амплитудный спектр показывает, насколько в двух кривых велики амплитуды компонент фиксированной частоты, т. е. как распределено число слоев примерно одинаковой мощности и песчанистости в двух разных разрезах.

Таким образом, все введенные нами в рассмотрение трансформанты метода теории случайных функций несут в себе ценную информацию о взаимозависимости как между отдельными ординатами исходной регистрограммы, так и между регистрограммами, характеризующими разрезы осадочных толщ. Более того, эти функции способны не только тонко описать структуру исходных переменных на основе ряда количественных мер, но и отметить характерные детали функционирования природного механизма. Для каждого исходного вектора (регистрограммы) ЭВМ по желанию исследователя может вычислить функцию взаимной корреляции, вещественную и мнимую части взаимного спектра, функции когерентности и разности фаз. Вся эта информация или ее часть выводится на печать в виде соответствующих графиков.

Третий этап. Классификация исходных регистрограмм по их спектрам с помощью метода главных компонент. Мы могли убедиться, что функции автокорреляции и спектрограммы, заключая в себе богатую информацию о структуре исходных переменных, по своей форме намного проще их. Более того, если в основе формирования структуры исходных регистрограмм лежит один и тот же механизм явления, то конфигурация этих регистрограмм в деталях может и различаться, однако функции автокорреляции и спектры в этом случае должны быть идентичны. Последний факт создает благоприятные предпосылки для классификации совокуп-

ности исходных регистрограмм, характеризующих интервалы разрезов, вскрываемых скважинами. Понятно, что после классификации увязать и проследить однотипные интервалы можно быстрее и точнее.

Для классификации были выбраны спектрограммы. Разумеется, что в качестве классификационных признаков могли быть взяты и другие трансформанты метода теории случайных функций или многомерная совокупность этих трансформант. Однако такое обилие данных, подлежащих обработке, намного осложнило бы алгоритм и сильно увеличило машинное время.

Спектрограммы, вычисленные для всех M исходных диаграмм и относящиеся к выделенным частям разреза, хранятся в памяти ЭВМ в виде матрицы F размером $(n/2 \times M)$. Появление числа $n/2$ объяснимо: преобразование (VI.1) однозначно только до граничной частоты ω_k, где $k = n/2$, поскольку этой точке соответствует частота Найквиста (т. е. частота, с которой выполняются дискретные выборки). Затем матрица F подвергается преобразованиям метода главных компонент (см. главу II). Понятно, что этот метод позволяет вскрыть общность в спектрограммах и по разным ее формам разбить исследуемую совокупность спектрограмм на ряд разных классов. Спектрограммам одного класса соответствуют идентичные регистрограммы, характеризующие те или иные части разрезов. Именно такие части разреза можно сопоставить и проследить по площади.

Таким образом, если методы теории случайных функций позволяют провести свертку исходной информации, то метод главных компонент дает возможность выполнить классификацию результатов свертки (в частности, спектров) и при этом из большой совокупности спектров выделить группы наиболее схожих между собой. Это достигается на основе количественных критериев, свидетельствующих о мере сходства спектрограмм как внутри групп, так и между ними. Кроме того, метод главных компонент позволяет получить небольшой ряд обобщенных спектров (главных компонент), которые с тем или иным весом присутствуют в каждом исходном спектре.

Все характеристики метода главных компонент, необходимые для классификации и идентификации спектрограмм (собственные числа, отношения вида $\lambda_i / \sum \lambda_i$, собственные векторы U, матрицы нагрузок W), а также графики непосредственно главных компонент (обобщенных спектров) вычисляются ЭВМ с помощью соответствующей процедуры и выводятся на печать.

Четвертый этап. Выделение закономерных (периодических) компонент из исходных регистрограмм. В связи с изучением карбонатных разрезов обсуждался вопрос о причине возникновения циклических компонент в распределении различных свойств осадка по разрезу. Природа циклических явлений в терригенных толщах может быть иной, но присутствие их обнаруживается достаточно часто [69, 78]. Функционирование геогенератора

(природного механизма), определяющего терригенную осадочную толщу, носит как закономерный, так и случайный характер, обусловливая тем самым соответствующее распределение и чередование в разрезе слоев разного литологического состава и мощности. При этом детерминированные черты проявляются в том, что свойства и параметры любого фиксированного слоя есть следствие предыдущих и причина последующих. Случайность вносит определенное искажение в детерминированный порядок чередования свойств и параметров слоев. Так, колебательному функционированию геогенератора соответствует периодическое изменение свойств слоев по вертикали разреза (что может быть обнаружено по спектро- и коррелограммам). Следовательно, поведение свойств слоев в терригенном разрезе также согласуется с моделью (V.1).

Число q доминирующих периодических компонент в регистрограмме и их частоты ω_j могут быть установлены при анализе спектрограмм. Оценки амплитуд β_j и фазового сдвига φ_j гармоник определяются по методу наименьших квадратов. ЭВМ выводит на печать параметры каждой гармоники, их распределение по разрезу, распределение суммы гармоник и компоненту e_h.

РАСЧЛЕНЕНИЕ ОСАДОЧНОЙ ТОЛЩИ И СОСТАВЛЕНИЕ КОРРЕЛЯЦИОННОЙ СХЕМЫ

Анализировались диаграммы СП, характеризующие терригенные разрезы раннемелового возраста (суходудинская свита), вскрытые глубокими скважинами в Енисей-Хатангском прогибе. Ряд скважин находится в пределах одной площади, что позволило проследить внутриплощадную корреляцию; другие скважины располагаются на разных площадях, что дало возможность прокоррелировать разрезы разных площадей. Геологическая характеристика разрезов суходудинской свиты по керну из-за крайне малого его выноса очень бедная и, конечно, не дает достаточно полного представления о литологическом составе переслаивающихся пород и об их стратиграфии. Мощность свиты в пределах рассматриваемых площадей 600—700 м. Свита заключена между двумя опорными границами, хорошо прослеживаемыми регионально: подошвой яковлевской свиты и кровлей нижнехетской. Суходудинская свита сложена песчаниками различной глинистости, алевролитами с незначительными прослоями глинистых пород. Ряд пластов являются промышленно нефтегазоносными.

Суходудинское время отличалось резко меняющейся физико-географической обстановкой, поэтому даже в пределах одной площади имели место разные режимы осадконакопления. Это обусловило специфическое чередование песчано-глинистых пород в разрезах, которые, как правило, плохо сопоставляются между собой. Чередование слоев, как отмечают многие исследователи [42], хорошо контролируется диаграммами ПС, но корреляция их также затруднена.

Разрезы скважин и диаграммы ПС были разбиты на отрезки по 128 м, последовательное перекрытие которых было выполнено по схеме, представленной на рис. VI.1. В итоге получилось 72 части, охарактеризованные квантованными диаграммами каротажа: шаг квантования равнялся 1 м нормальной мощности разреза. Таким образом составилась матрица X из $M = 72$ векторов размером n, где n — число отсчетов, снятых с диаграммы, приходящейся на один интервал; в данном случае $n = 2^k$, где $k = 7$, следовательно, $n = 128$. Для каждого отрезка диаграммы каротажа из матрицы X, пользуясь оценкой (VI.1), мы получили 72 спектрограммы.

Метод главных компонент позволил классифицировать это многообразие спектрограмм. В табл. VI.1 приведены нагрузки (коэффициенты корреляции) между главными компонентами (обобщенными спектрами) и исходными спектрами. По нагрузкам вся совокупность спектров разбивается на пять групп, в каждой из которых формы спектрограмм можно считать идентичными. Проведя обратную операцию — восстановление спектров по трем главным компонентам, мы получили интегральные спектры для каждой из пяти групп (рис. VI.3). Интегральным спектром аппроксимируется исходный спектр конкретной группы. Насколько они близки, показывают приведенные в табл. VI.1 нагрузки и кривые на рис. VI.3. Идентичность конфигурации спектрограмм не вызывает сомнений.

Таблица VI.1

Результаты группировки спектрограмм по методу главных компонент

Группа	Количество интервалов, попадающих в группу	Главные компоненты		
		I	II	III
I	29	0,97	0,00	0,00
II	13	0,91	—0,31	0,00
III	5	0,73	—0,46	0,36
IV	6	0,84	—0,41	—0,22
V	19	0,94	0,30	0,00

Диаграммы ПС, достаточно хорошо отражая изменение по разрезу литологического состава и мощности пластов, а также особенности их чередования, содержат в себе информацию о динамике осадконакопления. Очевидно, что аналогичная информация, только в более компактном и легко обозримом виде, заключается и в спектрах. Таким образом, пять типов интегральных спектров соответствуют пяти различным режимам осадконакопления, которые имели место при формировании суходудинской осадочной толщи. Узкополосность интегральных спектров показывает, что изменение значений ПС по разрезу является квазипериодическим с медленно меняющейся амплитудой. Такая конфигурация обу-

словлена соответствующим распределением литологического состава пород по разрезу.

Анализируя интегральные спектры, можно заметить, что спектры I и II групп, а также III и IV попарно сходны между собой. Поэтому целесообразно говорить лишь о трех типах спектрограмм (в дальнейшем обозначаемых A, B, C) и, следовательно, о трех режимах осадконакопления.

Тип A. Высокие значения ординат спектра концентрируются в низкочастотной области, следовательно, диаграммы каротажа

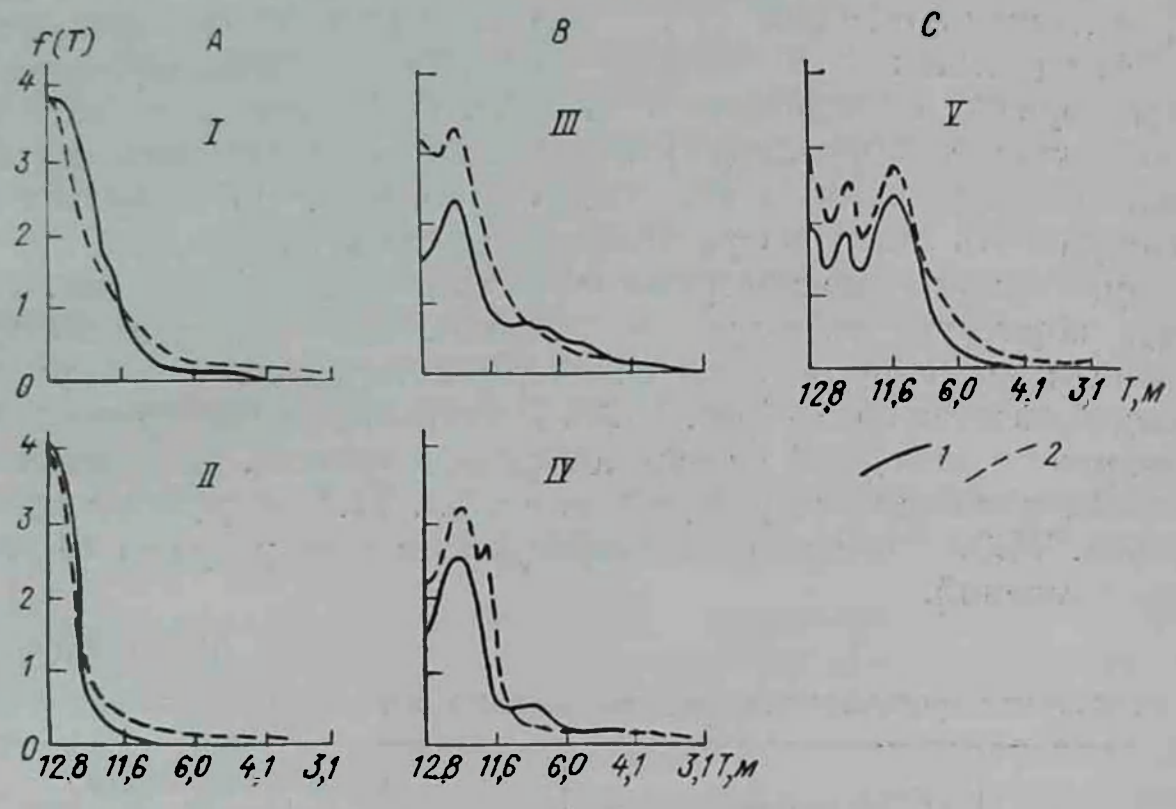


Рис. VI.3. Сопоставление исходных (*1*) и интегральных (*2*) спектров, восстановленных по главным компонентам.
I— *V*— группы (см. табл. VI.1).

содержат гармоники с длинным периодом. Таким образом, изменением свойств пород по разрезу управлял природный механизм (геогенератор), функционирование которого носило колебательный характер. Длительность этих колебаний соизмерима с мощностью рассматриваемого интервала (128 м).

Тип B. Конфигурацию диаграмм каротажа обусловливают гармоники не только длинного, но и среднего периода ($T \approx 21$ м), т. е. диапазон колебаний динамической системы здесь был значительно шире.

Тип C. Здесь интегральный спектр охватывает широкий диапазон частот. Структура диаграмм каротажа, соответствующих такому спектру, должна быть наиболее дифференцированной. Осадочный разрез, дающий подобную диаграмму каротажа, имеет сложное циклическое строение, при котором на циклы изменения физических свойств низкого порядка (длинного периода) наложен целый ряд циклов более высокого порядка (короткого периода).

Последние обусловливают чередование слоев песчано-глинистых пород мощностью не более 5 м (полуцикл). Следовательно, геогенератор в этом случае был осцилляционного типа с широким спектром колебаний.

Итак, выделив три типа спектров и сопоставив их, мы можем отметить, что каждый последующий тип спектра содержит в себе черты предыдущего, но вместе с тем обогащен частотами более высокого порядка. Унаследование показывает, что на протяжении всего времени формирования суходудинской свиты часть основных параметров, определивших функционирование геогенератора, оставалась неизменной. Иногда процесс осложнялся новыми колебаниями, которые накладывались на прежние. Все это получает отражение в конфигурации (структуре) диаграмм каротажа, которые являются суперпозицией разных по периоду гармоник (значения периодов можно установить по спектрограммам). Очевидно, что каждая гармоника фоссилизирует определенные черты функционирования природного механизма и, следовательно, в целом работа геогенератора носит сложный колебательный характер; это лишний раз подтверждает фундаментальные положения теории литогенеза [78].

При сопоставлении разрезов различных площадей (рис. VI.4) можно заметить, что строение Мессояхской, Северо-Соленинской и Пеляткинской достаточно сходно. Здесь в начале валанжина отмечается режим типа А, для которого характерно чередование песчано-глинистых пластов значительной мощности (соизмеримой с рассматриваемым интервалом). Это чередование может возникнуть в нормально-морских условиях при медленных тектонических колебаниях [42].

К концу валанжина море испытывает регрессию, в результате чего начинается обмеление морского бассейна. Но по динамике процесса перечисленные площади отличаются от Озерной. В первом случае диаграмма ПС на этой части разреза дает спектр типа С, который свидетельствует о том, что на фоне длиннопериодных колебаний устанавливались средне- и короткопериодные, т. е. функционирование геогенератора имело сложный колебательный характер. Во втором случае диаграмма дает спектр типа В. Этот режим охватывал здесь длительный временной диапазон. Таким образом, толща осадков на Озерной площади формировалась в несколько иных тектонических условиях.

Заметим, что смена режима осадконакопления наблюдается не только в вертикальном, но и в горизонтальном плане. Целесообразность именно такой корреляции интервалов с разными индексами вытекает из детального анализа взаимных корреляционных функций, взаимных спектров и особенно функций когерентности.

Детальное рассмотрение всех спектров показывает, что амплитуда гармоники с периодом 42,7 м значительна во всех диаграммах. Следовательно, колебание, обусловившее смену литологического состава пород с найденным периодом, было одним из

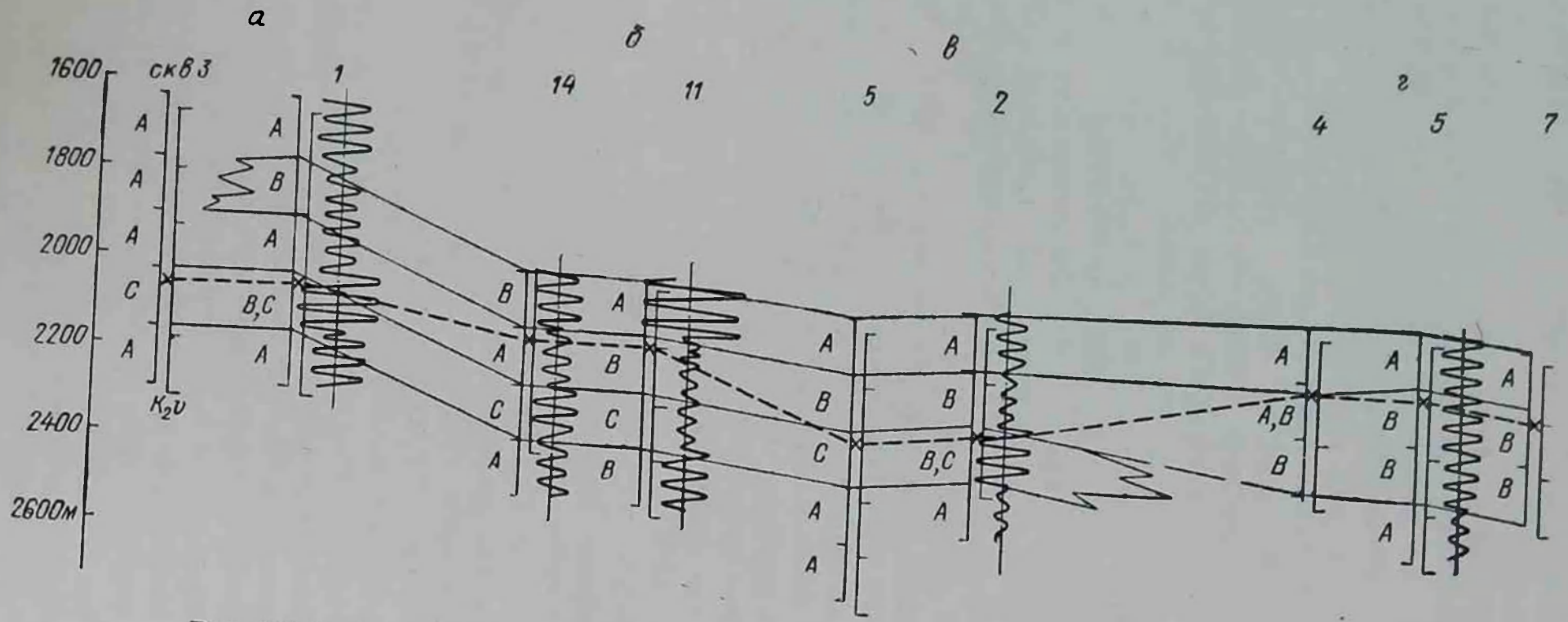


Рис. VI.4. Схема поинтервального сопоставления разрезов суходудинской свиты.
Площади: *а* — Мессояхская, *б* — Северо-Соленинская, *в* — Пеляткинская, *г* — Озерная.
Крестиками показана отметка кровли пласта СД-III по Г. Д. Гинсбургу.

основных. Поведение этой гармоники по разрезу свидетельствует, что за равный промежуток времени суммарная мощность песков и глин была примерно одинаковой. Действительно, на границах двух соседних отрезков диаграмм выделенная гармоника не претерпевает существенных разрывов, т. е. колебание этого порядка не прерывалось и запечатлелось в виде периодического изменения литологического состава пород без перерыва во всем диапазоне мощности рассматриваемых разрезов. Тем самым параметр φ в модели (V.1) является строго фиксированным, а распределение амплитуды (параметра β), вероятно, подчиняется нормальному закону. Так, на Озерной площади амплитуда колебания постепенно нарастала от подошвы суходудинской свиты к ее кровле, модель распределения свойств осадка с глубиной имеет вид

$$y_h = \beta_0 + \sum_{j=1}^{q} e^{-\alpha_j h} \cos(\omega_j h - \varphi_j),$$

где α — коэффициент затухания гармоники.

В разрезах других площадей отмечается спорадическое изменение амплитуды колебания. Это еще раз подтверждает, что режим осадконакопления на Озерной площади существенно отличался от режима соседних площадей.

Подобного типа построения, выполненные традиционными методами, дают, как известно, только границы какого-либо маркирующего пласта. Так, на рис. VI.4 показана кровля пласта СД-III, прослеженная Г. Д. Гинсбургом. Видно, что этот пласт попадает в разные режимы осадконакопления, чего, очевидно, быть не должно. Такие результаты типичны при использовании обычных методов составления корреляционных схем, основанных на прослеживании похожих по морфологии импульсов кривых ПС от пластов однотипной литологии. Чтобы избежать ошибок такого рода, целесообразно сначала выделить в толще зоны с одинаковым режимом осадконакопления и лишь затем отыскивать в них маркирующие пласты.

Итак, перечислим полученные результаты.

1. Предложен автоматизированный метод сопоставления диаграмм каротажа на основе комбинации двух статистических методов: теории случайных функций и главных компонент.

2. Применение метода при обработке и интерпретации диаграмм ПС позволило установить, что для суходудинской свиты характерны три достаточно четко различающихся между собой режима осадконакопления: A, B и C. Каждый режим дал особую структуру разреза, т. е. определенную вариацию песчанистости (глинистости) и изменение мощности слоев и пачек, что нашло отражение в соответствующей структуре диаграмм каротажа. Метод позволяет найти различие в структурах диаграмм и затем эффективно отделить один режим осадконакопления от другого.

3. Результаты использования метода свидетельствуют, что при формировании суходудинской свиты особую роль играли колебательные явления. Поэтому основные черты функционирования природного механизма (геогенератора), обусловившего исследуемую осадочную толщу, достаточно полно описываются в терминах спектрального анализа.

4. Метод позволил из достаточно большого числа колебательных явлений, совокупность которых обусловила строение разреза и, следовательно, определенную конфигурацию диаграмм каротажа, выделить одно и проследить характер его развития по разрезу. При этом показано, что за один и тот же временной интервал на разных участках региона в суходудинское время накапливались примерно одинаковые мощности осадков. По размаху (амплитуде) выделенного колебания можно провести геотектоническое районирование региона.

5. Метод целесообразно использовать в условиях, когда при наличии геофизических материалов существует острый дефицит керна.

6. В данном примере рассматривались лишь диаграммы ПС. Однако возможности программы шире: ее можно использовать при анализе комплекса геолого-геофизических материалов, т. е. в многомерном случае.

НЕКОТОРЫЕ ИНТЕГРАЛЬНЫЕ ПАРАМЕТРЫ ПРИ ОПИСАНИИ И ИЗУЧЕНИИ КОНФИГУРАЦИИ РЕГИСТРОГРАММ

Особенности конфигурации исходных регистрограмм можно оценивать и по-иному. Так, дисперсия показывает средний разброс квадрата ординат кривой от ее среднего уровня. Другим примером может служить форма эмпирической функции автокорреляции. Ее резкое падение с ростом сдвига τ и появление отрицательных корреляций при небольших значениях τ свидетельствуют о том, что исходная кривая имеет знакопеременную осцилляцию; медленное снижение функции автокорреляции указывает на относительно плавный ход исходной кривой. Таким образом, интуитивно ясно, что вычисление производных того или иного порядка в характерных точках абсциссы функции автокорреляции может дать ценную информацию, притом количественную, об особенностях конфигурации регистрограммы. Вместе с тем очевидно, что не только перечисленные характеристики могут нести информацию об особенностях конфигурации анализируемых кривых. Целесообразно поэтому ввести ряд других параметров и найти пути их оценки. Рассмотренные ниже параметры могут быть полезны как при идентификации семейства анализируемых кривых, так и при решении практических геологических задач.

Предлагаемые ниже параметры исходной кривой по своей сути являются функционалами, заданными на множестве реализаций случайной функции. Следовательно, эти параметры есть случай-

ные величины, для которых имеет смысл говорить о распределении вероятностей. Однако непосредственно аналитический вид распределения удается получить крайне редко, поэтому при решении практических задач ограничиваются поиском каких-либо характеристик распределения, например, математического ожидания и дисперсии, т. е. двух первых моментов. Даже эти простые параметры конфигурации кривой заключают в себе достаточно богатую геологическую информацию.

Приведем формулы математического ожидания [71, 90], так как они являются относительно простыми и могут быть использованы в практических расчетах. Вопрос о вероятностных параметрах случайных функций, отражающих особенности их конфигурации, поставлен сравнительно давно в теории связи [54]. Примечательно, что в грави-, магнито- и сейсморазведке также имеется

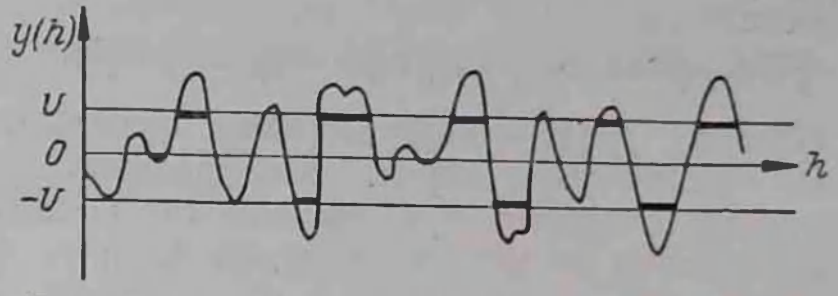


Рис. VI.5. Схема, поясняющая смысл параметра $l(U)$.

ряд решений этого вопроса [3, 24]. Однако в промысловой геофизике подобные исследования проводятся в несколько ином плане [31], хотя аналогия каротажных кривых с другими случайными функциями напрашивается достаточно явно.

1. Длительность пребывания регистрограммы выше заданного уровня. Рассчитывается функционал $l(U)$ — суммарная длина отрезков прямой, параллельной оси скважины, которые вырезаются исходной регистрограммой, когда высота (уровень) прямой равен значению U, т. е. этот функционал характеризует длину «опорной» линии кривой $y(h)$ на уровне U (рис. VI.5).

Геологическая трактовка параметра $l(U)$ очевидна. Так, если в качестве кривой $y(h)$ взять диаграмму ПС для песчано-глинистого разреза, то значение $l(U)$ покажет суммарную мощность либо пластов глин, либо пластов песчаников в зависимости от уровня U. Если рассматривается кривая пористости по разрезу (или какой-либо ее геофизический аналог, например диаграмма НГМ), то функционал $l(U)$ есть суммарная мощность пластов с пористостью выше заданной. Очевидно, что такой функционал может быть использован при подсчете запасов углеводородов, где определение общей мощности продуктивных пластов с заданными коллекторскими свойствами является одной из основных задач.

Ограничимся рассмотрением формулы математического ожидания параметра $l(U)$ в промежутке $[0, L]$. Положим исходную

функцию (регистрограмму) $y(h)$ тождественно равной $e(h)$ — стационарному нормальному процессу со средним $\mathrm{M}e(h)=0$ и с непрерывной автокорреляционной функцией

$$r(\tau)=\mathrm{M}e(h)\,e(h+\tau).$$

Тогда

$$\mathrm{M}l(U)=\left[1-\Phi\left(U/\sqrt{r(0)}\right)\right]L, \qquad \text{(VI.7)}$$

где

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp\left(-\frac{t^2}{2}\right)dt$$

— табулированная функция [71].

Выше в качестве математической модели регистрограммы применялась формула (V.1). Ограничимся здесь более простым выражением модели:

$$y(h)=\beta\cos(\omega h+\varphi)+e(h)=y_\beta(h)+e(h). \qquad \text{(VI.8)}$$

Таким образом, $y_\beta(h)$ — периодическая функция, а $e(h)$ — как и ранее, стационарная нормальная функция.

Математическое ожидание функционала $l(U)$ на единичном интервале функции $y_\beta(h)$ дается формулой [90]:

$$\mathrm{M}l_\beta(U)=\frac{1}{2\pi}\int_0^{2\pi}y_\beta(h)\,dh=\begin{cases}(1/\pi)\arccos(U/\beta) & \text{при } |U|\leqslant\beta;\\ 0 & \text{при } |U|>\beta;\\ 1 & \text{при } U<-\beta.\end{cases} \qquad \text{(VI.9)}$$

Для смешанной модели, т. е. для $y(h)$,

$$\mathrm{M}l(U)=1-\frac{1}{2\pi}\int_0^{2\pi}\Phi\left(\frac{U-\beta\cos\varphi}{\sqrt{r_e(0)}}\right)d\varphi.$$

Если отношение $\beta^2/r(0)$ достаточно мало, то

$$\mathrm{M}l(U)\approx 1-\Phi\left(\frac{U}{\sqrt{r_e(0)}}\right)+\frac{\beta^2}{4\sqrt{r_e(0)}}\frac{1}{\sqrt{2\pi r_e(0)}}\exp\left[-\frac{U^2}{2r_e(0)}\right], \qquad \text{(VI.10)}$$

где $r_e(0)$ — значение функции автокорреляции в точке $\tau=0$, рассчитанной по функции $e(h)$.

2. Среднее арифметическое отклонение ординат регистрограммы от оси стационарности. Введем в рассмотрение функционал вида

$$G_a=\frac{1}{L}\int_0^L|y(h)|\,dh, \qquad \text{(VI.11)}$$

заданный на интервале $[0, L]$. Для случайной нормальной функции $e(h)$ математическое ожидание

$$\mathrm{M}G_a=\sqrt{(2/\pi)\,r(0)}. \qquad \text{(VI.12)}$$

Надо заметить, что это выражение справедливо и для суммарной функции вида $y(h) = y_\beta(h) + e(h)$ при малом отношении $\beta^2/[2\sqrt{r_e(0)}]$. Однако для чисто гармонической компоненты

$$MG_a = \frac{\beta}{2\pi}\int_0^{2\pi} |\cos\varphi|\, d\varphi = \frac{2\beta}{\pi}\int_0^{\pi/2} \cos\varphi\, d\varphi = \frac{2\beta}{\pi}. \qquad \text{(VI.13)}$$

3. Среднее квадратическое отклонение ординат регистрограммы от оси стационарности. Рассмотрим на интервале [0, L] функционал вида

$$G_q^2 = \frac{1}{L}\int_0^L y^2(h)\, dh. \qquad \text{(VI.14)}$$

Математическое ожидание этого функционала в предположении, что исходная функция тождественна $e(h)$, дается формулой

$$MG_q^2 = r(0) = \sigma_{e(h)}^2. \qquad \text{(VI.15)}$$

Для чисто гармонической компоненты

$$MG_q^2 = \frac{1}{2\pi}\int_0^{2\pi} \beta^2 \cos^2\varphi\, d\varphi = \frac{\beta^2}{2}; \qquad \text{(VI.16)}$$

для суперпозиции $e(h)$ и гармоники $y_\beta(h)$

$$MG_q^2 = \beta^2/2 + r_e(0).$$

Эта формула справедлива, когда $y_\beta(h)$ и $e(h)$ ортогональны друг другу.

Обратим внимание, что отношение

$$\sqrt{MG_q^2}/MG_a = \sqrt{\pi/2} \approx 1{,}25,$$

когда $y(h) \equiv e(h)$; аналогично имеем

$$\sqrt{MG_q^2}/MG_a = \pi/(2\sqrt{2}) \approx 1{,}11,$$

когда $y(h) \equiv y_\beta(h)$; получаем

$$\sqrt{MG_q^2}/MG_a = \sqrt{\beta^2\pi/[4r_e(0)] + 1{,}57},$$

когда анализируется суперпозиция двух функций $y_\beta(h)$ и $e(h)$.

Следовательно, если эмпирические расчеты дают определенные значения приведенного отношения, то можно распознать, с каким классом функций мы имеем дело, и осуществить тестирование исходных функций (регистрограмм).

4. Число пересечений исходной кривой некоторого уровня. Ясно, что чем больше пересечений на единичном интервале испытывает исходная функция, тем выше ее частота. Этот функционал

может иметь практическое приложение. Так, многие исследователи отмечали, что кривые бокового или простого электрического микрозондирования в зонах повышенной трещиноватости карбонатных пород имеют более высокочастотную характеристику. Хотя эта зависимость качественная, параметр числа пересечений уровня может описать ее количественно.

Итак, математическое ожидание числа пересечений функцией $e(h)$ уровня на высоте U (для единичного интервала)

$$\mathrm{M}m(U) = \frac{1}{\pi}\sqrt{-\frac{r''(0)}{r(0)}}\exp\left[-\frac{U^2}{2r(0)}\right]. \qquad \text{(VI.17)}$$

В частности, если $U = 0$, то

$$\mathrm{M}m(0) = \frac{1}{\pi}\sqrt{-\frac{r''(0)}{r(0)}}. \qquad \text{(VI.18)}$$

Для функции $y_\beta(h)$ среднее число пересечений кривой с уровнем U (на единицу длины)

$$\mathrm{M}m(U) = \begin{cases} 2/T & \text{при } |U| < \beta; \\ 1/T & \text{при } |U| = \beta; \\ 0 & \text{при } |U| > \beta, \end{cases} \qquad \text{(VI.19)}$$

что геометрически очевидно [T — период функции $y_\beta(h)$].

Для смеси гармоники $y_\beta(h)$ с функцией $e(h)$ имеется результат Райса [90] достаточно сложного вида, поэтому ограничимся асимптотической формулой, справедливой лишь для малого уровня гармонической компоненты:

$$\mathrm{M}m(0) \approx \frac{1}{\pi}\frac{\sqrt{-r_e''(0)+\beta^2\omega^2/2}}{\sqrt{r_e(0)+\beta^2/2}}; \qquad \text{(VI.20)}$$

очевидно, что в этом случае автокорреляционная функция

$$r_y(\tau) = r_\beta(\tau) + r_e(\tau) = (\beta^2/2)\cos(\omega\tau) + r_e(\tau). \qquad \text{(VI.21)}$$

5. Среднее число выбросов функции выше заданного уровня (на единичном интервале). Если в качестве функции $y(h)$ выбрать изменение пористости по разрезу, то этот функционал будет характеризовать среднее число пластов на определенном интервале с пористостью выше уровня U.

Для функции $e(h)$

$$\mathrm{M}\nu(U) = \frac{1}{2\pi}\sqrt{\frac{-r''(0)}{r(0)}}\exp\left[-\frac{U^2}{2r(0)}\right]. \qquad \text{(VI.22)}$$

Этот функционал в 2 раза меньше, чем $\mathrm{M}\,m(U)$, что геометрически очевидно.

6. Длительность пребывания функции выше заданного уровня в течение одного выброса. Поделив функционал $l(U)$ на $\nu(U)$, получим

$$M v(U) = \pi \frac{\sqrt{r(0)}}{\sqrt{-r''(0)}} \left[1 - \Phi\left(\frac{-U}{\sqrt{r(0)}}\right)\right] \exp \frac{U^2}{2r(0)}. \quad \text{(VI.23)}$$

Если для примера взять функцию изменения пористости по разрезу, то этот функционал даст оценку мощности отдельного пласта заданной пористости.

Для расчета приведенных функционалов необходимо располагать значениями в пуле вторых производных от автокорреляционных функций. Дискретное дифференцирование эмпирических коррелограмм, особенно многократное, процесс нежелательный, так как его результаты не очень точны. Поэтому целесообразно сначала провести аппроксимацию эмпирической коррелограммы. Для этого используют ряд аналитических формул (табл. VI.2); приведенные в таблице функции спектральной плотности получены от функций автокорреляции с помощью преобразования Фурье.

Вопрос о том, какому аналитическому выражению отдать предпочтение при аппроксимации эмпирической коррелограммы, доста-

Таблица VI.2

Функции автокорреляции $r(\tau)$ и спектральной плотности $f(\omega)$

$r(\tau)$	$f(\omega)$
$e^{-\alpha\lvert\tau\rvert}$	$\frac{1}{\pi} \cdot \frac{\alpha}{(\alpha^2+\omega^2)}$
$e^{-\alpha\lvert\tau\rvert}\cos(\beta\tau)$	$\frac{\alpha(\alpha^2+\omega^2+\beta^2)}{\pi\lvert(\omega^2-\alpha^2-\beta^2)^2+4\alpha^2\omega^2\rvert}$
$e^{-\alpha\lvert\tau\rvert}\left[\cos(\beta\tau)+\frac{\alpha}{\beta}\sin(\beta\lvert\tau\rvert)\right]$	$\frac{1}{\pi} \cdot \frac{2\alpha(\alpha^2+\beta^2)}{\lvert(\omega^2-\alpha^2-\beta^2)+4\alpha^2\omega^2\rvert}$
$e^{-\alpha\lvert\tau\rvert}\left[\cos(\beta\tau)-\frac{\alpha}{\beta}\sin(\beta\lvert\tau\rvert)\right]$	$\frac{1}{2\pi} \cdot \frac{\omega^2}{\lvert(\omega^2-\alpha^2-\beta^2)+4\alpha^2\omega^2\rvert}$
$e^{-\alpha^2\tau^2}$	$\frac{1}{2\pi\sqrt{\pi\alpha}} \exp\left[-\left(\frac{\omega}{2\alpha}\right)^2\right]$
$e^{-(\alpha\tau)^2}\cos(\beta\tau)$	$\frac{1}{4\alpha^2\sqrt{\pi}}\left\{\exp\left[-\frac{(\omega+\beta)^2}{4\alpha^2}\right]+\exp\left[-\frac{(\omega-\beta)^2}{4\alpha^2}\right]\right\}$
$\begin{cases}\sum_{j=0}^{m} C_j \cos\left(\frac{j\pi}{T}\tau\right) & \text{при } \lvert\tau\rvert<T;\\ 0 & \text{при } \lvert\tau\rvert>T\end{cases}$	$T\sum_{j=0}^{m}(-1)^j C_j \frac{\omega T \sin(\omega T)}{\pi(\omega T)^2-(j\pi)^2}$

точно сложный [90]. При оценке функционалов, как уже отмечалось, необходимо знать значения производных от автокорреляционных функций в точке $\tau = 0$; отсюда напрашивается естественный вывод: аналитическое выражение хорошо должно аппроксимировать эмпирическую коррелограмму в точках, располагающихся около оси ординат. Однако при этом надо помнить, что не все приведенные в табл. VI.2 функции автокорреляции имеют вторую производную в точке $\tau = 0$, хотя при определенном подборе параметров β и α автокорреляционной функции все представленные аналитические функции достаточно близко аппроксимируют эмпирические коррелограммы, особенно вблизи оси ординат.

Вызывает интерес аналитическая автокорреляционная функция вида

$$r(\tau) = \sigma^2 e^{-\alpha\tau} [\cos(\beta\tau) + (\alpha/\beta)\sin(\beta|\tau|)], \qquad \text{(VI.24)}$$

соответствующая стационарному решению дифференциального стохастического уравнения второго порядка

$$\frac{d^2y}{dh^2} + 2g\frac{dy}{dh} + m^2y = e(h), \qquad \text{(VI.25)}$$

где g и m — постоянные, связанные с α и β следующим образом:

$$\alpha = m;\ \beta = \sqrt{g^2 - m^2};\ \sigma^2 = \pi c/(2gm^2) = r(0); \qquad \text{(VI.26)}$$

как и раньше, $e(h)$ — случайная стационарная функция, спектр которой обладает свойствами «белого шума», т. е.

$$S_e(\omega) = c = \text{const}. \qquad \text{(VI.27)}$$

Уравнение подобного типа получается, когда при отклонении системы от положения равновесия, характеризуемом величиной $y(h)$, возникает «восстанавливающая сила» $m^2y(h)$, пропорциональная $y(h)$, и сила торможения $2g(dy/dh)$, пропорциональная скорости изменения отклонения $y(h)$. Для устойчивости подобной системы необходимо, чтобы постоянная g была положительна. При этом в случае отсутствия «возмущающей силы» $e(h)$ «выходная» функция $y(h)$ будет иметь затухающий апериодический или колебательный характер в зависимости от того, будет постоянная g больше или меньше, чем m.

Как видно из выражения (VI.24), функция автокорреляции кроме дисперсии σ^2 зависит от двух параметров: α и β. Параметр α характеризует быстроту убывания корреляционной связи между ординатами исходной функции $y(h)$; отношение α/β показывает степень сходства исходной регистрограммы с синусоидой: при малом значении α/β ординаты, взятые через промежутки длины $\omega\beta$, оказываются сильно коррелированными и реализация становится похожей на гармоническую компоненту, при большом значении α/β периодичность затушевана [71].

Само дифференциальное уравнение (VI.25) практически важно, так как функционирование многих механизмов (динамических систем), и в том числе природных, может быть описано в рамках этого уравнения. Следовательно, и его стационарное реше-

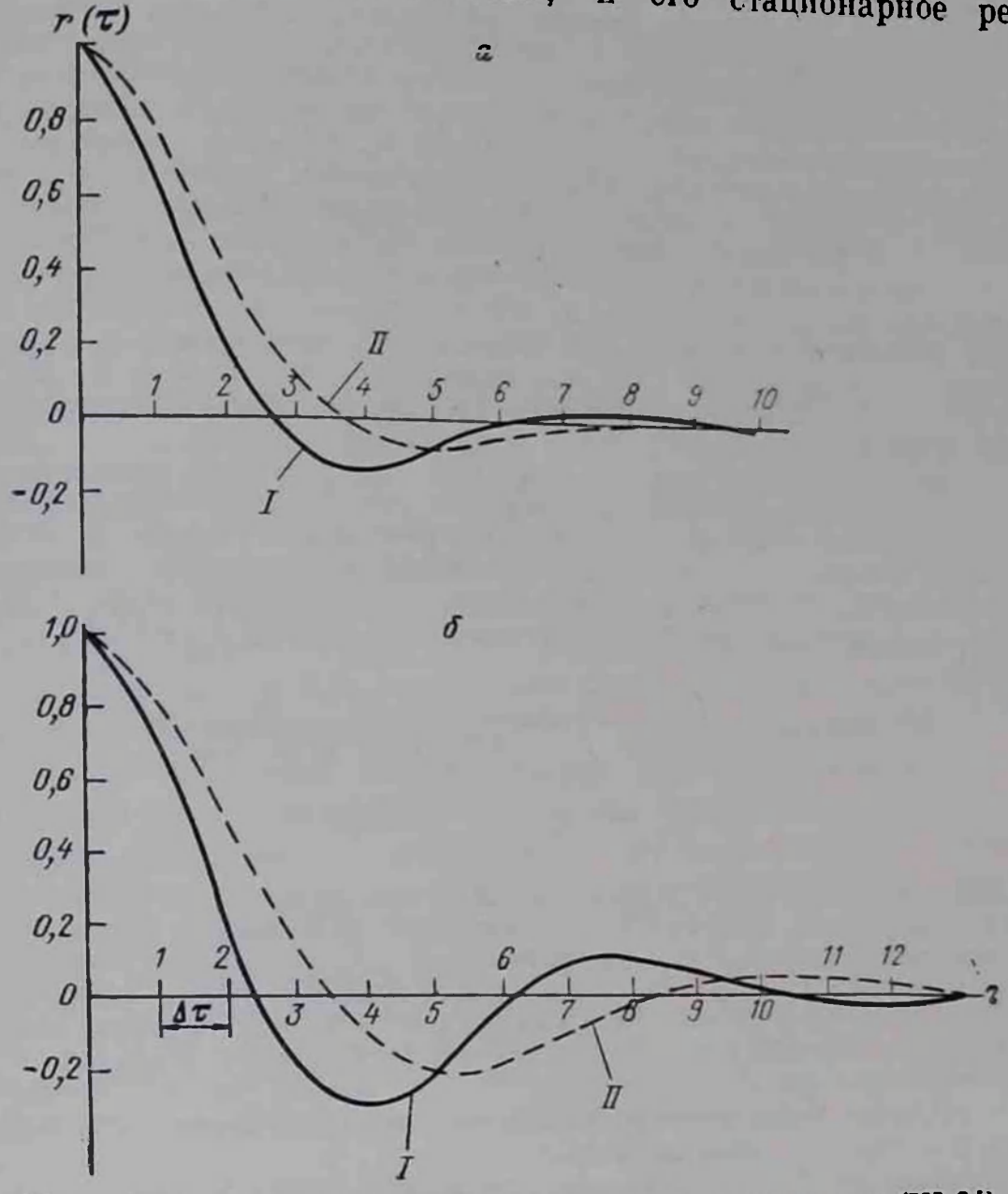


Рис. VI.6. Коррелограммы, описываемые выражением (VI.24) при разных значениях параметров α и β (палетки).

Рис.	$\Delta\tau$	α	β
а	10	0,05	0,08 (*I*); 0,06 (*II*)
	1	0,5	0,8 (*I*); 0,6 (*II*)
б	10	0,03	0,08 (*I*), 0,06 (*II*)
	1	0,3	0,8 (*I*); 0,6 (*II*)

ние — корреляционная функция вида (VI.24) — представляет практический интерес. Для быстрой оценки коэффициентов α и β эмпирических коррелограмм, которые удовлетворительно аппроксимируются выражением (VI.24), составлены палетки (рис. VI.6).

Чтобы показать, как приведенные выше результаты могут быть использованы при решении практических задач, рассмотрим следующий пример.

Исследователи, изучающие разрезы карбонатных пород позднемелового возраста Северо-Восточного Предкавказья, приводят разные значения наиболее высокой межзерновой и блоковой пористости: 10, 15, 20% [16, 76]. Указывается также, что в разрезе могут быть встречены пласты и с большей пористостью (в процессе бурения такие породы разрушаются и не дают керна). Заметим, что наличие высокопористых интервалов даже небольшой мощности существенно изменяет оценку запасов карбонатных коллекторов.

Таким образом, возникает следующий вопрос: какой может быть зафиксированная в единичном интервале разреза общая мощность коллекторов, пористость которых выше, например, 10 (15; 20)%. Интересно также оценить среднюю мощность отдельного пласта высокопористых коллекторов и среднее число таких пластов в единичном интервале разреза. Литературные источники [62, 63] и результаты наших исследований (см. главы II и V) показывают, что средние значения пористости пород как для отдельных скважин, так и для площади в целом близки и даже в общем совпадают. Поэтому примем для первого расчета 5%, для второго 7%. Стандартное отклонение рассчитано нами по оценке блоковой пористости, проведенной А. М. Нечаем [63] с использованием геофизических исследований скважин (НГМ и БМ); принимаем $\sigma^2 = 5{,}29$.

Функции автокорреляции (см. рис. V.1) содержат ряд гармонических компонент, наиболее значительной является гармоника с периодом 80 м. Оценка квадрата амплитуды этой гармоники $\hat{p}^2 = 0{,}48\,\sigma^2$, откуда $\hat{p}^2 = 2{,}54$.

Общую мощность коллекторов в единичном интервале разреза оценим по формулам (VI.10), это будет справедливо, так как величина $\hat{p}^2/\sigma^2$ достаточно мала. Для коллекторов с пористостью 10% и выше (при М $K_п = 5\%$) получаем

$$Ml\,(10) = 1 - \Phi\left(\frac{10-5}{2{,}3}\right) + \frac{2{,}54}{4\sqrt{5{,}29}} \cdot \frac{1}{\sqrt{2\pi \cdot 5{,}29}} \times$$

$$\times \exp\left[-\frac{(10-5)^2}{2 \cdot 5{,}29}\right] = 0{,}0119.$$

Следовательно, в 100 м разреза общая мощность пластов с пористостью 10% плп выше равна 1,19 м. При тех же условиях и М$K_п = 7\%$ аналогичная мощность составит 8,7 м. При уровне $U = 15\%$ в первом случае $l\,(U) = 1{,}63 \cdot 10^{-6}$, во втором $1{,}85 \cdot 10^{-3}$. Полученные значения крайне низки и проводить аналогичные расчеты для уровня пористости 20% не имеет смысла. Следовательно, вероятность зафиксировать в карбонатных разрезах интервалы коллекторов с пористостью даже 15% (не говоря уже о 20%) крайне мала. Если и есть такие прослои, то они настолько маломощны, что практического интереса не представляют.

Гармоника указанной амплитуды при расчете функционала $l\,(U)$ вносит малый вклад, поэтому не будет большой ошибки, если исходную реализацию кривой пористости по разрезу положить тождественной функции $e\,(h)$. Соответствующая ей эмпирическая функция автокорреляции может быть аппроксимирована выражением вида (VI.24). Палеточным методом находим, что коэффициенты $\alpha = 0{,}035$ и $\beta = 0{,}133$.

Для оценки среднего числа выбросов функции $e\,(h)$ за уровень $U = 10\%$ на единичном интервале наблюдения (или, что то же самое, для оценки среднего числа пластов заданной пористости) найдем вторую производную функции автокорреляции:

$$\frac{d^2}{d\tau^2} r\,(\tau)\,|_{\tau=0} = -\sigma^2\,(\alpha^2 + \beta^2).$$

По формуле (VI.22) имеем (при М$K_п = 7\%$)

$$M_{1/7}\,(U) = \frac{1}{2\pi}\sqrt{\frac{5{,}29\,[0{,}035^2 + 0{,}133^2]}{5{,}29}}\exp\left[-\frac{(10-7)^2}{2 \cdot 5{,}29}\right] = 0{,}0093,$$

т. е. на 100 м разреза надо ожидать примерно один выброс.

Оценка функционала (VI.23) дает

$$Mv(U)=\frac{\pi}{0{,}137}\left[1-\Phi\left(\frac{10-7}{2{,}3}\right)\right]\exp\frac{-(10-7)^2}{2\cdot 5{,}29}=0{,}52.$$

Роль введенных в рассмотрение функционалов достаточно ясна. Они позволяют не только описать поведение исходных случайных функций по разрезу, вскрыть особенности их конфигурации, но и получить ряд численных параметров, имеющих вполне определенный геологический смысл. Более того, функционалы представляют собой своеобразные меры, сравнение которых весьма полезно при решении ряда задач.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Предложенные два метода описания конфигурации регистрограмм, хотя и используют одни и те же понятия теории вероятностей и математической статистики, имеют разное назначение. Метод спектрального анализа вскрывает структуру диаграмм, показывая тем самым, какие гармонические составляющие в ней доминируют. Этот метод не только позволяет идентифицировать и рассортировать серию регистрограмм, конфигурация которых на глаз не различается, но и дает основание воссоздать механизм явления, обусловившего подобную серию кривых.

Метод описания конфигурации регистрограмм с помощью интегральных параметров более формален. Он дает оценку конфигурации диаграммы в целом на основе определенных численных стохастических мер, причем каждая из них имеет достаточно четкую геологическую трактовку. Имея ряд параметров, описывающих конфигурацию регистрограмм с разных сторон, можно по этой многомерной совокупности классифицировать их, используя тот или иной алгоритм задачи распознавания образов.

ЗАКЛЮЧЕНИЕ

При комплексной обработке и интерпретации геолого-геофизической информации целесообразно применять подход, существенно отличный от общепринятого. Этот подход базируется на статистико-вероятностных методах математической теории оптимального эксперимента. Исходная информация как бы расщепляется на две основные части: детерминированную (закономерную) и случайную (стохастическую). Этапу расщепления информации предшествует этап ее свертки (оценка средних, стандартов, коэффициентов корреляции, факторных нагрузок, спектров и т. п.). При свертке вскрываются характерные черты информации, контрастно подчеркивается, что есть общего (закономерного) в совокупности наблюдений, а что частного (случайного).

Рассматривая определенным образом дисперсии геолого-геофизических показателей, мы выяснили, какие из геологических факторов воздействуют на геофизические характеристики, а какие не вызывают их ощутимого изменения (главы I, II, V). При анализе коэффициентов корреляции между геолого-геофизическими показателями установлено, что эти коэффициенты дают возможность построить регрессионные уравнения, связывающие одни характеристики с совокупностью других, т. е. они позволяют составить многомерные уравнения регрессии (главы I, II). Более того, определенная трансформация матрицы коэффициентов корреляции приводит к матрице нагрузок (глава II), по которой можно сравнительно быстро оценить характер и тесноту взаимозависимостей между наблюдениями разной геолого-геофизической природы и, что самое главное, вскрыть геологические факторы, предопределяющие согласованное (взаимосвязанное) поведение исходных геофизических показателей.

Анализ периодограмм (спектров) позволяет увидеть, из каких основных элементов складывается конфигурация исходных геолого-геофизических кривых, описывающих разрезы осадочных толщ. Такой анализ весьма полезен при исследовании распределения коллекторов разного типа, а также при расчленении и корреляции осадочной толщи.

Надо отметить, что разные математические модели не составляют друг другу конкуренции. Наоборот, каждая модель помогает

обнаружить свои специфические стороны сложного геологического явления. Так, при оценке емкости карбонатных коллекторов мы выбрали модель регрессии по методу главных компонент. Разные типы пористости были оценены с помощью факторной модели (глава II). Изучение периодического поведения пористости по разрезу карбонатных отложений было осуществлено на основе полициклической модели (глава V).

Использование математических моделей существенно расширило границы интерпретации геолого-геофизических наблюдений. Появилась возможность извлекать из геофизических данных геологическую информацию генетического типа, что общепринятыми способами в принципе не предусматривается. Так, периодические компоненты диаграмм отражают колебательные движения дна бассейна седиментации (главы IV, V). Выделенная из комплекса геофизических наблюдений первая главная компонента тесно связана с межзерновой пористостью первичного происхождения (глава II). Казалось бы, незначительное генетическое условие — зависимость особенностей последующего слоя от свойств предыдущего — позволило построить и применить такие модели, которые дали возможность разбить общую пористость на составные части: первичную и вторичную (глава IV).

Очевидно, что генетическая информация крайне важна при решении разнообразных задач нефтегазовой геологии и геофизики. На генетической основе можно коррелировать разрезы скважин (глава VI), разделять общую пористость на ее составные элементы (главы IV, V), изучать формирование коллекторов (главы II, V), выяснять условия образования осадочных толщ (главы V, VI).

Нетрадиционные методы оказались плодотворными потому, что при анализе диаграмм каротажа обращается внимание не только на параметры их отдельных импульсов (как это делается при общепринятых методах обработки), но и на всю диаграмму в целом.

СПИСОК ЛИТЕРАТУРЫ

1. *Альпин Л. М.* К теории электрического каротажа буровых скважин. М., 1958. 87 с.

2. *Андерсон Т.* Введение в многомерный статистический анализ. М., Физматгиз., 1963. 500 с.

3. *Андреев Б. А., Клушин И. Г.* Геологическое истолкование гравитационных аномалий. Л., Гостоптехиздат, 1962. 495 с.

4. *Багов М. С., Цой В. И., Ремнев Б. Ф.* Оценка физических свойств кернов трещиноватых пород. — «Труды Грозн. нефт. науч.-иссл. ин-та, 1961, вып. 10, с. 161—171.

5. *Багров Н. А.* О некоторых особенностях корреляционного анализа и его применение к прогнозам погоды. — «Метеорология и гидрология», 1968, № 1, с. 17—27.

6. *Баринова О. И., Холин А. И., Элланский М. М.* Некоторые аспекты использования математических методов классификации геологических объектов и оценки их параметров по косвенным данным. — «Изв. вузов. Нефть и газ», 1969, № 11, с. 3—7.

7. *Белоусов В. В.* Основные вопросы геотектоники. М., Госгеолтехиздат, 1954. 606 с.

8. *Бокс Дж., Дженкинс Г.* Анализ временных рядов, прогноз и управление. М., «Мир», 1974. 406 с.

9. *Болтышев Н. Н.* Литологические особенности верхнемеловых отложений Малгобекско-Вознесенской площади. — В кн.: Геология и нефтегазоносность Восточного Предкавказья. Вып. 17. М., «Недра», 1964, с. 191—204.

10. *Брагина Л. П.* Вопросы методики изучения нефтегазоносности карбонатных коллекторов верхнемеловых отложений ЧИ АССР и Ставропольского края по комплексу промыслово-геофизических исследований. Автореф. дис. Грозный, 1969. 23 с.

11. *Булач М. Х.* О трещиноватости меловых отложений Чечено-Ингушской АССР в связи с изучением коллекторских свойств. — «Труды Всесоюз. нефт. науч.-иссл. геол.-развед. ин-та», 1961, вып. 165, с. 89—128.

12. *Булач М. Х., Деч В. Н.* Факторы, определяющие изменчивость геолого-геофизических показателей, характеризующих карбонатные породы. — Там же, 1974. вып. 351, с. 69—82.

13. *Вальд А.* Последовательный анализ. М., Физматгиз, 1960. 328 с.

14. *Вапник В. Н., Червоненкис А. Я.* Теория распознавания образов. М., «Наука», 1974. 415 с.

15. *Васильев В. М.* Комплексное использование существующих методов оценки емкости трещинных коллекторов на примере верхнемеловой залежи нефти месторождения Малгобек-Вознесенка (ЧИ АССР). — «Нефтегаз. геология и геофизика», 1968, № 10, с. 13—15.

16. *Везирова А. Ф., Вендельштейн Б. Ю.* Влияние нерастворимого остатка на формирование емкости верхнемелового карбонатного коллектора. — Там же, 1968. № 12, с. 20—25.

17. *Винер Н.* Кибернетика. М., «Сов. Радио», 1958. 215 с.

18. *Вистелиус А. Б.* Материалы к литостратиграфии продуктивной толщи Азербайджана. М., Изд-во АН СССР, 1961. 157 с.

19. *Вистелиус А. Б.* Фазовая дифференциация палеозойских отложений Среднего Поволжья и Заволжья. М.—Л., Изд-во АН СССР, 1963. 204 с.

20. *Вистелиус А. Б., Романова М. А.* Об основных факторах, определяющих состав современных песков Заунгузских Каракумов. «Докл. АН СССР», т. 188, 1969, № 1, с. 1317—1320.

21. *Гмид Л. П.* Литолого-петрографические исследования карбонатных коллекторов нефти и газа. — «Труды Всесоюз. нефт. науч.-исслед. геол.-развед. ин-та», 1968, вып. 264, с. 44—62.

22. *Гмид Л. П., Звоницкая И. В.* Литолого-петрографические исследования карбонатных пород и формирование в них пористости. — Там же, 1965, вып. 242, с. 49—73.

23. *Гмид Л. П., Леви С. Ш.* Атлас карбонатных пород-коллекторов. Л., «Недра», 1972. 80 с. (Труды Всесоюз. нефт. науч.-исслед. геол.-развед. ин-та, вып. 313).

24. *Гольцман Ф. М.* Статистические модели интерпретации. М., «Наука», 1971. 328 с.

25. *Горюнов И. И.* Некоторые результаты методических исследований трещинных коллекторов по их электрическому сопротивлению. — «Труды Всесоюз. нефт. науч.-исслед. геол.-развед. ин-та», 1962, вып. 193, с. 148—168.

26. *Горюнов И. И.* Геофизические исследования трещиноватости горных пород. — Там же, 1968, вып. 264, с. 134—150.

27. *Горюнов И. И., Молотков Л. А.* Электрическое сопротивление трещиноватой породы и его зависимость от коллекторских свойств. — Там же, 1961, вып. 165, с. 237—247.

28. *Горюнов И. И., Кузнецов Ю. И., Деч В. Н.* Геофизические исследования закономерностей развития трещиноватости горных пород. — Там же, 1965, вып. 242, с. 273—289.

29. *Губерман Ш. А.* Использование обучающих программ для решения геологических задач. — «Труды Моск. ин-та нефтехим. и газ. пром-сти», 1964, вып. 62, с. 3—15.

30. *Губерман Ш. А., Извекова М. Л.* О выделении горизонтов и определении водонефтяных контактов в многопластовых залежах нефти по геофизическим данным. — «Докл. АН СССР», т. 205, 1972, № 3, с. 570—574.

31. *Губерман Ш. А., Овчинникова М. И.* О некоторых возможностях использования статистических характеристик геологических разрезов. — «Изв. АН СССР. Сер. геофиз.», 1964, № 7, с. 1021—1030.

32. *Гублер Е. В., Генкин А. А.* Применение непараметрических критериев статистики в медико-биологических исследованиях. Л., «Медицина», 1973. 141 с.

33. *Дахнов В. Н.* Интерпретация результатов геофизических исследований разрезов скважин. М., «Недра», 1972. 365 с.

34. *Дебрант Р.* Теория интерпретации результатов геофизических методов исследования скважин. М., «Недра», 1972. 288 с.

35. *Деч В. Н.* Опыт применения вероятностной модели при интерпретации геолого-геофизических скважинных исследований трещинных коллекторов нефти и газа. — «Геология нефти и газа», 1975, № 3, с. 61—68.

36. *Деч В. Н., Кноринг Л. Д.* Опыт применения вероятностных моделей при интерпретации геолого-геофизических данных, характеризующих трещинные коллекторы. — «Труды Всесоюз. нефт. науч.-исслед. геол.-развед. ин-та», 1973, вып. 326, с. 64—80.

37. *Дженкинс Г., Ваттс Д.* Спектральный анализ и его приложения. М., «Мир», 1972. Вып. 1. 316 с. Вып. 2. 287 с.

38. *Заборовский А. И.* Электроразведка. М.—Л., Гостоптехиздат, 1943. 444 с.

39. *Зверев Г. Н.* Организующая система автоматической непрерывной интерпретации промыслово-геофизических данных на цифровых вычислительных машинах. — В кн.: Математические методы в газонефтяной геологии и геофизике. М., «Недра», 1972, с. 81—95.

40. *Изотова Т. С.* К вопросу оценки пористости трещинно-кавернозных пород. — В кн.: Методика выделения и оценки коллекторов методами промысловой геофизики и гидродинамическими исследованиями скважин. М., 1965, с. 63—80. (Всесоюз. науч.-исслед. геол.-развед. нефт. ин-т).

41. *Калачева В. Н.* О распределении карбонатных (трещинных) коллекторов и их покрышек на локальных поднятиях Балтийской синеклизы. — «Труды Всесоюз. нефт. науч.-исслед. геол.-развед. ин-та», вып. 351, с. 124—135.

42. *Карцева Г. Н., Ронкина З. З., Шаровская Н. В.* Сопоставление юрских и меловых отложений западной и восточной частей Енисей-Хатангского прогиба. — В кн.: «Енисей-Хатангская нефтегазоносная область. Л., 1974, с. 33—37. (Науч.-исслед. ин-т геологии Арктики).

43. *Кендалл М. Дж., Стьюарт А.* Статистические выводы и связи. М., «Наука», 1973. 899 с.

44. *Кноринг Л. Д.* Математические методы при изучении механизма образования тектонической трещиноватости. Л., «Недра», 1969. 86 с.

45. *Кноринг Л. Д., Деч В. Н.* Изучение колебательных движений методами скрытых периодичностей. — «Сов. геология», 1972, № 5, с. 81—93.

46. *Кноринг Л. Д., Деч В. Н.* К вопросу об определении пористости карбонатных пород. — «Докл. АН СССР», т. 202, 1972, № 4, с. 819—822.

47. *Крамбейн У., Грейбилл Ф.* Статистические модели в геологии. М., «Мир», 1969. 397 с.

48. *Крамбейн У., Кауфман М., МакКеммон Р.* Модели геологических процессов. М., «Мир», 1973. 150 с.

49. *Кузнецов Ю. И., Булач М. Х.* Влияние некоторых геологических факторов на физические свойства горных пород. — «Труды Всесоюз. нефт. науч.-исслед. геол.-развед. ин-та», 1970, вып. 290, с. 95—112.

50. *Кузнецов Ю. И., Позиненко Б. В.* Скорость распространения упругих колебаний в изотропных трещиноватых пористых средах. — Там же, 1964, вып. 228, с. 190—196.

51. *Кулинкович А. Е.* О критериях тесноты связи между физическими свойствами. — В кн.: Методика геофизических исследований. Киев, «Наукова думка», 1972, с. 75—79.

52. *Кулинкович А. Е.* Основы машинной интерпретации каротажных диаграмм. Киев, «Наукова думка», 1974. 188 с.

53. *Курьянов Б. Ф., Медведева Л. Е.* Гармонический анализ стационарных случайных процессов. Вып. 8. М., 1970. 47 с. (ВЦ Моск. ун-та).

54. *Левин Б. Р.* Теоретические основы статистической радиотехники. М., «Сов. радио», 1969. 507 с.

55. *Лоули Д., Максвелл А.* Факторный анализ как статистический метод. М., «Мир», 1967. 144 с.

56. *Мамяшин В. Г., Сонич В. Н., Романов Е. А.* О роли полевых шпатов в формировании свойств полимиктовых песчаников. — «Труды Зап.-Сиб. науч.-исслед. геол.-развед. нефт. ин-та», 1975, вып. 98, с. 156—167.

57. *Математическая* геология. Реферативный систематический указатель основной литературы по 1968 год. Ред. А. Б. Вистелиус. Л., 1969. 246 с. (Отд. библиогр. АН СССР).

58. *Математические* методы в газонефтяной геологии и геофизике. М., «Недра», 1972. 207 с. Авт.: М. М. Элланский, А. И. Холин, Г. Н. Зверев и др.

59. *Методика* изучения трещиноватости горных пород и трещинных коллекторов нефти и газа. Л., «Недра», 1969. 129 с. (Труды Всесоюз. нефт. науч.-исслед. геол.-развед. ин-та, вып. 276).

60. *Налимов В. В.* Теория эксперимента. М., «Наука», 1971. 207 с.

61. *Нечай А. М.* Оценка продуктивности и коллекторских свойств трещиноватых карбонатных пород. — «Прикл. геофизика», 1960, вып. 26, с. 70—80.

62. *Нечай А. М.* Вопросы количественной оценки вторичной пористости трещиноватых коллекторов нефти и газа. — «Прикл. геофизика», 1964, вып. 38, с. 201—212.

63. *Нечай А. М.* Изучение трещинных коллекторов методами промысловой геофизики. — «Развед. геофизика», 1969, № 36, с. 111—127.

64. *Обухов А. М.* О статистических ортогональных разложениях эмпирических функций. — «Изв. АН СССР, Сер. геофиз.», 1960, № 3, с. 432—439.

65. *Паасикиви Л. Б., Дукальская И. М.* Возможность нефтеобразования в карбонатных формациях ордовика и силура Прибалтики. — В кн.: Нефтепроизводящие свиты в карбонатных формациях». Л., «Недра», 1971, с. 203—259.

66. *Пирсон С. Д.* Справочник по интерпретации данных каротажа. М., «Недра», 1966. 413 с.

67. *Ронов А. Б.* Некоторые общие закономерности развития колебательных движений материков. — В кн.: Проблемы геотектоники. М., Госгеолтехиздат, 1961, с. 118—164.

68. *Рутман А. М.* Вопросы количественной интерпретации материалов промысловой геофизики с помощью многомерного регрессионного анализа и ЦЭВМ. Автореф. канд. дис. Баку, 1973. 26 с.

69. *Рухин Л. Б.* Основы литологии. Л.—М., Гостоптехиздат, 1953. 671 с.

70. *Ручкин А. В.* Об эффективности метода «двух растворов» при изучении трещинных коллекторов. — «Нефтегаз. геология и геофизика», 1973, № 4, с. 49—52.

71. *Свешников А. А.* Прикладные методы теории случайных функций. М., «Наука», 1968. 463 с.

72. *Семенов А. С.* Влияние структуры на удельное сопротивление агрегатов. — «Материалы Всесоюз. нефт. науч.-исслед. геол.-развед. ин-та». 1948, сб. 12, с. 43—61.

73. *Сербуленко М. Г.* Линейные методы разделения потенциальных полей. — В кн.: Приложение некоторых методов математики к интерпретации геофизических данных. Новосибирск, «Наука», 1967, с. 3—67.

74. *Серебренников М. Г., Первозванский А. А.* Выявление скрытых периодичностей. М., «Наука», 1965. 244 с.

75. *Сиротенко О. Д.* Многомерный регрессионный анализ как метод исследования связей урожая сельхозкультур с гидрометеорологическими факторами. — «Метеорология и гидрология», 1969, № 12, с. 3—9.

76. *Смехов Е. М.* Теоретические и методические основы поисков трещинных коллекторов нефти и газа. Л., «Недра», 1974. 199 с.

77. *Сохранов Н. Н.* Обработка и интерпретация промыслово-геофизических данных на ЦЭВМ. — «Прикл. геофизика», 1967, вып. 54, с. 92.

78. *Страхов Н. М.* Основы теории литогенеза. В 3-х т. М., Изд-во АН СССР. Т. 1. 1960. 212 с. Т. 2. 1960. 574 с. Т. 3. 1962. 550 с.

79. *Тхостов Б. А., Везирова А. Д., Вендельштейн Б. Ю.* Современная методика подсчета запасов нефти в трещинных карбонатных коллекторах. М., «Недра», 1972. 221 с.

80. *Уилкс С.* Математическая статистика. М., «Наука», 1968. 632 с.

81. *Урбах В. Ю.* Математическая статистика для биологов и медиков. М., «Медицина», 1963. 208 с.

82. *Урбах В. Ю.* Дискриминантный анализ: основные идеи и приложения. — «Стат. методы классификации», 1969, вып. 1, с. 79—173. (Моск. гос. ун-т).

83. *Усенко Ю. Н.* О влиянии некоторых факторов на интервальное время и его связь с пористостью пород. — В кн.: Методика геофизических исследований на Украине. Киев. «Техника», 1975, с. 144—152.

84. *Федоров В. В.* Теория оптимального эксперимента. М., «Наука», 1971. 312 с.

85. *Халфин Л. А.* Информационная теория интерпретации геофизических исследований. — «Докл. АН СССР», т. 122, 1958, № 6, с. 1007—1011.

86. *Ханин А. А.* Породы — коллекторы нефти и газа и их изучение. М., «Недра», 1969. 366 с.

87. *Харман Х. Х.* Современный факторный анализ. М., «Статистика», 1972. 392 с.

88. *Хейсканен К. П.* Некоторые черты динамики осадконакопления в средне- и верхнеятулийском бассейне Центральной Карелии. — «Сов. геология», 1964. № 12, с. 58—59.

89. *Химмельблау Д.* Анализ процессов статистическими методами. М., «Мир», 1973. 957 с.

90. *Хусу А. П., Витенберг Ю. Р., Пальмов В. Л.* Шероховатость поверхностей (теоретико-вероятностный подход). М., «Наука», 1975. 344 с.

91. *Шарапов И. П.* Применение математической статистики в геологии. М., «Недра», 1971. 245 с.

92. *Шеффе Г.* Дисперсионный анализ. М., Физматгиз, 1963. 628 с.

93. *Эшби У. Р.* Введение в кибернетику. М., ИЛ, Физматгиз, 1963. 628 с.

94. *Mattax C. C., Kyte J. K.* Inhibition oil recovery from fractured water-drive reservoir. — Soc. Petrol. Eng. J., 1962, vol. 6, p. 131—135.

95. *Preston F. W., Henderson I. H.* Faurier series characterization of cyclies sediments for stratigraphic correlation (Simposium of cyclies sedimentation, Kansas). — Bull. State Geol. Surv., Kansas, 1964, vol. 2, N 169, p. 415—425.

96. *Schwarzacher W.* An application of statistical timeseries analysis of a limestone-shale seguence. — J. Geol., 1964, vol. 72, N 2, p. 195—213.

ИБ № 921

Виктор Николаевич Деч
Лев Давыдович Кноринг

НЕТРАДИЦИОННЫЕ МЕТОДЫ КОМПЛЕКСНОЙ ОБРАБОТКИ И ИНТЕРПРЕТАЦИИ ГЕОЛОГО-ГЕОФИЗИЧЕОКИХ НАБЛЮДЕНИЙ В РАЗРЕЗАХ СКВАЖИН

Редактор издательства Л. Г. Ермолаева. Переплет художника Н. И. Абрамова. Технический редактор И. Г. Сидорова. Корректор В. Н. Малахова. Сдано в набор 27.01.78. Подписано к печати 20.07.78. М-18182. Формат 60 × 90/16. Бумага тип. № 1. Гарнитура обыкновенная новая. Печать высокая. Печ. л. 12. Уч.-изд. л. 13,17. Тираж 1700 экз. Заказ 786/407. Цена 85 коп.

Издательство «Недра». Ленинградское отделение. 193171, Ленинград, С-171, ул. Фарфоровская, 12. Ленинградская типография № 6 Союзполиграфпрома при Государственном комитете СССР по делам издательств, полиграфии и книжной торговли. 196006, Ленинград, М-6, Московский пр., 91.

85 коп.

• НЕДРА •