

2. Чупров И.Ф. Теоретические и технологические основы теплового воздействия на залежи аномально вязких нефтей и битумов: дис. ... д-ра техн. наук. – Ухта, 2009. – С. 67.

3. Aziz Kh., Gontijo J.E. A simple analytical model for simulating heavy oil recovery by cyclic steam in pressure-depleted reservoirs // Paper SPE 13037 presented at the 59th Annual Technical Conference and Exhibition. – Houston, September 16–19, 1984. 13037-MS.

4. Малофеев Г.Е. О механизме притока нефти при паротепловой обработке скважин // Нефт. хоз-во. – 1986. – № 6. – С. 38–40.

5. Урсегов С.О. Обоснование оптимальных параметров термических технологий разработки крупных месторождений высоковязких нефтей Тимано-Печорской нефтегазоносной провинции: дис. ... канд. техн. наук. – Ухта, 2007. – С. 73.

УДК 622.276

## ДЕРЕВЬЯ РЕШЕНИЙ КАК ЭФФЕКТИВНЫЙ МЕТОД АНАЛИЗА И ПРОГНОЗИРОВАНИЯ

**О.Н. Пичугин, Ю.З. Прокофьева**

(ЗАО "КОНКОРД", г. Москва),

**Д.М. Александров**

("Бузачи Оперейтинг Лтд.", г. Актау)

В настоящее время в отечественной нефтедобывающей отрасли сложилась устойчивая тенденция снижения эффективности геолого-технических мероприятий (ГТМ), которое связывают, как правило, с ухудшением структуры остаточных запасов. Этот объективный фактор, без сомнения, существенно снижает результативность ГТМ. Однако нельзя не упомянуть и о методических аспектах проектирования методов воздействия на пласт. Так, в одной известной книге [1] есть очень верный тезис: "...Как всякие большие системы, объекты нефтегазодобычи требуют использования целой иерархии моделей – от дифференциальных до интегральных, от детерминированных до адаптивных..." . К сожалению, сегодня для решения сложных задач прогнозирования ГТМ в основном используются только детерминированные модели. Действующие регламентные документы не дают четких рекомендаций и схем проектирования ГТМ. Более того, "современная мода" на использование 3D симуляторов сводит к минимуму применение простых подходов, интегральных и адаптивных моделей. Таким образом, становится понятным наличие другой, уже субъективной причины, обуславливающей низкую эффективность геолого-технических мероприятий, – это проблемы методического характера.

За последние десятилетия накоплен большой опыт реализации ГТМ, который содержит в себе полезные знания и дает возможность применения формализованных подходов для получения правил и рекомендаций. Современное направление анализа данных за рубежом получило емкое название "datamining" [2] – добыча знаний из "сырых" данных. Практический опыт применения различных инструментов анализа позволил остановиться на деревьях решающих правил [3, 4] – простом, прозрачном, но в то же время эффективном методе машин-

ного обучения. В соответствии с этим алгоритмом иерархический процесс построения дерева начинается с поиска наиболее селективного правила для корневой вершины и далее продолжается для дочерних ветвей, пока не будут выполнены определенные терминальные условия.

Рассматриваемый подход обладает следующими важными достоинствами по сравнению с другими нелинейными методами прогнозирования:

- высокая скорость обучения;
- извлечение правил на естественном языке;
- интуитивно понятная классификационная модель;
- возможность выявления причин успешности/неуспешности мероприятий;
- возможность экспертной оценки полученных деревьев решений.

Для демонстрации возможностей использования данного инструмента в работе рассмотрен опыт бурения и ввода новых добывающих скважин. Фонд скважин разделен на 3 различных типа: основной, уплотняющий и горизонтальный одного объекта. Выборка основного фонда состоит из 356 скважин, уплотняющего – 356 скважин, горизонтального – 114 скважин. В работе использовались классификационные алгоритмы, в связи с этим каждая выборка скважин была разбита на 2 группы – "успешную" и "неуспешную". Поскольку средние значения входного дебита нефти в зависимости от фонда скважин сильно различались, по каждому из них были приняты свои критерии успешности: по основному фонду – 11 т/сут, по уплотняющему – 8 т/сут, по горизонтальному – 13 т/сут. Критерии подбирались таким образом, чтобы выборка делилась примерно на две равные части. В таблице приведены средние значения по фондам скважин и группам успешности.

**Среднее значение входного дебита нефти (т/сут) по группам успешных и неуспешных для каждого фонда**

Фонды	Успешные	Неуспешные
Основной	31,4	5,3
Уплотняющий	15,7	4,5
ГС	21,8	8,4

**Основной фонд**

Рассмотрим пример дерева решений, полученного в результате обучения на выборке скважин основного фонда (рис. 1).

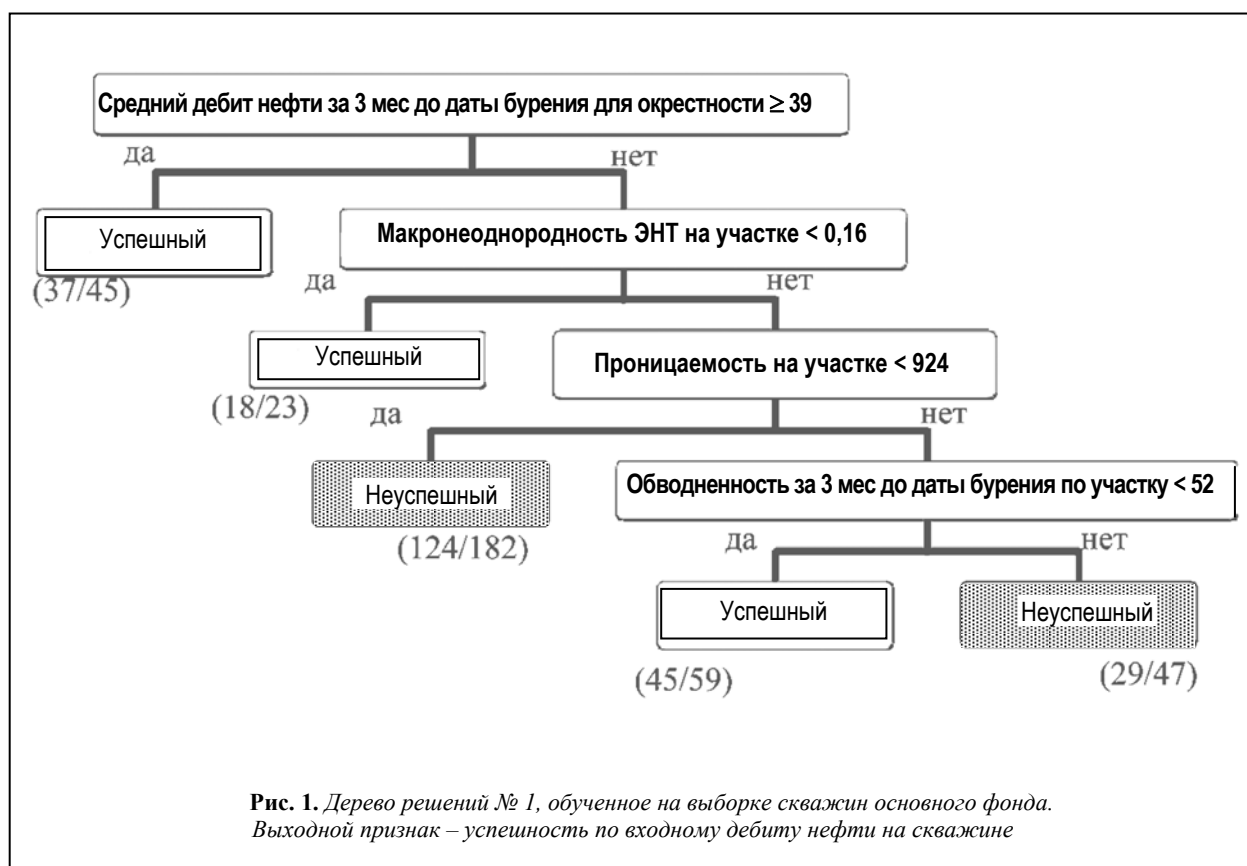
Корневым признаком дерева является "Средний дебит нефти для окрестности" с критерием 39 т/сут. Он делит выборку на две основные ветви. Уже на этом уровне селекции наблюдается существенное различие в успешности групп скважин, попавших в разные ветви. Так, из 45 скважин, пробуренных в зоне относительно высоких дебитов, значительная доля оказалась успешной. Это подтверждает достоверность, представленная в дереве (37 из 45 скважин).

Ветвь группы скважин с низким средним дебитом нефти по окрестности продолжает деление по признаку "Макронеоднородность на участке". В условиях пониженных дебитов, но при низких значениях макронеоднородности (< 0,16) прогнозируется высокий уровень успешности (достоверность 18 из 23). Высокая макронеоднородность оказывает преимущественно отрицательное влияние на эффективность скважин. Однако достоверность правила при этом остается невысокой (167

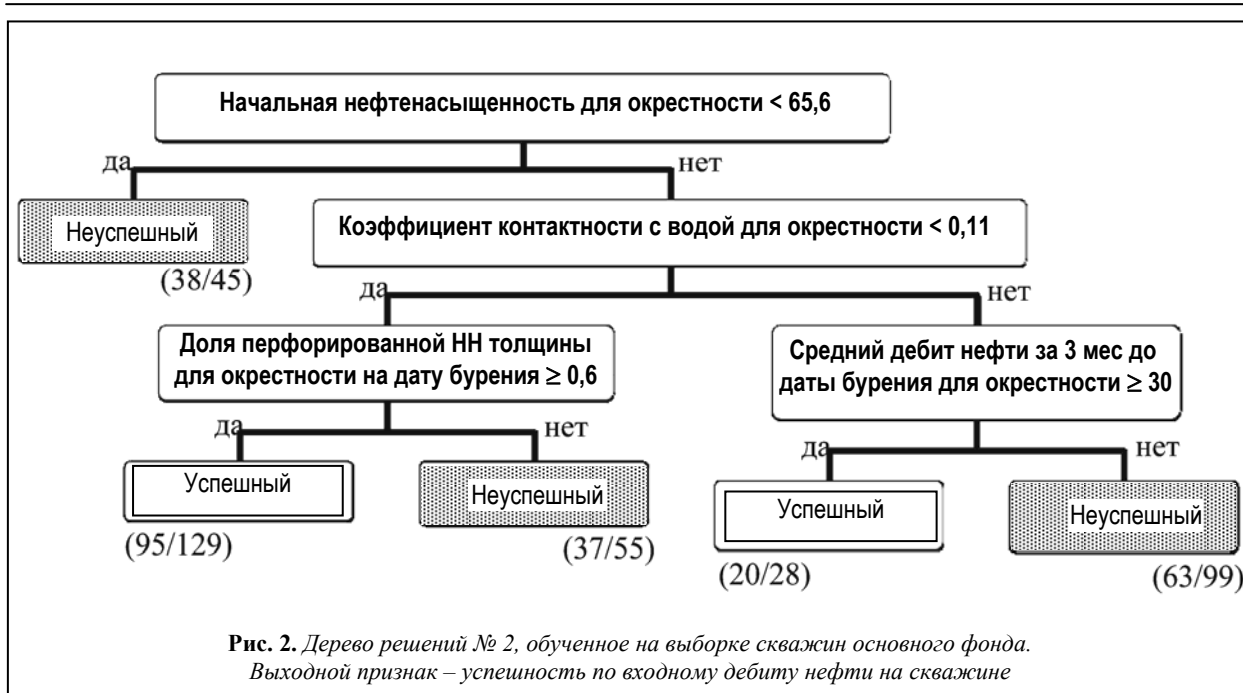
из 288 скважин, 58 %), поэтому группа скважин с высокой макронеоднородностью далее делится признаком "Проницаемость на участке" с граничным значением 924 мД. Так, при высоких значениях макронеоднородности для сохранения успешности скважин необходима повышенная проницаемость (> 924 мД). При низких значениях проницаемости и большой неоднородности коллектора прогнозируется получение низких дебитов нефти.

В условиях повышенной проницаемости для неоднородных участков рассматривается признак "Средняя обводненность по участку", который формирует следующее правило в дереве. Высокая обводненность участков ( $\geq 62\%$ ) указывает на вероятность неблагоприятного исхода. В случае умеренной обводненности (< 62 %) основная часть скважин рассматриваемой ветви (45 из 59) признается успешной.

Рассмотрим еще одно дерево, построенное по той же выборке основного фонда (рис. 2). Здесь, в отличие от первого варианта, корневой узел формирует признак, характеризующий запасы в окрестности скважины, а именно "Начальная нефтенасыщенность в окрестности". В условиях пониженной нефтенасыщенности (< 65,6) практически все скважины признаются неуспешными. Достоверность этого узла довольно высока (85 %), поэтому дальнейшая селекция продолжается по смежной ветви с преимущественно успешными скважинами.



**Рис. 1.** Дерево решений № 1, обученное на выборке скважин основного фонда. Выходной признак – успешность по входному дебиту нефти на скважине



В зонах с повышенной нефтенасыщенностью на участке в иерархии правил появляется следующий селективный признак "Коэффициент контактности с водой". При низких значениях контактности ( $< 0,11$ ) основная часть скважин признается успешной, высокие значения контактности указывают на вероятные проблемы. Поскольку достоверность правила в каждой из веток не слишком высока, обе ветви подлежат дальнейшему делению. В условиях низкой контактности с водой ( $< 0,11$ ) успешность скважины увязывается с высоким значением величины "Доля перфорированной толщины скважин окрестности" ( $\geq 0,6$ ). При повышенных значениях контактности с водой ( $\geq 0,11$ ) успешность скважин сопрягается с признаком "Средний дебит нефти для окрестности". Благоприятными являются значения дебита более 30 т/сут, при меньших значениях и высокой контактности с водой большая часть скважин группы признается неуспешной.

Таким образом, представленные два дерева решений позволяют уточнить иерархию влияния факторов, определить критерии успешности работы скважин, а также условия, когда высокие дебиты нефти на скважине маловероятны. Было выявлено, что основными факторами, влияющими на эффективность основного фонда, являются геологические признаки и характеристики запасов. Так, наиболее благоприятными являются условия низкой макронеоднородности и высокой проницаемости при наличии повышенной начальной нефтенасыщенности и низких значениях коэффициента контактности запасов с водой.

#### Уплотняющий фонд

Далее рассмотрим пример дерева решений, обученного уже на выборке скважин уплотняющего фонда (рис. 3).

Корневым признаком дерева здесь является "Средний дебит жидкости по участку" с пороговым значе-

нием 108 т/сут. Скважины со значением признака  $\geq 108$  т/сут с очень высокой достоверностью признаются успешными, иначе – неуспешными, но с достоверностью всего лишь 52 %. По данной группе скважин формируется еще одно правило: для достижения успешности генерируется требование "Пластовое давление по окрестности"  $\geq 19,6$  атм. В условиях низкого пластового давления почти все скважины являются неуспешными и дальнейшее деление ветви прекращается. Ветвь с более высоким уровнем пластового давления продолжает делиться за счет признака "Песчанистость в окрестности" с пороговым значением 0,66. Так, скважины в условиях повышенного пластового давления и песчанистостью  $\geq 0,66$  признаются успешными с довольно большой достоверностью. В варианте с низким значением песчанистости появляется еще один селективный признак геологического типа – "Эффективная нефтенасыщенная толщина на участке". Благоприятными для успешности скважин являются значения признака  $\geq 27$  м, в случае низких толщин велика вероятность неудачи (более 70 %). Последним селективным признаком становится параметр "Средний дебит нефти по участку". Он отбраковывает скважины со значениями среднего дебита нефти  $< 9,6$  т/сут.

Рассмотрим дерево № 2, построенное по выборке уплотняющего фонда (рис. 4). Здесь, так же как и в дереве № 1, во главе стоит признак, характеризующий окрестность с точки зрения технологии, а именно "Средний дебит нефти по участку". Практически все скважины со значением признака  $\geq 31$  т/сут являются успешными, и эта ветвь последующему делению не подлежит. У 48 % скважин другой ветви сохраняется вероятность положительного исхода. Следующим показателем, разделяющим эту группу скважин по признаку успешности, является "Пластовое давление для окрестности" с граничным значением 20 атм.

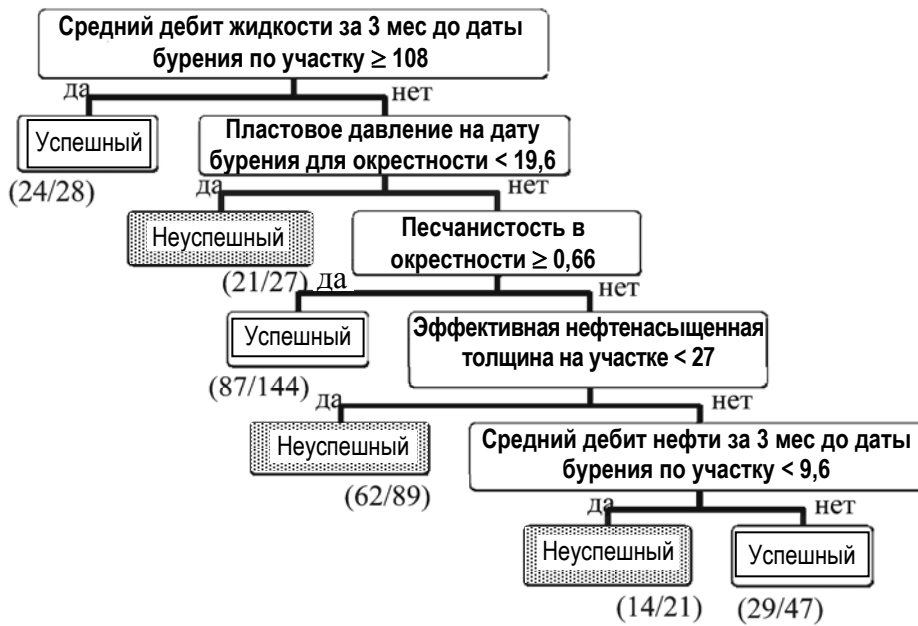


Рис. 3. Дерево решений № 1, обученное на выборке скважин уплотняющего фонда. Выходной признак – успешность по входному дебиту нефти на скважине



Рис. 4. Дерево решений № 2, обученное на выборке скважин уплотняющего фонда. Выходной признак – успешность по входному дебиту нефти на скважине

Аналогично дереву № 1 в условиях низких дебитов окрестных скважин повышающим шансы уплотняющего фонда условием признается высокий уровень пластового давления. Большинство скважин с пониженным давлением признаются неуспешными, и на этом данная ветвь дерева заканчивается. Следующим правилом ветки скважин с относительно высоким давлением выступает геологический при-

знак "Коэффициент контактности с водой". Благоприятным для уплотняющего фонда является значение контактности  $> 0,37$ . На первый взгляд это кажется не вполне логичным, однако принимая во внимание то, что зоны с повышенной контактностью, как правило, характеризуются относительно стабильным энергетическим состоянием, данный критерий становится вполне осмысленным. Достоверность

группы скважин с высокой контактностью на данном ответвлении довольно большая, поэтому дальнейшая селекция продолжается по смежной ветви.

Следующим признаком в иерархии правил выступает параметр системы разработки – "Расстояние до ближайшей нагнетательной скважины". Практически все скважины с близким расположением нагнетательных скважин (< 117 м) оказываются успешными. Очевидно, это связано с повышенным риском обводнения. При значениях признака  $\geq 117$  м основная часть скважин является успешной, однако параметр "Прокачанный нефтенасыщенный поровый объем по участку" продолжает селекцию успешных скважин. При значениях признака < 0,01 достигается максимальная эффективность на скважинах уплотняющего фонда. Большая часть скважин с высокими значениями прокачанного нефтенасыщенного порового объема признаются неуспешными в силу высокого риска обводнения.

Необходимо отметить, что основной проблемой объекта, скважины которого подлежат анализу, является пониженное пластовое давление. В связи с этим успешность уплотняющих скважин в значительной степени определяется факторами, характеризующими энергетическое состояние пласта: пластовое давление, средний дебит участка, коэффициент контактности запасов с водой и др.

#### Горизонтальный фонд

Теперь рассмотрим пример дерева решений, обученного на выборке скважин горизонтального фонда (рис. 5).

Корневым признаком дерева № 1 является параметр, характеризующий систему разработки "Плотность сетки скважин в окрестности". В случае с горизонтальным фондом данный признак оказывает существенное влияние на эффективность скважин: при

плотности сетки  $\geq 7,8$  га/скв. основная часть скважин признается успешной. Дальнейшая селекция осуществляется за счет признака "Накопленная обводненность участка". Практически все скважины со значением признака  $\geq 64$  % оказались успешными, с другой стороны, 86 % ГС на участках с низкой обводненностью оказались неуспешными. Обратная ситуация наблюдалась по основному фонду, для которого высокое значение обводненности окрестных скважин являлось неблагоприятным признаком. Это объясняется тем, что горизонтальные скважины принято бурить там, где основной фонд неэффективен.

При довольно плотной сетке скважин ( $\geq 7,8$  га/скв.) ГС признаются преимущественно неуспешными. Однако достоверность совсем невелика (56 %), поэтому данная ветвь продолжает делиться по правилу "Коэффициент контактности с водой для окрестности"  $\geq 0,24$ . Значит, для горизонтальных скважин благоприятны высокие значения контактности с водой. Ветвь группы скважин с высокой контактностью далее делится признаком расположения горизонтального ствола внутри пласта – "Расстояние от кровли пропластка до ствола скважины (нормированное на толщину пропластка)" с пороговым значением 0,34. Это правило (в условиях высокой контактности с водой) свидетельствует о том, что более благоприятная область размещения ствола расположена ближе к кровле коллектора, а именно в первой верхней трети части пропластка. При более низком расположении велика опасность обводнения скважины.

Таким образом, на горизонтальном фонде важную роль играют параметры сетки и расположение скважин внутри пласта. Рассмотрим следующее дерево, обученное на выборке горизонтальных скважин (рис. 6).

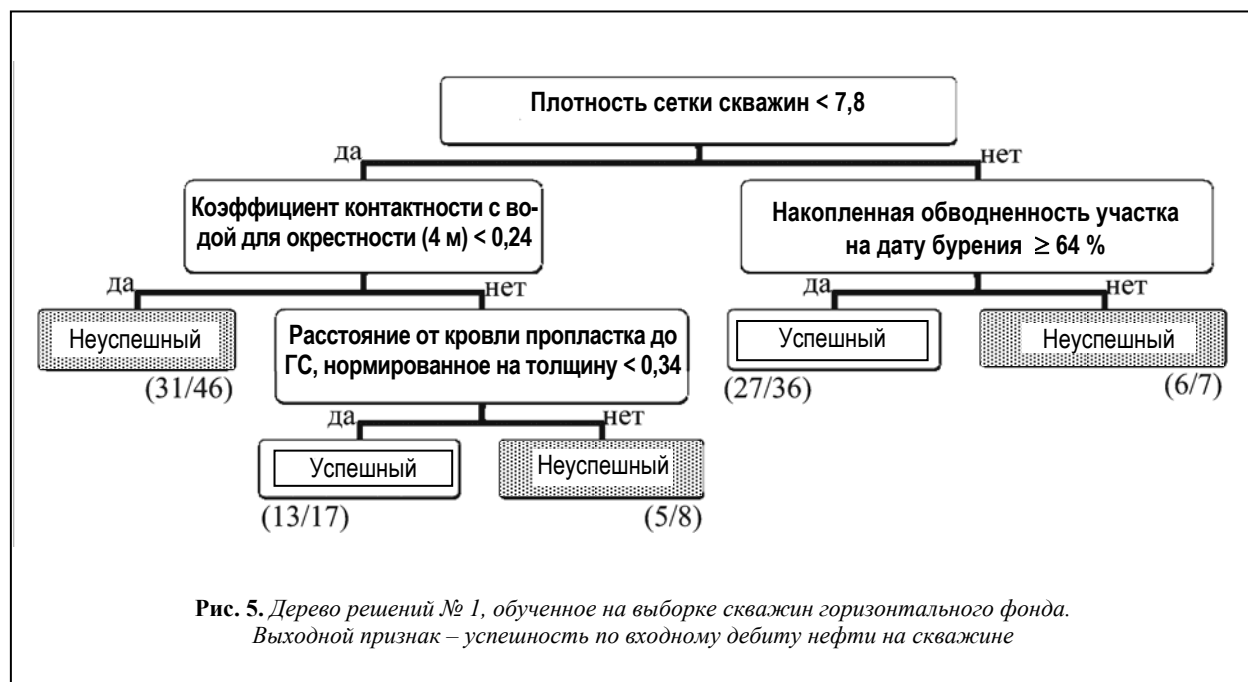


Рис. 5. Дерево решений № 1, обученное на выборке скважин горизонтального фонда. Выходной признак – успешность по входному дебиту нефти на скважине



Рис. 6. Дерево решений № 2, обученное на выборке скважин горизонтального фонда. Выходной признак – успешность по входному дебиту нефти на скважине

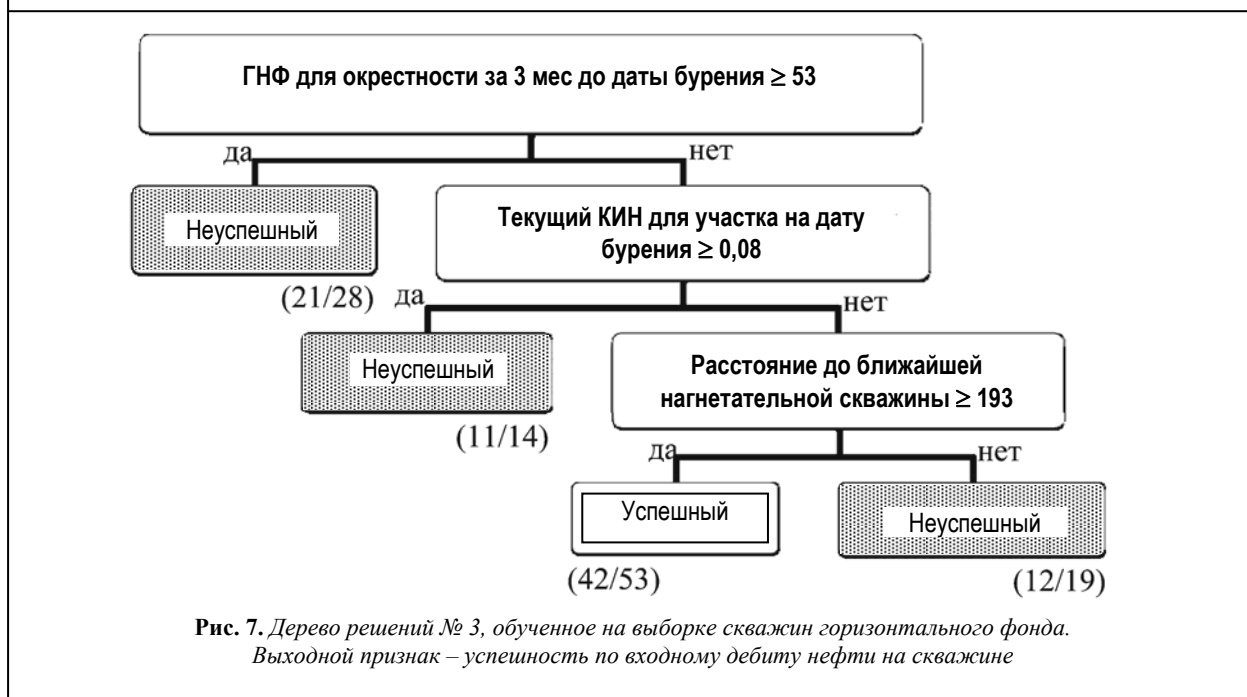


Рис. 7. Дерево решений № 3, обученное на выборке скважин горизонтального фонда. Выходной признак – успешность по входному дебиту нефти на скважине

Корневым признаком дерева № 2 является параметр запасов – "Текущий КИН для окрестности". Основная часть скважин со значениями  $< 0,1$  являются успешными, при значениях  $\geq 0,1$  преимущественно неудачными. То есть участки с довольно большой выработкой запасов неблагоприятны для бурения горизонтальных скважин. В таком случае скважины могут быть эффективны лишь при выполнении следующего правила: "Забойное давление окрестных скважин"  $\geq 30,5$  атм, т. е. при хорошем энергетическом состоянии участка.

Ветвь группы скважин с относительно низким текущим значением КИН далее делится с помощью

параметра системы разработки – "Расстояние до ближайшей нагнетательной скважины". Значительная доля скважин, располагающихся вблизи нагнетательных ( $< 171$  м), признаются неудачными. Расстояние  $\geq 171$  м является благоприятным, 68 % скважин в данных условиях являются успешными. Для достижения более высокой достоверности ветвь дерева подлежит дальнейшей декомпозиции. Следующее правило на рассматриваемой ветке – "Длина скважины"  $\geq 158$  м. Практически все скважины с большой длиной оказались успешными. Стоит отметить, что скважины с длиной  $< 158$  м также могут быть эффективными, однако только при выполне-

нии условия: "Средневзвешенная нефтенасыщенность по длине скважины"  $\geq 0,7$ . Все 100 % скважин с более низкой нефтенасыщенностью по длине ствола скважины признаются неуспешными.

Рассмотрим дерево № 3, обученное на выборке горизонтальных скважин (рис. 7). Во главе дерева № 3 стоит технологический признак – "ГНФ для окрестности".

При больших значениях ГНФ ( $\geq 53$ ) скважины с большой достоверностью признаются неуспешными. В данном случае этот признак выступает как маркер энергетического состояния: высокий газонефтяной фактор свидетельствует о плохом подержании пластового давления.

Ветвь группы скважин с низким ГНФ продолжает делиться признаком "Текущий КИН для участка". Снова большинство скважин с высокой выработкой запасов (текущий КИН  $\geq 0,08$ ) признаются неуспешными. В случае с низкими значениями признака появляется последнее правило в дереве – "Расстояние до ближайшей нагнетательной скважины" с граничным значением 193 м. При расстоянии  $\geq 193$  м имеем очень высокий показатель успешности, в то время как при более близком расположении нагнетательных скважин возникают существенные риски для горизонтального бурения.

Итак, благодаря использованию деревьев решений, установлено, что на данном объекте для скважин горизонтального фонда благоприятны большие длины стволов при не слишком уплотненной сетке скважин при большом расстоянии до нагнетательных скважин ( $> 170$  м), а также довольно высокая контактность запасов с водой, но в условиях проводки ствола в первой верхней трети пропластка.

Стоит отметить, что одиночные деревья решений не в силах охватить всего многообразия причинно-следственных связей, поэтому для большего понимания исследуемых процессов необходимо использовать целые комитеты деревьев решений или леса решений ("decisionforest") [5]. Очевидно при этом, что и прогнозные свойства таких комитетов обладают более высоким потенциалом по сравнению с одиночными деревьями. На рис. 8 представлен пример результата обучения модели с помощью другой разновидности деревьев решений – "Boosting Trees" [6] – на выборке горизонтальных скважин того же исследуемого объекта.

Обучение деревьев решений, в отличие от множества интеллектуальных систем прогнозирования, является способом формализованного анализа, который нацелен на извлечение новых знаний, правил и критериев. Деревья решений могут использоваться как для "работы над ошибками", так и для прогноза эффективности новых мероприятий.

Данный метод успешно применяется для прогнозирования ГРП, зарезки боковых стволов, бурения уплотняющих и горизонтальных скважин, перево-

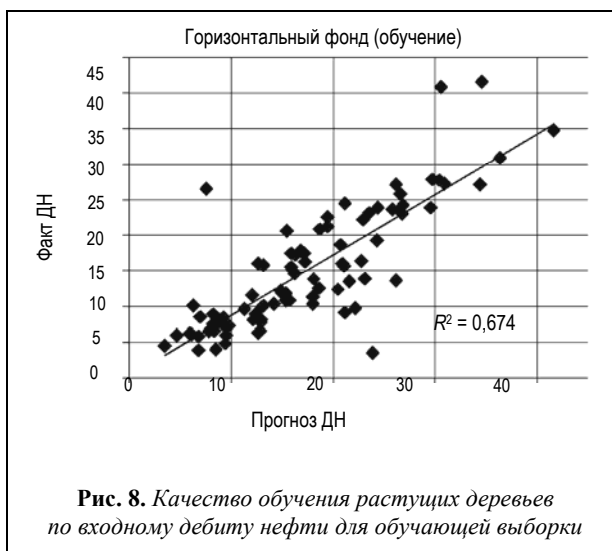


Рис. 8. Качество обучения растущих деревьев по входному дебиту нефти для обучающей выборки

дов с объекта на объект. Работы по прогнозированию ГТМ выполнялись и выполняются ЗАО "КОНКОРД" на различных крупных месторождениях Западной Сибири: Ватьеганское, Покачевское, Тевлинско-Русскинское (ОАО "ЛУКОЙЛ"); Вынгапуровское, Суторминское (ОАО "Газпром нефть"), а также на месторождениях Казахстана: Каракудук, Кумколь, Северные Бузачи.

Экспертные оценки наряду с практическим опытом свидетельствуют о том, что применение данной системы прогнозирования на месторождениях позволяет повысить успешность мероприятий на 15...25 %.

Представленный в работе методический подход получил высокую оценку ЦКР Роснедра [7] и был рекомендован к использованию при проектировании ГТМ.

#### ЛИТЕРАТУРА

1. Мирзаджанзаде А.Х., Хасанов Р.Н., Бахтизин Р.Н. *Этюды о моделировании сложных систем нефтедобычи. Нелинейность, неравновесность, неоднородность.* – Уфа: Гилем, 1999. – 464 с.
2. Zangl Georg, Hannerer Josef. *Data Mining: Applications in the Petroleum Industry.* Round Oak Publishing, 2003.
3. Quinlan R. *C4.5 Programs for Machine Learning.* Morgan Kaufmann, San Mateo, California, 1993.
4. Пичугин О.Н., Соляной П.Н., Фатихова Ю.З. *От "работы над ошибками" – к прогнозированию эффективности мероприятий // Нефть. Газ. Новации.* – 2012. – № 3.
5. Dieterich T.G. *An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees: Bagging, Boosting, and Randomization.* Machine Learning, 2000. – Vol. 40. – P. 139–157.
6. *STATISTICS Methods and Applications, StatSoft, Inc.* (2013).
7. *Протокол заседания нефтяной секции ЦКР Роснедра от 03.12.2009 г. № 4752 "Методические подходы к обоснованию выбора скважин-кандидатов для проведения геолого-технических мероприятий с использованием деревьев решений". URL: http://www.naen.ru/files/4752.pdf*