

Soft computing for qualitative and quantitative seismic object and reservoir property prediction

Part 3 Evolutionary computing and other aspects of soft computing

Программы и алгоритмы для качественной и количественной интерпретации данных сейсморазведки и оценки свойств коллекторов

Часть 3, Эволюционные вычисления и другие аспекты программ и алгоритмов

Фред Аминзаде (Fred Aminzadeh)*

Перед вами третья из серии обзорных статей по программам и алгоритмам, применяемым в нефтяной отрасли. Здесь мы рассмотрим эволюционные вычисления и смежные вопросы, как то: генетические алгоритмы (ГА), геновую инженерию, геном, ДНК, искусственную жизнь и возникновение разума. Мы начнем с краткого обзора технологии эволюционных вычислений. Будет также дан обзор применения ГА в разведочной и промысловой геофизике (РПГ). Далее отдельно рассматриваются некоторые недавние примеры реализации генетических алгоритмов в разных областях РПГ. В их число входят программы оптимизации добычи, характеристики коллекторов и оценки проницаемости. Мы также укажем области сейсморазведки, в которых использование ГА (программа Genome) и программ, основанных на «теории сложности», может стать более эффективным.

Введение

Технологии эволюционных вычислений затрагивают широкий спектр взаимосвязанных алгоритмов. Среди них следует отметить ГА, геновую инженерию, геном, ДНК, искусственную жизнь и возникновение разума. Эти технологии уже сейчас оказывают сильное воздействие на многие области. Говоря конкретно, геном человека уже находит практическое применение в биологических науках (в медицине и фармацевтической отрасли). На рис. 1, предоставленном Министерством энергетики США (проект «Геном человека»), показана связь ДНК и жизни.

Некоторые из этих методов уже используются, независимо или в сочетании с другими алгоритмами, во многих сферах наук о земле, в разведке углеводородного сырья и решении производственных задач. Большая часть ГА используется как средство эффективной оптимизации.

Их также применяли для выявления и формулировки понятий или правил, особенно в тех случаях, когда необходимо рассмотреть большой объем информации. В некоторых программах концепции, подобные геному или ДНК, использованы для классификации геологических формаций, анализа сейсмограмм или моделирования процессов осадконакопления. Это – наиболее перспективные области применения эволюционных вычислений в разведке углеводородного сырья.

Далее в этом Введении приведен обзор эволюционных вычислений, а также некоторых приложений этих методов в нефтяной отрасли. Остальная часть статьи посвящена некоторым недавним примерам из разных областей РПГ и содержит краткое описание теории сложности, которая, как ожидается, найдет широкое применение в разведочной геофизике.

Что такое эволюционные вычисления?

Эволюционные вычисления основаны на тех же принципах, которые действуют в природе («выживает самый приспособленный» - Чарльз Дарвин). Концепции ГА и «искусственной жизни» впервые представил Холланд в 1975 г. GA (Holland, 1975). Он улучшил понимание природных процессов приспособления, создав искусственные системы с такими же свойствами, как у систем природных. Ныне ГА, несмотря на их кажущуюся простоту, становятся одним из самых перспективных подходов к работе со сложными системами. Чисто аналитические алгоритмы оптимизации, будучи весьма эффективными, требуют строгих ограничений (например, дифференцируемости). ГА дают приемлемое решение задач оптимизации со многими целями и работают в конфликтных ситуациях. Они высокоэффективны, просты в применении и устойчивы при неединственном решении. Далее излагаются основы ГА (Rennard, 2002, в сокращении).

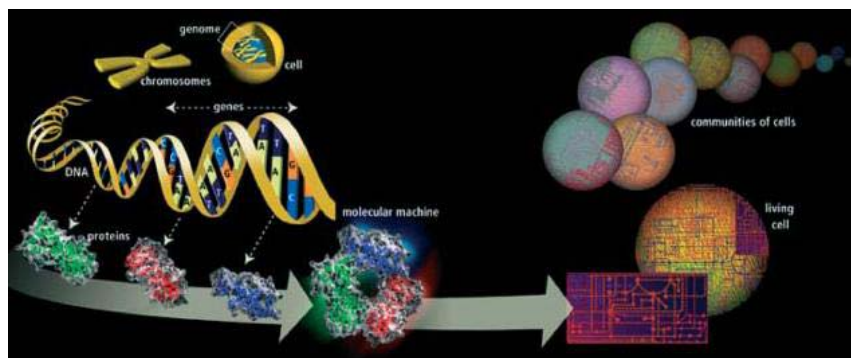


Рис. 1. От ДНК к жизни.
Министерство энергетики США,
проект «Геном»
(www.doegenome.gov)

* dGB-USA, 1 Sugar Creek Center Blvd., Suite 935, Sugar Land TX 77478, USA, fred.aminzadeh@dgb-group.com

Программы и алгоритмы

Эволюция

Отправимся на 45 млн. лет назад и исследуем «протокита» базилозавра. У него была почти независимая голова и передние лапы (рис. 2А). Задние конечности редуцировались до маленьких плавников, требовавших больших усилий для движения в воде. Чтобы плавать и охотиться эффективнее (быстрее и с большей точностью движений), должны были произойти сразу два события: 1) укорочение «плечевой» части, при фиксации локтя; и 2) удлинение пальцев, которые составят основу плавника (рис. 2А). Движения таких особей были быстрее и точнее, они жили дольше и оставляли больше потомства. Другие улучшения в аэродинамике (например, срастание головы с туловищем и усиление хвостового плавника) приспособили их к ограничениям, налагаемым водной средой. Это процесс приспособления столь совершенен, что в наши дни наблюдается паразитичное сходство акулы, дельфина и подводной лодки. Но еще раньше была хрящевая рыба Chondrichthyes, жившая в девонское время (400 млн. лет назад), задолго до появления первых млекопитающих, чьи китообразные потомки живут и ныне. Таким образом, дарвиновский механизм, лежащий в основе ГА, запускает оптимизационный процесс приспособления, включающий воспроизводство, отбор и изменение (мутации).

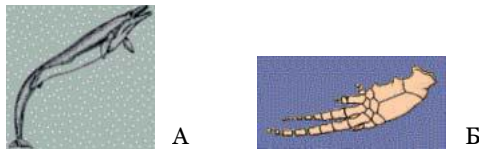


Рис. 2. Базилозавр, древний кит (А). Боковой плавник дельфина обыкновенного: два пальца гипертрофированы в ущерб остальным частям (Б).

Эволюция и ГА. Генетический пул (совокупность генов) данной популяции потенциально содержит решение (или улучшенное решение) любой конкретной задачи приспособления. Это решение не «реализовано», так как комбинация генов, на которой оно основано, разделена между несколькими особями. Только объединение различных геномов может привести к решению. Например, укорочение «лапы» и удлинение пальцев базилозавра требует наличия двух «генов». Особи с таким геномом нет, но в процессе воспроизводства и скрещивания возникает новая комбинация генов, и новая особь может унаследовать «хороший ген» от обоих родителей. В работе Эммеха (Emmehche, 1994) показано как мутация (которая сама по себе не ведет ни к какому оптимальному решению) в сочетании с холландовской рекомбинацией генов, значительно улучшает возможности алгоритма и обычно приводит к оптимуму.

Работа ГА. Каждая группа связанных признаков кодируется «хромосомой». «Гены» кодируют механизм активации или дезактивации признака. Пусть в мире есть четыре базилозавра; рассмотрим их глобальный генетический пул и «хромосомы», кодирующие длины органов передних плавников («лап» и пальцев»).

Они кодируются четырьмя генами: первые два отвечают за «лапу», остальные - за пальцы. В нашем представлении генома кружок на синем фоне обозначает активацию признака; крестик на зеленом - дезактивацию. Идеальный геном (короткие лапы и длинные пальцы) выглядит так: $\text{■} \times \text{■} \times \text{■} \times \text{■}$. Генетический пул популяции включает геномы А: $\text{■} \times \text{■} \times \text{■} \times \text{■}$, В: $\text{■} \times \text{■} \times \text{■} \times \text{■}$, С: $\text{■} \times \text{■} \times \text{■} \times \text{■}$ и D: $\text{■} \times \text{■} \times \text{■} \times \text{■}$. Заметим, что А и В ближе всего к родительским геномам, а D – к оптимуму, нужно лишь несколько удлинить пальцы. Мера соответствия или вероятность воспроизводства вычисляется путем начисления одного балла за каждый ген, соответствующий идеалу. Таким образом, совершенный геном получит четыре балла.

Рассмотрим цикл воспроизводства с четырьмя потомками. D попадетс четыре раза и даст четырех потомков. С попадетс дважды и даст двух потомков. А и В попадутс только по одному разу. Результаты скрещивания показаны в табл. 1. В ходе воспроизводства скрещивание происходит на произвольном месте (в середине генома для А', В' и С', сразу за первым геном для D'). Наличие связи между степенью приспособления и вероятностью воспроизводства приводит к росту общей приспособленности (качества) популяции. В нашем случае она увеличилась с 7 до 10.

Особь	Полученные гены	Геном	Качество	Вероятность воспроизводства
A'	A : $\text{■} \times \text{■} \times \text{■} \times \text{■}$, D : $\text{■} \times \text{■} \times \text{■} \times \text{■}$	$\text{■} \times \text{■} \times \text{■} \times \text{■}$	2	2/10=0.2
B'	B : $\text{■} \times \text{■} \times \text{■} \times \text{■}$, D : $\text{■} \times \text{■} \times \text{■} \times \text{■}$	$\text{■} \times \text{■} \times \text{■} \times \text{■}$	2	2/10=0.2
C'	D : $\text{■} \times \text{■} \times \text{■} \times \text{■}$, C : $\text{■} \times \text{■} \times \text{■} \times \text{■}$	$\text{■} \times \text{■} \times \text{■} \times \text{■}$	3	3/10=0.3
D'	C : $\text{■} \times \text{■} \times \text{■} \times \text{■}$, D : $\text{■} \times \text{■} \times \text{■} \times \text{■}$	$\text{■} \times \text{■} \times \text{■} \times \text{■}$	3	3/10=0.3
Итого			10	10/10=1

Табл. 1. Генетический пул популяции

В следующем цикле воспроизводства у С' и D' будет общий потомок, который унаследует желаемую комбинацию генов: D': $\text{■} \times \text{■} \times \text{■} \times \text{■}$ + С': $\text{■} \times \text{■} \times \text{■} \times \text{■}$.

Скрещивание – основа ГА, но имеются также и другие операции, такие как мутация. На практике искомое решение может и не содержаться в доступном генетическом пуле, даже если он велик. Мутации позволяют возникать новым сочетаниям генов, которые расширяют пул и тем самым увеличивают шанс найти оптимальное решение.

Приспособление и отбор: проблема масштаба.

Обычные методы ГА, подобные описанным выше, имеют два недостатка.

- 1) Процесс селекции «супер-особи» часто приводит к тому, что в популяции начинается конвергенция к этому оптимальному геному, что снижает разнообразие генетического пула, тормозя исполнение ГА.
- 2) В ходе исполнения ГА уменьшаются приращения меры приспособленности. В результате выравниваются вероятности воспроизводства, и выполнение ГА останавливается. Чтобы устранить эти затруднения, мера приспособленности преобразуется одним из способов: пересчет в окне, потенцирование, линейное преобразование или линейная нормализация. Ниже приводится схема типичного ГА:

1. Задача кодируется в виде бинарной строки.
2. Порождается популяция (генетический пул): случайно составленная группа возможных решений.

Программы и алгоритмы

Преимущества	Недостатки
Использует только оценки функций	Не может использовать градиенты
Легко изменятся для различных задач	Трудно учесть специфическую информацию
Хорошо работает с зашумленными функциями	Плохо определяет локальные экстремумы
Хорошо работает на больших, плохо определенных пространствах поиска	Плохо определяется условие завершения
Хорошо работает на многомодальных распределениях	Плохо работает на гладких одномодальных распределениях
Возвращает набор решений	Дополнительно требует методов поиска локальных экстремумов
Очень устойчив при затруднениях, при вычислении функции цели	Дополнительно требует методов поиска локальных экстремумов
Легко распараллеливается	

Табл. 2. Преимущества и недостатки ГА

3. Вычисляется мера приспособленности (расстояние от оптимума) каждой особи.
4. Отбор подходящих особей на основе их вклада в общую приспособленность популяции.
5. Скрещивание и мутация геномов.
6. Повтор процесса, начиная с шага 3.

Разумеется, ГА применимы не ко всем типам задач. В табл. 2 (Carter, 2003) приведены некоторые достоинства и недостатки ГА. Чтобы преодолеть некоторые из них, внедряются новые методы кодирования, основанные на биологической ДНК. Кодирование геномом или ДНК применимо для извлечения информации из большого объема данных. В ДНК имеется много избыточной информации, что важно для извлечения полезных знаний. Кроме того, эта технология допускает перекрытие представлений генов и не имеет ограничений по точкам скрещивания и один и тот же тип мутации может применяться к любой позиции.

В этом случае длина хромосомы может меняться и любой фрагмент ДНК легко вставить и/или удалить. Роль генома состоит в сохранении информации, необходимой для формирования облика особи. Делать это следует так, чтобы при применении операции скрещивания порождался жизнеспособный потомок, наследующий значительную часть генома. При конструировании генома следует также сохранять как можно больше информации о связях между генами, путем собирания связанных генов в одной хромосоме и даже в одной части хромосомы. Структура хромосомы может быть любой, если она помогает сохранить важные соотношения между генами.

Примеры применения ГА в науках о Земле, разведке и добыче нефти

Большинство приложений ГА в области нефтяных месторождений или в науках о земле ограничивается решением обратных задач или используется как способ оптимизации. Вот несколько примеров таких программ:

Эксплуатация месторождений: Довольно типичный вариант использования ГА описан в (Sen et al., 1995). ГА используются для генерирования случайных полей проницаемости по данным о дебете и потоке маркирующего вещества. Значения проницаемости в узлах сетки (параметры согласования) кодируются строкой двоичных чисел. Работа обычно ведется с популяциями численностью более 200

начальные значения присваиваются случайным образом. Буш и Картер (Bush and Carter, 1996) разработали модифицированный ГА, включавший нестандартные двоичное кодирование и стратегию отбора. Это подход был опробован на модельном разрезе с тремя параметрами: проницаемостью песков, проницаемостью глинистых сланцев и амплитудой сброса. Эти параметры кодировались двоичной строкой переменной длины. Группа Герейро (Guerreiro et al., 1998) успешно использовала ГА для оценки свойств гетерогенных коллекторов путем подбора по шести параметрам (положение прямоугольного включения, длины его сторон, значения пористости вне и внутри включения) профилей выхода маркирующего вещества в модели гетерогенной среды с соотношением составляющих 1:4. В работах Моагих и Джонсона и Роджерса (Mohaghegh, 2003; Johnson and Rogers, 2003) описано применение ГА для оптимизации добычи (см. ниже). Другие вопросы методов, основанных на ГА, и их применение рассмотрены также в работах Вонга (Wong, 2002) и Никравеша и др. (Nikraves et al, 2003).

Петрофизические и геологические программы – Хванг и др. (Huang et al, 1998) продемонстрировали применение ГА в сочетании с нейронно-нечетким анализом для интерполяции и анализа данных каротажа. Никравеш и др. (Nikraves et al, 1999) предложили применять нейронно-нечеткую генетическую модель для получения и моделирования данных в науках о Земле и исследовании нефтяных коллекторов. Кроме того, предлагается использование моделей, сочетающих нейросетевой подход, нечеткую логику и ДНК, для извлечения информации из сейсмических данных, для совмещения данных каротажа и сейсморазведки, и для оценки пористости (и проницаемости) с применением сейсмического картирования. Сафтик и Велич (Saftic and Velic, 2000) изучили тонкую стратиграфическую структуру миоценовых осадков. Поттер и Корбетт (Potter, and Corbett, 2000) использовали генетическую интеграцию петрофизических и геофизических данных для улучшения оценок ключевых параметров.

Программы для сейсморазведки. Стоффа и Сен (Stoffa and Sen, 1991) применяли ГА для решения 2D обратной задачи сейсморазведки. Они же (Sen and Stoffa, 1992) привели примеры инверсии сейсмического сигнала с основанным на ГА быстрым просмотром пространства моделей. Сторк и Кусума (Stork and Kusuma, 1992) использовали генетический метод для ввода статических поправок в зашумленные данные. Маллик и Фрейзер (Mallick and Frazer, 1995) применили ГА к решению обратной задачи сейсморазведки со стартовой моделью по данным до суммирования. Рот и Холлинджер (Roth and Hollinger, 1998) использовали ГА для оценки скоростей P- и S-волн по характеристикам

Программы и алгоритмы

затухания волн Рэлея и волн в волноводах. Метод основан на предположении, что у поверхности земли коэффициент Пуассона настолько велик, что позволяет считать волну в волноводе акустическим сигналом. Шнайдер (Snieder, 1997) применил ГА для решения обратных задач. Барнс (Barnes, 1997) применил генетическую классификацию для анализа атрибутов комплексной сейсмической трассы. Ма (Ma, 1998) использовал алгоритм «закаливания» для расчета акустического импеданса по сейсмическим данным после суммирования и прослеживания границ.

МакКормак и др. (McCormack et al., 1999) дают обзор применения ГА в разведке и добыче. Маллик (Mallick, 1999, 2001) предложил уточненный, основанный на ГА, метод решения обратной задачи по данным до суммирования. ГА, в сочетании с Байесовским подходом, априорной информацией о модели и физическими основами решения прямой задачи, применялись для расчета синтетических трасс. Синтетические трассы сравнивались с наблюдаемыми, чтобы получить приближенные оценки граничных значений апостериорных плотности вероятности (АПВ) в пространстве моделей. Рассматривая эти АПВ, интерпретатор может выбирать модели, которые лучше всего описывают конкретные геологические условия и позволяют точно прогнозировать сейсмический разрез. Джи и др. (Ji et al, 2000) дали количественные оценки устойчивости обратной задачи при изменении амплитуды в зависимости от скорости (медленности). Порсани и Урсин (Porsani and Ursin, 2000) применили ГА при деколонволюции и оценке формы сигнала. Дункан и Латкивич (Duncan and Latkiewicz, 2002) рассмотрели использование кодирования геномом для оценки внутренних параметров разреза путем «декодирования» данных сейсморазведки (см. ниже). Кин и др. (Qin et al, 2003) описывают в терминах ГА метод анализа скоростей, основанный на картировании ОГТ.

Свежие примеры использования ГА в РПГ

Далее рассмотрим несколько свежих примеров применения ГА для оптимизации добычи, описания коллекторов и оценки проницаемости.



Рис. 3. Эксплуатационные (x), нагнетательные (пронумерованы) и планируемые нагнетательные (#/x) скважины. Нумерация, как и положение в строке битов, произвольна.

Оптимизация добычи

В этом разделе методы ГА применяются для оптимизации закачки воды путем поиска наилучшей комбинации скважин. Изложение ведется в соответствии с работой Джонсона и Робертса (Johnson and Rogers, 2003), в которой ГА применен как поисковая машина в нейронной сети, которую предполагалось использовать для оптимизации проекта закачки воды на опытном месторождении. Цель состояла в том, чтобы указать такое положение куста нагнетательных скважин (от одной до четырех), при котором месторождение которое давало бы максимальную чистую прибыль в течение семи лет. С использованием стандартной программы моделирования коллекторов, была создана база данных из 550 результатов моделирования при различных комбинациях из 25 возможных местоположений нагнетательных скважин. Первые запросы к базе данных относились к оптимальным сценариям максимизации чистой прибыли за три и семь лет. Применение ГА при решении этой задачи рассматривается ниже.

Применение ГА. 25 положений скважин, образующих множество переменных задачи, представляются в ГА в виде строки из 25 битов, каждый из которых может принимать значения 0 или 1. В этой постановке пространственное положение каждой скважины задано явным образом. Порядок скважин в строке задан их номерами на рис. 3. Поиск начинается с набора из 100 комбинаций скважин; для каждой комбинации вычисляется функция цели. Из старого поколения комбинаций с использованием трех механизмов – отбора, воспроизводства и мутации - формируется новое, также из 100 комбинаций.

Отбор. Этот механизм определяет, какие члены текущего поколения будут отобраны для передачи, в той или иной форме, признаков следующему поколению. Чтобы исключить потерю самых подходящих комбинаций из-за случайностей отбора и скрещивания, три комбинации с наибольшими рангами переносятся в следующее поколение без изменений. Остальные 97 позиций заполняются в ходе процесса с отбором родителей. В ходе отбора за один раз отбирается пара родительских комбинаций, из которых формируется комбинация-потомок. Этот процесс повторяется, пока не будут порождены все потомки. Одна и та же комбинация может быть у обоих членов родительской пары; в таком случае потомок оказывается просто клоном родителя.

Воспроизводство (Скрещивание). Чаще всего воспроизводство происходит в форме одноточечного скрещивания. Комбинации-потомки формируются путем разрыва родительских комбинаций в случайно выбранной точке строки битов и соединении сегментов от каждого родителя. При создании новых комбинации из «обрывков» старых, важной становится близость положений в строке. Это значит, что близость скважин в строке должна отражать одно или более соотношений, существующих в физической задаче, представленной данной строкой. Чтобы разрушить недопустимые сочетания, используется другой механизм воспроизводства – однородное скрещивание. В этом случае значение каждого бита в строке-потомке определяется независимо от всех остальных битов. Вероятность смены значения

Программы и алгоритмы

можно установить так, чтобы дать преимущество более подходящему родителю (если такой есть); но в рассматриваемом эксперименте вероятность смены значения всегда принималась равной 0.5.

Мутация. Мутация – способ поддержать разнообразие генов популяции, путем произвольного изменения значений битов в комбинациях-потомках в некотором темпе (часто – обратно пропорциональном численности популяции). Высокие темпы мутации могут подавить эффект от скрещивания, низкие – ограничивают проникновение «новизны» в популяцию.

Обычно оптимизация завершения работы основана на неком признаке схождения к единственному лучшему решению. В данном случае цель состоит скорее в порождении набора решений, близких к оптимальному, чем в поиске единственного лучшего решения. Эта цель достигается увязкой критерия завершения со средним качеством популяции, а не с качеством наилучшей отдельной комбинации. Поиск прекращается либо когда среднее качество популяции не улучшается на протяжении пяти поколений, либо когда сменилось некоторое максимальное количество (25) поколений. Максимальные оценки каждого поколения сохраняются. Комбинации с наивысшими оценками из каждого поколения образуют множество решений, близких к оптимальному. На результат поиска влияет случайный выбор, производимый в ходе решения. Для повышения устойчивости окончательный результат формируется по результатам 10 поисков, каждый из которых проводится с новой инициализацией датчика случайных чисел.

Далее составлялся наилучший сценарий, полученный каждым методом. Он состоял из набора сценариев, близких к оптимальным, которые можно было анализировать, пытаясь понять, какие физические причины делают их оптимальными для решения конкретной задачи управления. Например, рассмотрев 25 лучших комбинаций из поиска ANN-GA, можно отметить, Скв. 7 фигурирует во всех (100%) комбинациях, а далее с большим отрывом следуют Скв. 24 (32%), Скв. 11 (28%) и Скв. 16 (24%). Можно сделать вывод, что Скв. 7 оказывает более сильное воздействие на соседние эксплуатационные скважины, которые важны для добычи и по истечении семилетнего срока. Частота других скважин может выражать известную рекомендацию повышать давление по краям залежи.

Другой важный аспект заключается в неопределенности, вносимой самой программой моделирования коллекторов. С учетом того, что имеется много альтернатив, «близких к оптимальным», важно знать некие показатели того, насколько сильны изменения, связанные с выбором одной из альтернатив. Хотя в общем случае в этом и состоит основная проблема, можно предложить некоторые подходы, основанные на линеаризации, для включения в алгоритм ГА элементов анализа неопределенности на различных этапах процесса оптимизации. Очень простым подходом будет дать каждой скважине ранг в соответствии с относительной точностью определения физических свойств вблизи нее. Оптимизируемая функция цели должна содержать условия санкции (исключения), основанное на величине этого ранга, который отражает риск, связанный с включением данной скважины в сценарий.

Моделирование и характеристика коллекторов

В этом разделе изложение ведется в соответствии с работой Ромеро и Картера (Romero and Carter, 2003), в которой предлагается применять ГА для характеристики коллекторов. В типичной задаче моделирования/характеристики коллектора проблема состоит в задании начальных значений различных свойств коллектора в каждом блоке сетки (обычно около 20000 блоков). Размер модели часто определяется тремя факторами: требуемой геологической сложностью коллектора, мощностью вычислительных средств и временем, отведенным на выполнение исследования. Среди свойств коллектора, подлежащих определению:

- Пространственное положение
- Размер
- Пористость – доля общего объема, занятая флюидами
- Водо-/нефте-/газо-насыщенность – доля порового пространства, занятая каждым типом флюида
- Давление
- Проницаемость – легкость перемещения флюидов по горизонтали и по вертикали
- Глинистость – доля породы, не обладающей коллекторскими свойствами

Их значения должны отвечать некоторым физическим ограничениям, например: блоки сетки закрывают все пространство без перекрытия; водонефтяная смесь располагается выше участков, где есть только вода; давления должны соответствовать законам гидростатики. В настоящее время под характеристикой резервуара понимается процесс, в ходе которого для каждого блока сетки определяется четыре величины: горизонтальная и вертикальная проницаемость, пористость и глинистость. Другие величины (свойства разрывных нарушений, остаточная водонасыщенность, минимальная нефтенасыщенность, параметры зоны проникновения) также могут быть определены в ходе процесса характеристики свойств коллектора.

Определить по очень малому объему наблюдаемых данных значения примерно в 100000 точек, которые определяют начальные условия – весьма сложная задача. В данном исследовании размеры залежи, на которой было 11 эксплуатационных и 6 нагнетательных скважин, составляли 11 км x 3 км x 200 м. Таким образом, плотность измерений по площади составляла 1 измерение на 2 км².

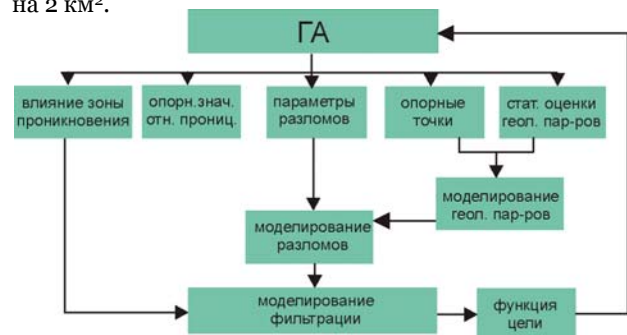


Рис. 4. Схема процедуры определения характеристик коллектора, используя «ключевые значения»

Программы и алгоритмы

Каждое измерение относится к блоку с площадью основания около 1 м^2 и высотой 20 м, таким образом мы имеем всего 510 замеров. Имеются также данные ежемесячных измерений скоростей фильтрации флюидов за четыре года (не более 1344 замеров в скважинах). Помочь в решении проблем, возникающих в связи с редкой сетью наблюдений, может учет плохо формализуемой информации, известной по геологическим и геофизическим данным. Распределение свойств в пространстве должно быть согласовано со статистическими параметрами геологических процессов, сформировавших залежь

Многие распределения свойств залежи (например, проницаемости и пористости), которые согласуются с результатами измерений скоростей фильтрации, либо не согласуются с другими данными, либо геологически бессмысленны. Таким образом, следует определить, какие распределения начальных условий согласуются со статистическими оценками геологических свойств и дают результат численного моделирования, удовлетворительно совпадающий с данными измерений. От попыток найти модель, дающую совпадение с результатами измерений, мы перешли к поиску моделей, которые с высокой вероятностью согласуются с имеющейся информацией. Во многих случаях мы ограничиваемся попытками построить наиболее вероятную модель.

В процессе характеристики залежи (рис. 4) определяются значения следующих основных величин:

- Величины, описывающие три свойства залежи (пористость, проницаемость и глинистость) по всему коллектору
- Некоторые величины, описывающие течение флюидов через разрывные нарушения
- Величины, описывающие зону проникновения каждой скважины

- Опорные значения относительной проницаемости.

Как показано на рис. 4, свойства залежи используются двояко, напрямую при моделировании фильтрации, а также при расчете свойств разрывных нарушений. Разрывные нарушения часто встречаются на месторождениях углеводородов, поэтому математические модели должны их включать. Способность разрыва нарушать фильтрацию – один из важнейших факторов, влияющих на общую погрешность модели. Разрывы влияют на модель фильтрации в залежи двояко. Они изменяют условия

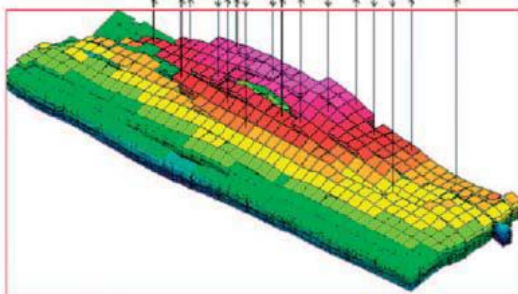


Рис. 5. Структура залежи с положениями скважин (Romero and Carter, 2003)

сопряжения элементарных потоков в осадках, поскольку смещения по плоскости разрыва приводят к полному или частичному наложению элементарных потоков. Другой механизм влияния на поток – изменение петрофизических свойств в зоне разрывного нарушения. Как мощность, так и проницаемость разломной зоны – наблюдаемые величины. Влияние зоны проникновения обычно состоит в понижении давления в зазоре скважины.

Описание залежи и план добычи

В данной работе рассматривается модельная залежь, чтобы избежать проблем с недостатком данных и неопределенностью строения. Строение модельной залежи довольно типично для тектонических ловушек Северного моря. Она ограничена крупным поперечным разломом. Как показано на рис. 5, внутри залежи плоскости разрывных нарушений падают как по направлению главного разлома, так и в противоположном. Чтобы получить правдоподобную модель, в ходе определения характеристики залежи и оценки соответствующих основных параметров, применен метод ГА. План добычи был реалистичным и соответствовал историческим данным по добыче. Среди упрощающих задачу допущений отметим следующие: вертикальные скважины, одинаковая для всех скважин стадия бурения, неизменные условия добычи. Графики добычи рассчитаны по реальным моделям и зашумлены нормально распределенной помехой.

Построение ГА. На рис. 4 показан ГА, управляющий процессом ввода на этапе описания залежи. Причины выбора ГА вместо обычных методов типа спуска были, в частности, таковы:

- Обширное пространство поиска
- Сложности с предварительной оценкой чувствительности задачи к изменению тех или иных величин
- Неединственность геометрических параметров при данном описании залежи
- Проблемы, связанные с нелинейностью, многомодальностью и плохой определенностью.
- Недостаток информации о поведении частных производных по всем переменным

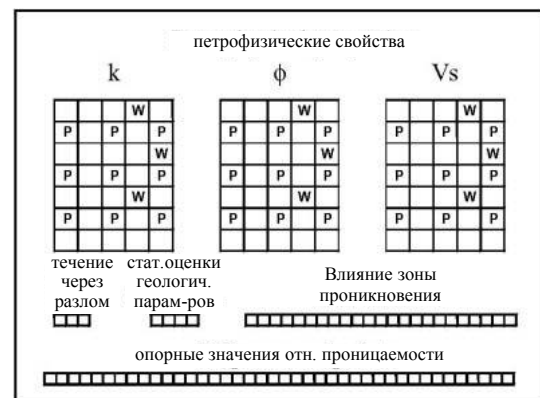


Рис. 6. Нестандартная структура генома

На рис. 6 показаны структуры геномов для различных параметров. Первые три хромосомы (отвечающие за свойства залежи) имеют одинаковую трехмерную структуру. Параметры разрывных нарушений кодируются тремя вещественными числами с известными пределами изменения. Статистические оценки геологических параметров кодируются четырьмя вещественными числами, также с известными пределами изменения. Влияние зоны проникновения и опорные значения относительной проницаемости кодируются соответственно 17-ю и 36-ю вещественными числами. Используются два оператора скрещивания. Для одномерных хромосом выбрано обычное скрещивание в k-той точке, причем вероятность взятия информации от данного родителя постоянна для каждой хромосомы. Для трехмерных хромосом разработан специальный оператор скрещивания, обеспечивающий переключение битов. Процесс оптимизации в ГА обычно состоит из стадии очень резкого убывания функции цели, который сменяется менее резким убыванием. Ошибка, связанная с дебетом воды, дает наибольший вклад в функцию. Сравнение с данными по скважинам (рис. 5) показало, что результаты хорошо согласуются с историей. Основные преимущества метода таковы:

- Алгоритмы дают семейство решений, из которых выбирается лучшее
- Можно получить несколько решений без повторного сопоставления с историей
- Простота реализации, устойчивость (малая вероятность потери или порчи решения)
- Легко распараллеливается, поскольку эффективен по вычислительным затратам
- Необходимо умеренное число решений прямой задачи для получения хороших решений
- Невысокая чувствительность к начальным значениям параметров, что подходит для характеристики

Признаки классов в ГП при моделировании проницаемости коллекторов. Во второй части этой серии (Aminzadeh and Wilkinson, 2004) было показано как можно использовать нечеткую логику при оценке проницаемости, самой сложной задачи моделирования коллекторов. Задачу моделирования с использованием генетического программирования (ГП) решил Коца (Koza, 1992). Уилкинсон и др. (Wilkinson et al, 2003) так описывают моделирование с выделением классов на базе ГА:

Для каждой литологической разности ГП используется для выделения признаков классов, указывающих пределы изменения проницаемости. Признаки класса - это несколько критических значений для различных диапазонов проницаемости. Например, для песков два критических значения, разграничивающих пески с высокой, средней и низкой проницаемостью, будут 3.0 и 2.0. Для супесей, суглинков и глин пары таких критических значений будут соответственной (2.5, 1), (2.5, 1) и (1, 0). ГП пригодно лишь для выделения признаков деления на два класса. Для литологических разностей, где возможно деление на три класса, нужно два признака. Параметры ГП, используемые при выделении этих признаков, приведены в табл. 3. Оценка качества основана на рейтинге – доле обучающих данных, для которой классификация проведена верно.

В процессе «турнирного» отбора случайным образом выбираются четыре кандидата, из которых отбирается два победителя. При совпадении рейтингов двух кандидатов победитель определяется по среднеквадратической погрешности. В рамках этого исследования среднеквадратическая ошибка определяется с точностью 0.01%. Если два признака совпадают и по рейтингу и по среднеквадратической ошибке, победитель определяется случайным образом. Лучшие признаки объединяются в команды. Команда принимает решение большинством голосов. Поскольку число признаков нечетно, победитель есть всегда (нет совпадений). Пример команды приведен ниже:

$$f_1(\phi, v, \rho) = K \rho * v - \phi^2 / \rho K$$

$$f_2(\phi, v, \rho) = K \rho^v - \phi^2 / \rho + v + 1.2K$$

$$f_3(\phi, v, \rho) = K v * v - \phi^2 / \rho - (\phi + 2.3)K$$

Критическое значение (K3) = 2.5;
 Большинство: ($f_1 > K3$; $f_2 > K3$; $f_3 > K3$)

Цель	Создание поискового признака, который различает высокую, среднюю и низкую проницаемость
Функции	Сложение; вычитание; умножение; деление; абсолютная величина; преобразование данных
Данные	Пористость (ϕ), плотность (ρ), скорость ($1/v$) и константы
Оценка качества	Рейтинг, затем среднеквадратическая погрешность
Рейтинг	Доля правильно классифицированных данных в обучающей последовательности
Отбор	Турнир (4 кандидата/2 победителя)
Размер популяции	100,000
Максимальное число поколений	9,000,000
Максимальная длина	256 узлов
Генетические операторы	скрещивание с вер. 50% (95% с тем же строением), мутация с вер 95% (по блокам - 30%, с заданной скоростью - 65%)

Табл. 3. Параметры ГП

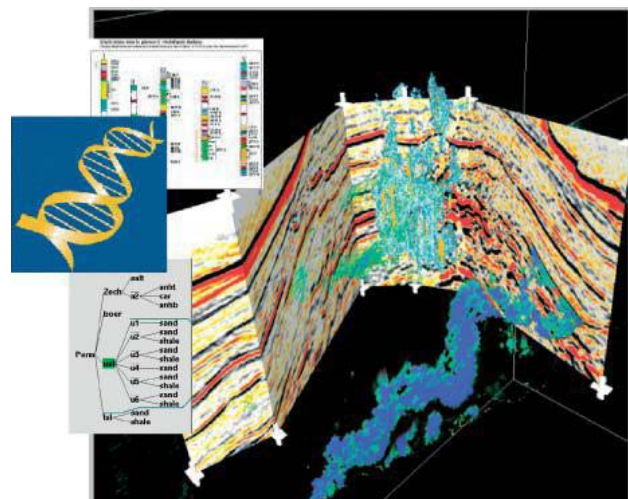


Рис. 7. Пример связи сейсмической волновой картины и ДНК (сейсмическая запись любезно предоставлена www.opendtect.org)

Программы и алгоритмы

Применение эволюционного

программирования и сходных технологий в будущем

В этом разделе рассматриваются две развивающиеся в настоящее время технологии – методы, основанные на геноме (ДНК) человека и «теория сложности» - которые могут в будущем оказать сильное воздействие на нефтяную отрасль.

Применение генома для анализа обнажений, геохимических и сейсмических данных

Голубая мечта исследователя – иметь возможность по малому объему данных прямых измерений (на обнажениях или шлифах) сделать конкретные выводы об истории развития и перспективности погребенных структур. Еще хотелось бы уметь получать такую же информацию косвенными методами - по сейсмическим данным, по скважинам или наземной геохимической съемке. Подходы, развитые при изучении генома человека и структуры ДНК, делают эти мечты более реальными.

Идея состоит в том, чтобы связать особенности сейсмических данных, результатов бурения или съемки биомаркеров в неких поисковых признаках («отпечатках пальцев») различных геологических режимов или процессов осадконакопления, приводящих к возникновению нефтегазосодержащих пород и залежей, пригодных к разработке. Вероятность успеха усилий на этом направлении повысится, если научиться создавать модели, из некоторых, представленных легко идентифицируемыми «генами» «кирпичей», которые можно нужным образом переставлять. «Кирпичи» позволят объяснить роль каждого «гена» в группе. Как в случае с осадконакоплением, процесс продолжается, если сочетание признаков («действующее лицо») изменяется, успешно противодействуя воздействию среды.

Вот как обстоят дела с «возникновением» или приспособлением: «кирпичи» одного уровня, объединяясь, образуют «кирпичи» более высокого уровня. Как показано в следующем разделе, «единая теорема», которую обещает теория сложности, дает шанс создать такие

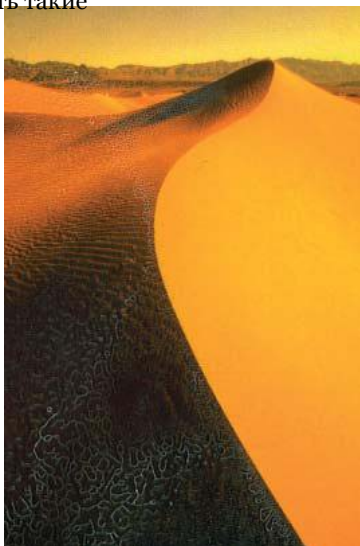


Рис. 8. Грань хаоса (Waldrop, 1992)

ДНК по разнотипным данным, несмотря на проблемы, связанные с масштабом, уровнем шума или достоверностью. Например, сопоставление наличия в сейсмических данных тех или иных деталей волновой картины тщательно построенной иерархической модели геологического строения можно осуществить, привязывая их к набору ДНК или геномов (рис. 7).

В науках о Земле и в геологоразведке имеются некоторые составляющие для этого типа анализа. Первые попытки уже сделаны (Duncan and Lat, 2002), но мы все еще далеки от полноценного использования этой технологии.

Теория сложности

Применение теории сложности - еще одна область, потенциально способная оказать в ближайшие годы значительное воздействие на методы исследования в целом. Теория сложности изучает взаимодействие независимых агентов, каждый из которых находится под воздействием остальных, и все вместе они воздействуют на целое. Организмы постоянно приспосабливаются друг к другу в процессе эволюции, объединяясь тем самым в тонко настроенные экосистемы. Самоорганизующиеся системы адаптивны. Они активно стараются обратить любые события к своей пользе. Им присущ своего рода динамизм, качественно отличающий их от статических объектов. В сложных системах больше спонтанности, больше беспорядка, чем в хаотических, они более живые, чем хаотические.

Теория хаоса потрясла самые основы науки, показав, что самые простые законы динамики могут приводить к очень сложным явлениям – рекам и фракталам. Сложные системы как-то находят способ научиться приводить порядок и хаос к равновесию. Лучше всего о теории сложности можно сказать так: это место встречи порядка и хаоса. На грани хаоса возникают новые идеи, а обновляющиеся генотипы постоянно подгрызают по краям существующее положение. Грань хаоса – это постоянно смещающееся поля боя между застоєм и анархией, единственное место, где сложная система может быть спонтанной, адаптивной и живой.

В замечательной книге Уолдроп (Waldrop, 1992) по теории сложности говорится: «...он чувствовал, что старые категории науки начинают растворяться. Вот-вот должна была родиться какая-то новая, единая наука. Это должна быть удивительная наука. Он был убежден, что она будет «строгой», как всегда была физика, и столь же основательно базироваться на законах природы. Но она не будет погоней за частицами, она будет заниматься потоками, изменениями и образованием и разрушением узоров. Вместо игнорирования всего негладкого и непредсказуемого она оставит в истории место для единичного и случайного. Она не будет наукой о простом, она будет наукой, ну-у..., о сложном.»

В теории сложности существует несколько родственных концепций. Например, по словам Уолдроп, иллюзорная «единая теория», на которой были сосредоточены исследования Института Санта Фе, стала давать некото-

рые результаты, установив некоторую связь между, казалось бы, не связанными областями и концепциями:

1. Классы конечных автоматов: I&II => «IV» => III
2. Динамические системы: Порядок => «Сложность» => Хаос
3. Вещество: Твердое => «Фазовый переход» => Жидкость
4. Вычисления: Останов => «Неразрешимая ситуация» => Продолжение

Как предлагал Аминзаде (Aminzadeh, 2000) теории хаоса и сложности следует рассмотреть в качестве возможных будущих подходов к созданию нового поколения средств обработки и анализа данных сейсморазведки. Они предлагают устойчивое и стабильное поведение в очень сложных и многомерных задачах.

Вопрос в том, сможем ли мы использовать эти идеи для лучшего понимания задач, с которыми мы сталкиваемся в разведке и добыче нефти

- Насколько больше информации мы получим по отражениям от разрывных нарушений?
- Будут ли физически значимыми границы эвапоритов и прочие границы со своеобразной волновой картиной (грань хаоса) и как установить их морфологию?
- Есть ли теория, способная описать «сверхпроницаемость»?
- Как лучше использовать поглощение высоких частот, обычное на крупных скоплениях углеводородов
- Связано ли образование нефтематеринских пород и отдача углеводородов с «гранью хаоса»?

Есть надежда, что будущие исследования дадут какие-то конкретные ответы на эти вопросы и позволят сделать скачок в применении нашей технологии.

Заключение

Мы дали краткий обзор эволюционных вычислений и ГА. Рассмотрено применение ГА в нефтяной отрасли. Мы пришли к выводу, что такие приложения пока лишь «скребутся» по поверхности области потенциального вклада эволюционных вычислений в исследования углеводородов. Также рассмотрены некоторые будущие возможности применения ДНК и теории сложности, показаны потенциальные области их применения в разведке и доразведке месторождений и в организации добычи.

Литература

- Aminzadeh, F. [2000] Challenges - Direct Future of Geophysics, *American Oil and Gas Reporter*, January, 123-132.
- Barnes, A. E., [1997] Genetic classification of complex seismic trace attributes. *67th Annual Meeting, EAGE*. Geophys., 1151-1154.
- Bush M.D. and Carter, J.N. [1996] Application of a Modified Genetic Algorithm to Parameter Estimation in the Petroleum Industry, In Dagli et al. (eds.), *Intelligent Engineering Systems through Artificial Neural Networks*

6, ASME, Press, York, 397-402.

- Carter, J. N. [2003] Genetic computing, in *Soft Computing and Intelligent Data Analysis* Eds., Nikravesh, M., Aminzadeh, F., Zadeh, L. A., Elsevier.
- Curtis, A. and Snieder, R. [1997] Reconditioning inverse problems using the genetic algorithm and revised parameterization. *Geophysics*, **62**, 1524-1532.
- Dorrington, K. P. et al. [2004] Genetic-algorithm/neural-network approach to seismic attribute selection for well-log prediction. *Geophysics* **69**, 212.
- Duncan, P. M. and Latkiewicz, C. [2002] Technology Unravels 'Genome Code' of 3D Data to Improve Quality, Speed of Seismic Exploration. *World Oil*, 99-107
- Emmeche C. [1994] *Garden in the Machine. The Emerging Science of Artificial Life*. Princeton University Press, 114
- Goldberg, D. E. [1989] *Genetic algorithms in search optimization and machine learning*. Addison-Wesley Publ. Co.
- Guerreiro, J.N.C., Barbosa, H.J.C., Garcia, E.L.M., Loula A.F.D, and Malta, S.M.C. [1998] Identification of Reservoir Heterogeneities Using Tracer Breakthrough Profiles and Genetic Algorithm. *SPE Reservoir Evaluation & Engineering*, 218-223.
- Holland J.H. [1975] *Adaptation in natural and artificial system*. Ann Arbor, University of Michigan Press.
- Huang, Y., Wong, P. M. and Gedeon, T. D. [1998] Neural-Fuzzy-Genetic-Algorithm Interpolator in Log Analysis. *60th Annual Meeting, EAGE*, Session: P106.
- Ji, Y., Singh, S. and Hornby, B. [2000] Sensitivity study using a genetic algorithm: inversion of amplitude variations with slowness. *Geophysical Prospecting* **48**, 1053-1074.
- Johnson, V. M. and Rogers, L. L. [2003] Applying soft computing methods to improve the computational tractability of subsurface simulation-optimization problem, in *Soft Computing and Intelligent Data Analysis*, Eds., Nikravesh, M., Aminzadeh, F., Zadeh, L. A. Elsevier.
- Koza, J. [1992] *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press, Cambridge, MA.
- Ma X. Q. [1998] An Accurate Seismic Inversion Approach Using Exact and Over-Parameterization methods. *SEG Extended Abstracts*.
- Mallick, S. [1995] Model-based inversion of amplitude-variation-with-offset data using a genetic algorithm. *Geophysics*, **60**, 939-954.
- Mallick, S. [1999] Some Practical Aspects of Prestack Waveform Inversion using a Genetic Algorithm: An example from the East Texas Woodbine Gas Sand. *Geophysics*, **64**, 2, 326-336.
- Mallick, S. [2001] Prestack waveform inversion using a genetic algorithm - The present and the future. *Recorder*, **26**, 6, 78-84
- McCormack, M. D., Stoitsits, R. F., Macallister, D. J. and

Алгоритмы и программы

- Crawford, K. D. [1999] Applications of genetic algorithms in exploration and production. *The Leading Edge* **18**, 6, 716-718.
- Mohaghegh, S. D. [2003] Virtual Magnetic resonance logs, a low cost Reservoir Description tool, in *Soft Computing and Intelligent Data Analysis*, Eds., Nikravesh, M., Aminzadeh, F., Zadeh, L. A., Elsevier.
- Nikravesh, M., Aminzadeh, F., Zadeh, L. A. [2003] *Soft Computing and Intelligent Data Analysis*. Elsevier.
- Porsani, M. and Ursin, B. [2000] Deconvolution and wavelet estimation by using a genetic algorithm. *70th Annual Meeting, SEG*, 2185-2188.
- Potter, D. and Corbett, P. [2000] Genetic petrophysics and data integration in pegasus - improved prediction of key parameters. *62nd Annual Meeting, EAGE*, Session:X0012.
- Qin, Y., Zhang, Z. and Xu, S. [2003] CDP mapping in tilted transversely isotropic (TTI) media. Part II: Velocity analysis by combining CDP mapping with a genetic algorithm. *Geophysical Prospecting*, 51, pp. 325-332.
- Rennard, J. P. [2000] www.rennard.org/alif
- Romero C.E. and Carter J.N. [2003] Using genetic algorithms Genetic Algorithm for Reservoir Characterization, in *Soft Computing and Intelligent Data Analysis*, Eds., Nikravesh, M., Aminzadeh, F., Zadeh, L. A. Elsevier.
- Roth, M. and Hollinger, K. [1998] Joint Inversion of Rayleigh and Guided Waves in High Resolution Seismic Data Using Genetic Algorithms. *SEG Extended Abstracts*.
- Saftic, B. and Velic, J. [2000] Genetic stratigraphic sequences in the upper Miocene sediments of the Sava basin. *62nd Annual Meeting, EAGE*, Session: P0033.
- Sen, M. K., and Stoffa, P. L. [1992] Rapid sampling of model space using genetic algorithms: Examples from seismic waveform inversion. *Geophys. J. International*, **108**, 281-292
- Sen, M.K., Datta-Gupta, A. K., Stoffa, P.L., Lake L.W., and Pope, G.A. [1995] Stochastic Reservoir Modeling Using Simulated Annealing and Genetic Algorithms. *SPE Formation Evaluation*, 49-55.
- Stork, C. and Kusuma, T. [1992] Hybrid genetic autostatics: A new approach for large amplitude statics with noisy data. *62nd Annual Meeting, SEG*, 1127-1131.
- Waldrop, M. M. [1992] *Complexity, the Emerging Science at the Edge of Order and Chaos*. Simon and Schuster, NY.
- Wilkinson, D., Yu, T., and Xie, D. [2003] *Genetic Programming Theory and Practice*. Kluwer Academic Publishers, 271 - 289.
- Wong, P. M., Aminzadeh, F. and Nikravesh, M. [2002] Soft Computing for Reservoir Characterization and Modeling. *Series Studies in Fuzziness and Soft Computing*, **80**, Physical Verlag, Springer, 2002.