

УДК 622.276

ПРИМЕНЕНИЕ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ ПЛАСТОВОГО ДАВЛЕНИЯ ПРИ РАЗРАБОТКЕ НЕФТЯНЫХ МЕСТОРОЖДЕНИЙ

Мартюшев Дмитрий Александрович¹,
martyushevdi@inbox.ru

Пономарева Инна Николаевна¹,
permpolitech@gmail.com

Захаров Лев Андреевич²,
lzakharov-ng@mail.ru

Шадров Тимур Артурович³,
shadrov.timur@yandex.ru

¹ Пермский национальный исследовательский политехнический университет, Россия, 614990, г. Пермь, пр. Комсомольский, 29.

² Филиал ООО «ЛУКОЙЛ-Инжиниринг» «ПермНИПИнефть» в г. Перми, Россия, 614000, г. Пермь, ул. Советской Армии, 29.

³ ООО «Диджитал фьюче системс», Россия, 614066, г. Пермь, шоссе Космонавтов, 111.

Актуальность исследования обусловлена тем, что появление искусственного интеллекта в нефтяной промышленности привело к увеличению его использования при разведке, разработке, добыче, проектировании месторождений и планировании управления, чтобы ускорить принятие решений, сократить затраты и время. Машинное обучение приобрело большую популярность в установлении взаимосвязи между сложными нелинейными наборами данных и продемонстрировало свое превосходство над методами регрессии в нефтяной инженерии с точки зрения ошибок прогнозирования данных большой размерности, вычислительной мощности и памяти. В представленной статье рассматривается применение машинного обучения для оценки его эффективности и потенциала для определения и прогнозирования значений пластового давления при разработке нефтяных месторождений по сравнению с обычными статистическими моделями нефтегазовой инженерии.

Цель: оценка возможностей расчета и прогнозирования пластового давления методом машинного обучения «случайный лес».

Объект: динамика пластового давления при разработке терригенных отложений нефтяных месторождений Пермского края.

Методы: методы вероятностно-статистического анализа и машинного обучения «random forest regression» («случайный регрессионный лес»).

Результаты. Предлагается новый метод прогнозирования пластового давления с использованием машинного обучения, основанный на непараметрической многомерной модели, связывающей показатели эксплуатации скважины во времени. Предлагаемый метод учитывает динамику показателей, характеризующих эксплуатацию скважин, а прогнозируемое пластовое давление хорошо коррелируется с измеренными с помощью гидродинамических исследований скважин значениями. Установлено, что метод машинного обучения «случайный лес» обеспечивает лучшую достоверность прогнозирования пластового давления, чем метод линейной регрессии. Перспективами дальнейшего развития является дополнительное «обучение» модели «случайного леса» и оценка возможности применения других методов машинного обучения для решения поставленной задачи, в том числе расширение набора факторов для более точного моделирования пластового давления.

Ключевые слова:

Искусственный интеллект, машинное обучение, многомерная модель, случайный лес, показатели эксплуатации скважин, текущее пластовое давление, корреляция.

Введение

На сегодняшний день большинство нефтяных месторождений России и Пермского края находятся на поздних стадиях разработки. Данные месторождения требуют оперативного контроля параметров разработки и проведения большого количества исследований продуктивных пластов [1]. Одним из важнейших показателей разработки является пластовое давление, которое в основном определяется в результате проведения гидродинамических исследований скважин (ГДИС). Точный прогноз пластового давления имеет широкое применение в нефтяной промышленности, особенно при оптимизации непрерывной добычи на месторождении, количественной оценке продуктив-

ности коллектора, корректировке стоимости добычи нефти и оценке ремонтных работ [2]. Контроль за его величиной – важнейшая задача эффективного мониторинга разработки нефтяных месторождений. Методы ГДИС являются основными в практике нефтепромыслового дела, их применение с целью контроля за энергетическим состоянием залежи в зонах дренирования скважин предписывается руководящими документами. Основным недостатком является необходимость остановки скважины, в ряде случаев на весьма продолжительное время, что приводит к так называемым недоборам добычи нефти. Кроме этого, представляется достаточно сложным выполнять сравнения пластовых давлений между собой во всех скважинах

ввиду одновременности проведения исследований, поскольку одновременно остановить весь фонд в условиях промысла невозможно [3].

В условиях современной нефтедобычи актуальной задачей является широкое применение цифровых технологий для решения различных проблем нефтегазового производства [4]. Их решение осложняет необходимость учета влияния геолого-технологических показателей на процесс разработки нефтяных и газовых месторождений. Стоит отметить, что даже хорошо изученные объекты разработки характеризуются значительной степенью неоднородности параметров пласта и технологических показателей, что значительно усложняет использование цифровых технологий для решения актуальных проблем нефтегазового производства [5]. В этой связи представляется актуальным исследование применимости методов вероятностно-статистического анализа и машинного обучения к решению данных проблем.

Появление искусственного интеллекта в нефтяной промышленности привело к увеличению его использования при разведке, разработке, добыче, проектировании месторождений и планировании управления, чтобы ускорить принятие решений, сократить затраты и время. Машинное обучение приобрело большую популярность в установлении взаимосвязи между сложными нелинейными наборами данных. Этот тип алгоритма машинного обучения продемонстрировал свое превосходство над методами регрессии в нефтяной инженерии с точки зрения ошибок прогнозирования данных большой размерности, вычислительной мощности и памяти [6, 7]. Это приводит к более быстрому принятию решений, что неизменно помогает сэкономить деньги, время и оборудование. Усовершенствованный и улучшенный процесс определения характеристик коллектора требует высокого уровня точности, устойчивого к любым ожидаемым или неожиданным изменениям, чтобы снизить потери с точки зрения человеко-часов и инвестиций [8].

Методы машинного обучения получают все большее применение в различных отраслях, в том числе и в нефтегазовой [9–15]. Все нефтедобывающие компании мира обрабатывают большой объем цифровой информации и темпы роста объема данных повышаются с каждым годом. Качество их обработки и интерпретации является основой принятия эффективных проектных и управленческих решений. В этой связи адаптация методов машинного обучения к нефтегазовой сфере с целью создания автоматизированных систем по контролю параметров эксплуатации нефтяных месторождений имеет большой потенциал [16–22].

Так, например, в компании ООО «Газпромнефть НТЦ» используют технологии машинного обучения для выявления причин сбоев в работе электроцентробежных насосов, а также выделяют для себя несколько приоритетных направлений с применением данных методов – поиск объектов-аналогов, восстановление исторических эксплуатационных данных, обработка данных исследований в режиме реального времени и др. Благодаря внедренной технологии компании уда-

лось снизить эксплуатационные затраты на 2 млн долл. [13].

На сегодняшний день накоплен значительный опыт по разработке нефтяных месторождений Пермского края, имеется обширная база результатов гидродинамических исследований и технологических показателей эксплуатации скважин, что позволяет исследовать применение цифровых технологий для решения различных производственных задач. При формировании большого количества технологической информации представляется возможным использование методов, основанных на сборе, систематизации, обработке и интерпретации данных, представленных в виде цифровых массивов.

В работе [23] описан подход, основанный на построении многоуровневых вероятностно-статистических моделей прогнозирования пластового давления. Использование разработанных многомерных математических моделей позволяет определить пластовое давление в любой период эксплуатации скважин без их остановки на исследование. Стоит отметить, что представленные модели не следует рассматривать в качестве альтернативы гидродинамическим исследованиям. Их применение целесообразно для экспресс-оценки пластового давления либо при невозможности остановки скважины на исследование по технологическим причинам.

Данная методика является наиболее достоверной и адаптированной из известных и поэтому представляется возможным ее применение и на других месторождениях Пермского края (России и мира), а также, учитывая опыт ее применения, исследование методов машинного обучения для определения значений пластового давления в режиме реального времени и для воспроизведения исторической работы скважин.

Принципиально же задачи, решаемые с помощью методов машинного обучения, можно разделить на следующие виды:

- 1) регрессия – предсказание конкретного числа на основе массива признаков или характеристик [24, 25];
- 2) классификация – определение по количеству и качеству признаков или характеристик категории объекта исследования [26];
- 3) кластеризация – объединение объектов в группы по общему признаку [27];
- 4) уменьшение размерности – сжатие массива характеристик объекта до меньшего количества признаков [28, 29].

Приведенные задачи повсеместно встречаются при анализе систем разработки нефтяных и газовых месторождений, его неотъемлемой частью является контроль за энергетическим состоянием объекта разработки, в ходе которого определяется параметр пластового давления. Для определения данного параметра на нефтяных месторождениях Пермского края описанные подходы ранее не использовались, в связи с чем актуально исследовать их применимость и оценить дальнейшие перспективы их применения и развития.

**Исходные данные для оценки
и прогнозирования пластового давления**

В качестве объекта исследования выбрано одно из перспективных нефтяных месторождений Пермского края (объект Бб). Исходные данные для построения моделей использованы с трех других нефтяных месторождений (объекты Бб) Пермского края, которые характеризуются значительным жизненным циклом эксплуатации и объемом промысловой информации. Данные месторождения являются хорошо изученными и имеют достаточное количество фактических замеров пластового давления. Основные сведения о разработке данных месторождений приведены в табл. 1.

Таблица 1. Основные сведения о нефтяных месторождениях, использованные для построения исходных моделей

Table 1. Basic information about oil fields used to build the original models

Параметр Parameter	Месторождение/Field		
	1	2	3
Количество скважин Number of wells	112	48	68
Количество ГДИ Number of wells tests	349	212	231
Начальное пластовое давление, МПа Initial reservoir pressure, MPa	21,2	23,4	22,5
Текущее пластовое давление, МПа Current reservoir pressure, MPa	11,1	9,5	7,2

Объяснение моделей машинного обучения всегда является важной темой исследований [30]. Простые модели машинного обучения, такие как линейная регрессия и деревья решений, легко понять и объяснить. Для линейной регрессии вклад каждой переменной определяется знаком и величиной ее коэффициента. Деревья решений можно интерпретировать путем визуализации внутренних узлов и ветвей. Однако сложные методы нелинейного машинного обучения, такие как опорная векторная регрессия, «случайные леса» и глубокие нейронные сети, трудны для понимания, даже если они всегда обеспечивают более высокую достоверность, чем более простые методы машинного обучения.

Для оценки и прогнозирования пластового давления использованы два метода: множественная линейная регрессия и «random forest regression» («случайный регрессионный лес»). Метод машинного обучения «случайный лес» получил широкое применение во многих сферах и отлично подходит для решения различного рода задач [10, 26]. Данный алгоритм машинного обучения впервые предложен американскими математиками Лео Брайманом и Адель Катлер и является одним из немногих универсальных алгоритмов. Его универсальность заключается в том, что он подходит для решения задач классификации, регрессии, кластеризации, поиска аномалий и т. д. Принципиально «случайный регрессионный лес» – это множество деревьев решений, в которых при решении задачи регрессии их ответы усредняются, что подходит для расчета параметра пластового давления.

Модель «случайного леса» описывается следующими характеристиками:

- количество решающих деревьев – от данного фактора зависит качество результата, однако при увеличении числа деревьев время настройки и работа модели также увеличиваются;
- максимальная глубина решающего дерева в модели – при увеличении данного фактора возрастает качество обучения, однако стоит учитывать, что при решении задач с большим количеством шумовых данных (выбросов) рекомендуется использовать неглубокие деревья;
- максимальное количество узлов решающего дерева (ширина) – при определении данного параметра необходимо учитывать, что он может привести к переобучению модели при небольшой глубине самого дерева;
- максимальное число признаков одного решающего дерева – при увеличении данного фактора время построения леса возрастает, а деревья становятся однообразными, для задач регрессии по умолчанию он равен $n/3$, где n – число деревьев.

Данные характеристики адаптированы под решение задач воспроизведения и прогнозирования значений пластового давления.

**Прогнозирование пластового давления
с помощью методов машинного обучения**

На первом этапе необходима предобработка и структуризация промысловых данных (дебит жидкости; коэффициент эксплуатации; забойное давление; начальное пластовое давление). Создана программа для ЭВМ «Squage» с целью автоматизации анализа промысловых данных и построения математических моделей, в основу алгоритмов которых легли вышеописанные методы.

Используя вероятностно-статистическую модель множественной линейной регрессии и метод машинного обучения «случайный регрессионный лес» воспроизведены исторические замеры пластового давления с целью проверки достоверности разработанных моделей.

Уравнение множественной линейной регрессии получено методом наименьших квадратов и имеет следующий вид:

$$P_{пл(t)} = 0,7548P_{пл(t-1)} + 0,0131 \frac{(Q_{ж(t)} - Q_{ж(t-1)})}{Q_{ж(t)}} + 0,207P_{заб(t)} - 0,00001T + 1,2851,$$

где $P_{пл(t)}$ – прогнозируемое пластовое давление; $P_{пл(t-1)}$ – пластовое давление, предшествующее прогнозу; $\frac{(Q_{ж(t)} - Q_{ж(t-1)})}{Q_{ж(t)}}$ – темп роста дебита жидкости (далее T_q) относительно предыдущего ГДИ; $Q_{ж(t)}$ – дебит жидкости в сутки в среднем за сяц; $P_{заб(t)}$ – забойное давление на текущую дату; T – время эксплуатации скважины.

В ходе проверки значимости коэффициентов уравнения линейной регрессии получены следующие результаты с помощью p -критерия (табл. 2).

Таблица 2. Значимость коэффициентов уравнения линейной регрессии

Table 2. Significance of the coefficients of the linear regression equation

Параметр/Parameter	p-критерий/p-criterion
Свободный член/Free member	0,000*
$P_{пл(t-1)}$	0,000*
T_q	0,005*
$P_{заб(t)}$	0,000*
T	0,000*

Примечание/Note: * – признак значимости/sign of significance.

В результате расчетов вычислено среднее абсолютное отклонение модели на входных данных, которое составило 0,821 МПа, при $R^2=0,757$.

Для построения модели «случайный регрессионный лес» использовались следующие параметры:

- количество деревьев – 200;
- максимальная глубина – 5;
- максимальное число признаков одним деревом – 3.

После обучения модели «случайного леса» рассчитаны коэффициенты значимости факторов. Значимость фактора в «случайном лесе» определяется его совокупной важностью для каждого решающего дерева, т. е. мерой уменьшения разнородности Джини (табл. 3).

Среднее абсолютное отклонение на входных данных модели «случайного леса» составило 0,812 МПа.

Для вышеописанных методов оценивались рабочие характеристики моделей при помощи подхода

перекрестной проверки («кросс-валидация»). Данная процедура подразумевает разделение выборки на равные части, далее каждая из частей выборки последовательно исключается (отложенная выборка) и строится модель на оставшихся данных, а на отложенной части проверяется величина ошибки. В результате данной проверки стандартное отклонение для модели линейной регрессии составило $1,071 \pm 0,14$ МПа, а для модели «случайного леса» – $1,018 \pm 0,17$ МПа. Данные величины указывают на стабильность моделей, что означает хорошую вероятность получения достоверного результата в случае, если ни одна из моделей ранее не «обучалась» на вводимых исходных данных.

Таблица 3. Важность факторов модели «случайного леса»

Table 3. Importance of factors in the «random forest» model

Фактор Factor	Коэффициент значимости Significance coefficient
$P_{пл(t-1)}$	0,815158
T_q	0,023523
$P_{заб(t)}$	0,132228
T	0,029091

Для оценки достоверности метода линейной регрессии и метода «случайный лес» построены зависимости фактических (1553 измерения) и рассчитанных замеров пластового давления (рис. 1, 2).

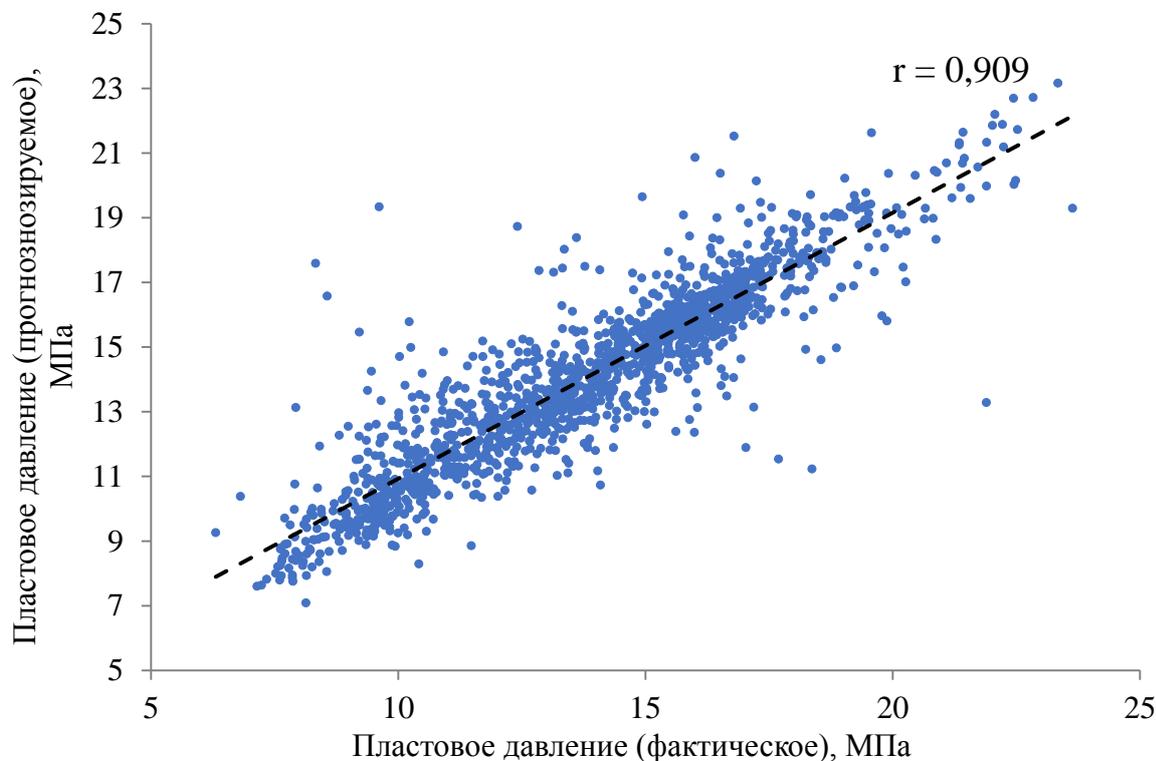


Рис. 1. Поле корреляции между фактическими и рассчитанными значениями пластового давления методом линейной регрессии

Fig. 1. Correlation field between actual and calculated values of reservoir pressure by linear regression method

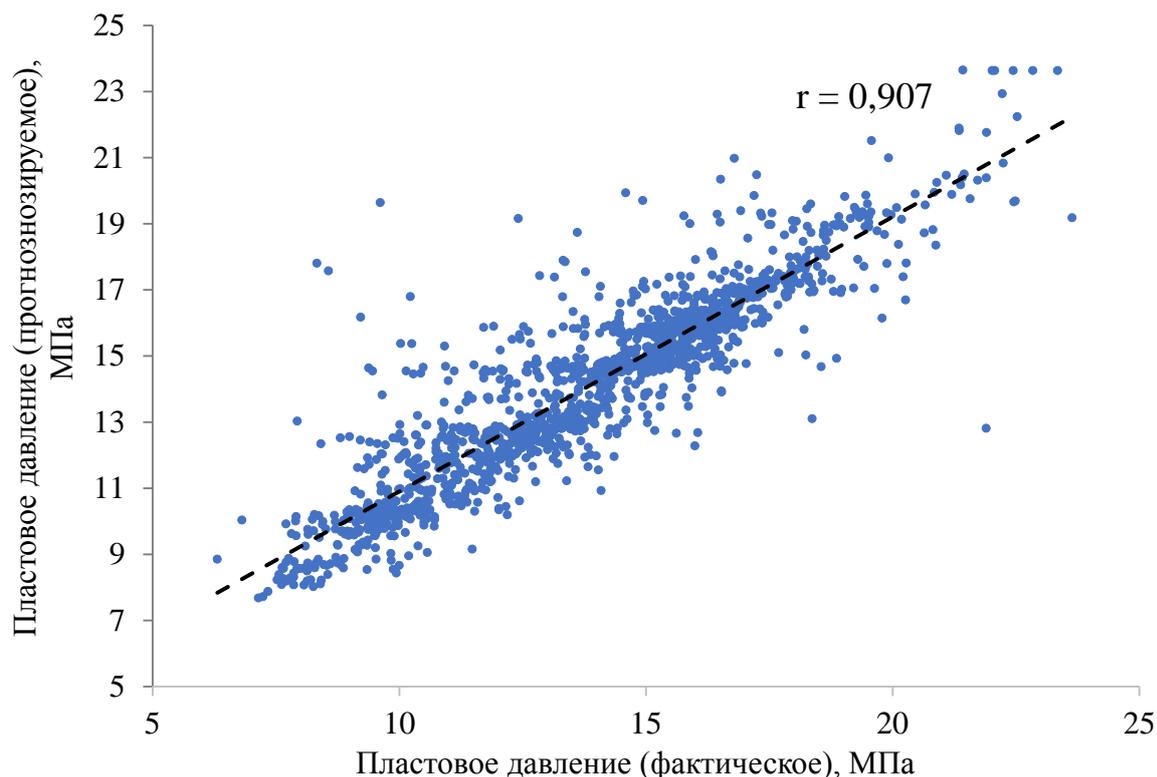


Рис. 2. Поле корреляции между фактическими и рассчитанными значениями пластового давления методом «случайного леса»

Fig. 2. Correlation field between the actual and calculated values of reservoir pressure by the «random forest» method

Анализируя представленные графики, можно сделать вывод, что в обоих случаях рассчитанные параметры пластового давления имеют «плотное» распределение с фактическими замерами, что говорит о хорошей сходимости результатов в целом по объекту исследования. Отклонения, полученные в результате применения методов линейной регрессии и «случайного леса» по всей исследуемой выборке, представлены в табл. 4.

Таблица 4. Отклонения в результате применения методов линейной регрессии и «случайного леса»

Table 4. Deviations due to linear regression and «random forest» method

Метод Method	Абсолютное отклонение от фактического измерения (среднее), МПа Absolute deviation from the actual measurement (average), MPa	Относительное отклонение от фактического измерения (среднее), % Relative deviation from the actual measurement (average), %
Линейная регрессия Linear regression	0,87	6,9
«Случайный лес» «Random forest»	0,86	6,8

Таким образом, можно отметить, что методы линейной регрессии и «случайного леса» имеют равное минимальное отклонение прогнозируемых значений пластового давления от фактических, что указывает

на эффективность и перспективность использования данных методов.

Учитывая «неоднородность» выборки и большой объем данных, необходимо провести поскважинное сопоставление результатов. С этой целью построены графики сопоставления результатов фактических и расчетных значений пластового давления (рис. 3–5). Выбор скважин для демонстрации полученных данных произведен таким образом, чтобы отразить наиболее полную картину применимости используемых методов.

Анализируя представленные графики сопоставления фактических и рассчитанных значений пластового давления, можно сделать вывод, что оба метода показывают хорошую сходимость с историческими данными при решении задачи воспроизведения «падающей» динамики исследуемого параметра. Однако в некоторых случаях метод «случайного леса» показывает лучшую сходимость. Так, например, на скв. 256 и 116 общий тренд пластового давления данным методом смоделирован ближе к факту. Особое внимание стоит обратить на расчет последнего замера пластового давления, так как он имеет наибольшую значимость при прогнозировании данного параметра. Высокая степень сходимости данной точки позволяет сделать вывод, что математическая модель достоверно отображает текущее энергетическое состояние скважин и объекта разработки. В этом ключе метод «случайного леса» также показывает лучшую сходимость по сравнению с методом линейной регрессией. Однако стоит отметить, что ни один из исследуемых

способов не смог смоделировать резкие перепады пластового давления (рис. 5). В этой связи необходи-

ма доработка методики контроля энергетического состояния залежи с учетом полученного опыта.

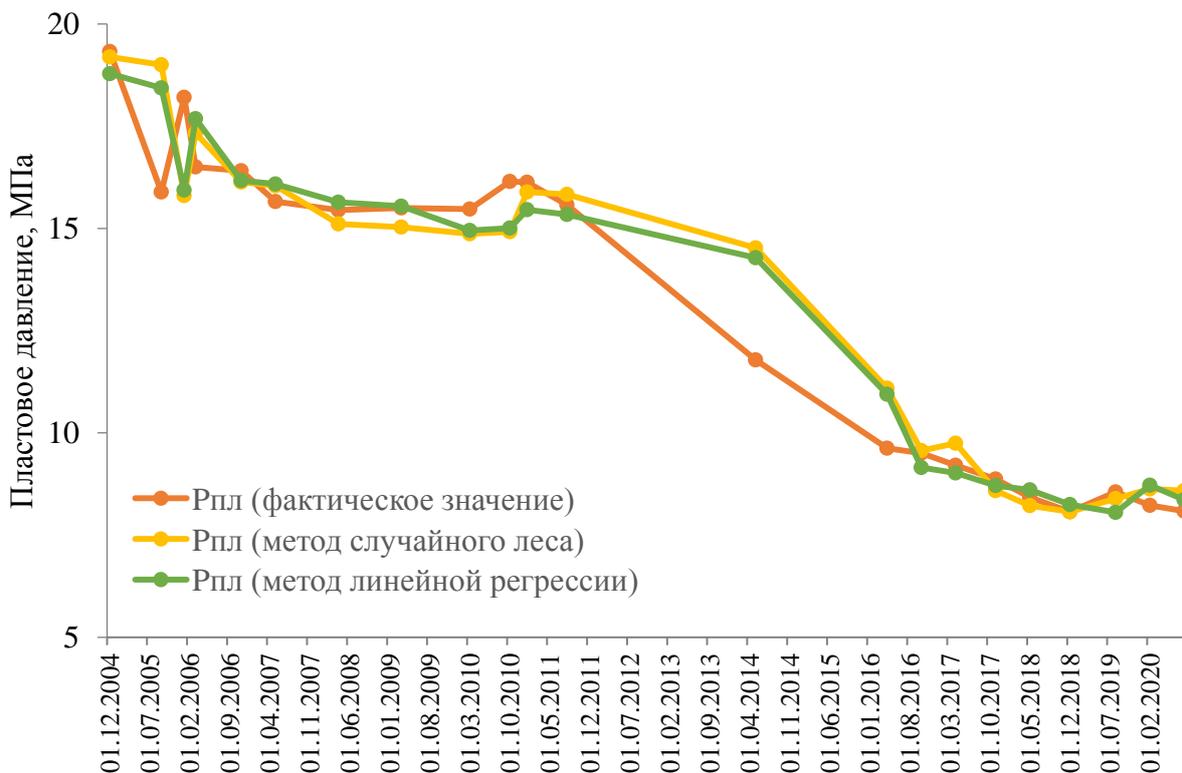


Рис. 3. Сопоставление фактических и рассчитанных значений пластового давления по скв. 256

Fig. 3. Comparison of actual and calculated values of reservoir pressure for well 256

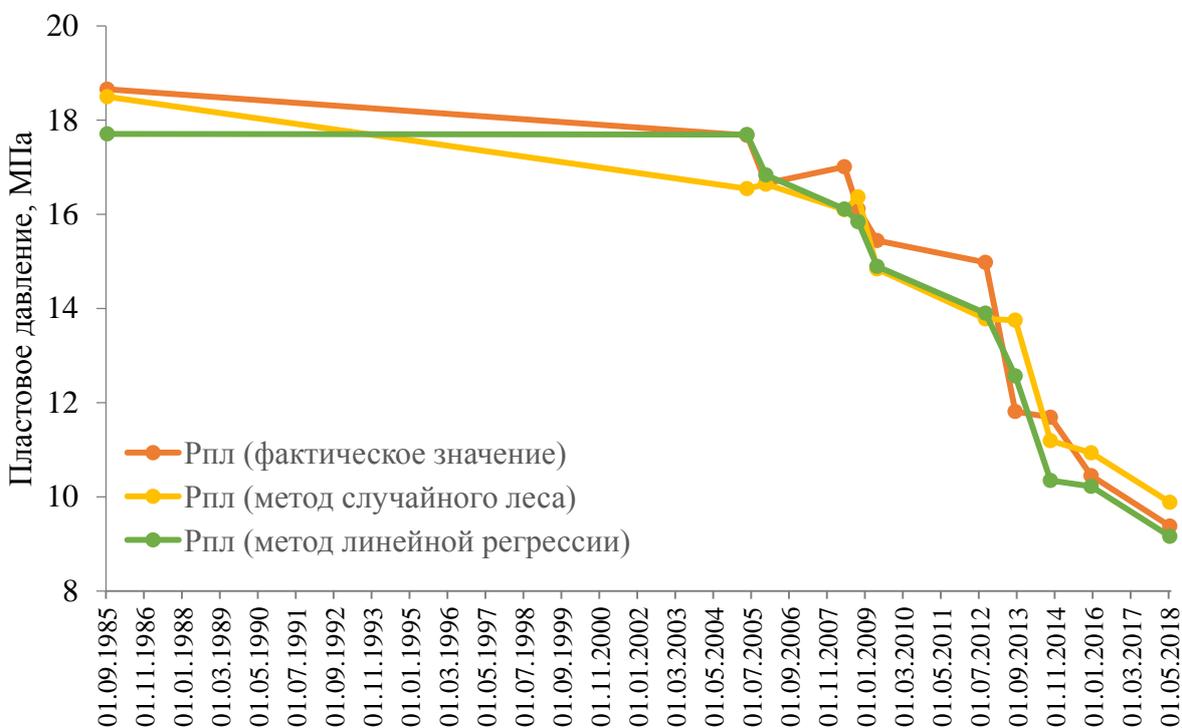


Рис. 4. Сопоставление фактических и рассчитанных значений пластового давления по скв. 116

Fig. 4. Comparison of the actual and calculated values of the reservoir pressure parameter for well 116

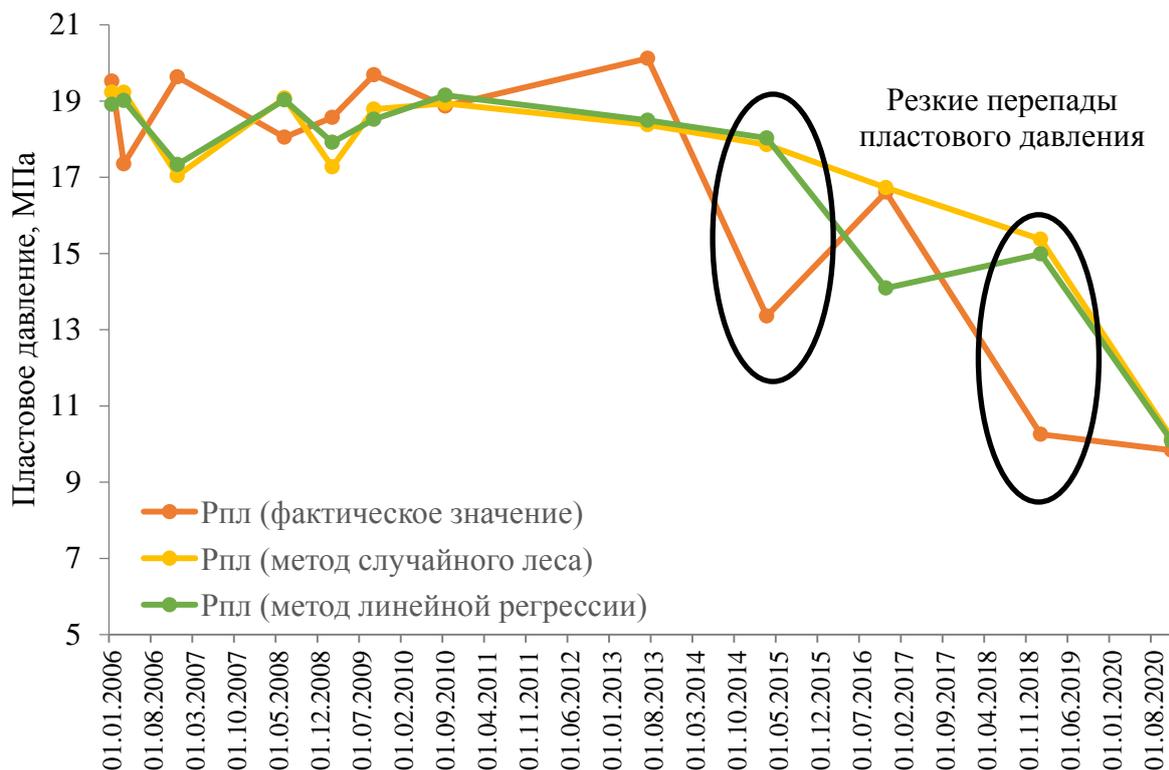


Рис. 5. Сопоставление фактических и рассчитанных значений пластового давления по скв. 45
 Fig. 5. Comparison of the actual and calculated values of the reservoir pressure parameter for well 45

В общем случае оба метода показали хорошие результаты по воспроизведению фактических значений параметра пластового давления и могут применяться для решения производственных задач с участием эксперта для оценки «выбросов» в получаемых данных. Также необходимо дополнительное «обучение» модели «случайного леса» и оценка возможности применения других методов машинного обучения для решения поставленной задачи, в том числе расширение набора факторов для более точного моделирования пластового давления.

Заключение

За прошедшие годы в нефтяной промышленности накопилось слишком много информации, поэтому алгоритмы машинного обучения, которые могут обрабатывать многомерные и сложные данные, являются гораздо более предпочтительными по сравнению с эмпирическими корреляциями и моделями линейной регрессии.

В представленном исследовании предлагается новый метод прогнозирования пластового давления с использованием метода машинного обучения, осно-

ванный на непараметрической многомерной модели, связывающей показатели эксплуатации скважины во времени. Предлагаемый метод учитывает динамику показателей, характеризующих эксплуатацию скважин, а прогнозируемое пластовое давление хорошо коррелируется с измеренными с помощью ГДИ значениями. В исследовании установлено, что метод машинного обучения «случайный лес» обеспечивает лучшую достоверность прогнозирования пластового давления, чем метод линейной регрессии. Кроме того, предлагаемый метод позволяет избежать утомительной процедуры калибровки коэффициентов по сравнению с методами, основанными на параметрических преобразованиях.

На основе рассчитанного значения пластового давления с помощью машинного обучения можно определить режим разработки залежи в данный момент, заблаговременно проектировать систему поддержания пластового давления или оценить ее эффективность, а также обосновано принимать дальнейшие рациональные решения по разработке нефтяных месторождений.

СПИСОК ЛИТЕРАТУРЫ

1. Молчанов А.А., Агеев П.Г. Внедрение новых технологий – надежный путь извлечения остаточных запасов месторождений углеводородов // Записки Горного Института. – 2017. – Т. 227. – С. 530–539.
2. Шамсиев М.Н. Интерпретация результатов термогазодинамических исследований вертикальных газовых скважин, несовершенных по степени вскрытия пласта // Георесурсы. – 2018. – Т. 20. – № 1. – С. 39–43.
3. Елесин А.В., Кадырова А.Ш., Никифорова А.И. Определение поля проницаемости пласта по замерам давления на скважинах с использованием сплайн-функции // Георесурсы. – 2018. – Т. 20. – № 2. – С. 102–107.
4. Бобб И.Ф. Международный опыт создания нефтегазовых ИТ-технологий для моделирования месторождений // Георесурсы. – 2018. – Т. 20. – № 3. – С. 193–196.
5. Белозеров И.П., Губайдуллин М.Г. О концепции технологии определения фильтрационно-емкостных свойств терриген-

- ных коллекторов на цифровой модели ядра // Записки Горного Института. – 2020. – Т. 244. – С. 402–407.
6. Jamiu Oyekan Adegbite, Hadi Belhaj, Achinta Bera. Investigations on the relationship among the porosity, permeability and pore throat size of transition zone samples in carbonate reservoirs using multiple regression analysis, artificial neural network and adaptive neuro-fuzzy interface system // Petroleum Research. – 2021. URL: <https://doi.org/10.1016/j.ptlrs.2021.05.005> (дата обращения 15.08.2021).
 7. Ponomareva I.N., Galkin V.I., Martyshev D.A. Operational method for determining bottom hole pressure in mechanized oil producing wells, based on the application of multivariate regression analysis // Petroleum Research. – 2021. URL: <https://doi.org/10.1016/j.ptlrs.2021.05.010> (дата обращения 15.08.2021).
 8. Belhouchet H.E., Benzagouta M.S., Dobbi A., Alquraishic A., Duplay J. A new empirical model for enhancing well log permeability prediction, using nonlinear regression method: Case study from Hassi-Berkine oil field reservoir – Algeria // Journal of King Saud University - Engineering Sciences. – 2021. – Vol. 33. – Issue 2. – P. 136-145.
 9. Wong P.M., Jian F.X., Taggart I.J. A critical comparison of neural networks and discriminant analysis in lithofacies, porosity and permeability predictions // Journal Petroleum Geology. – 1995. – V. 18. – P. 191–206.
 10. Nikraves M., Aminzadeh F. Ch. 1. Soft computing for intelligent reservoir characterization and modeling // Developments in Petroleum Science. – 2003. – V. 51. – P. 3–32.
 11. Применение методов искусственного интеллекта для выявления и прогнозирования осложнений при строительстве нефтяных и газовых скважин: проблемы и основные направления решения / А.Д. Черников, Н.А. Еремин, В.Е. Столяр, А.Г. Сбоев, О.К. Семенова-Чашина, Л.К. Фицнер // Георесурсы. – 2020. – Т. 22. – № 3. – С. 87–96.
 12. Лаломов А.В., Бочнева А.А., Чефранов Р.М. Разработка цифровой системы прогнозирования коренных источников золота по результатам шлифового опробования на примере Вагранского россыпного узла (Северный Урал) // Георесурсы. – 2020. – Т. 22. – № 2. – С. 67–76.
 13. Евсюткин И.В., Марков Н.Г. Глубокие искусственные нейронные сети для прогноза значений дебитов добывающих скважин // Известия Томского политехнического университета. Инжиниринг георесурсов. – 2020. – Т. 331. – № 11. – С. 88–95.
 14. Применение технологий глубокого обучения для изучения шлифов на примере Усинского месторождения нефти / Н.А. Попов, И.С. Путилов, А.А. Гуляева, Е.Е. Винокуров // Известия Томского политехнического университета. Инжиниринг георесурсов. – 2020. – Т. 331. – № 6. – С. 100–112.
 15. Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models / D. Asante Otchere, T.O. Arbi Ganata, R. Gholami, S. Ridhaa // Journal of Petroleum Science and Engineering. – 2020. – 108182.
 16. Harpreet Singh. Machine learning for surveillance of fluid leakage from reservoir using only injection rates and bottomhole pressures // Journal of Natural Gas Science and Engineering. – 2019. – V. 69. – 102933.
 17. Shahab D. Mohaghegh. Subsurface analytics: contribution of artificial intelligence and machine learning to reservoir engineering, reservoir modeling, and reservoir management // Petroleum Exploration and Development. – 2020. – V. 47. – Iss. 2. – P. 225–228.
 18. A new method of predicting the saturation pressure of oil reservoir and its application / Guoyi Yu, Feng Xu, Yingzhi Cui, Xiangling Li, Chujuan Kang, Cheng Lu, Siyu Li, Lin Bai, Shuheng Du // International Journal of Hydrogen Energy. – 2020. – V. 45. – Iss. 55. – P. 30244–30253.
 19. Pattern visualization and understanding of machine learning models for permeability prediction in tight sandstone reservoirs / Guoyin Zhang, Zhizhang Wang, Shahab Mohaghegh, Chengyan Lin, Yanan Sun, Shengjie Pei // Journal of Petroleum Science and Engineering. – 2020. – 108142.
 20. Teixeira A.F., Secchi A.R. Machine learning models to support reservoir production optimization // IFAC-PapersOnLine. – 2019. – V. 52. – Iss. 1. – P. 498–501.
 21. Predicting field production rates for waterflooding using a machine learning-based proxy model / Zhi Zhong, A.Y. Sun, Yanyong Wang, Bo Ren // Journal of Petroleum Science and Engineering. – 2020. – V. 194. – 107574.
 22. Khomechi E., Bemani A. Prediction of pressure in different two-phase flow conditions: Machine learning applications // Measurement. – 2020. – 108665.
 23. Галкин В.И., Пономарева И.Н., Мартюшев Д.А. Прогноз пластового давления и исследование его поведения при разработке нефтяных месторождений на основе построения многоуровневых многомерных вероятностно-статистических моделей // Георесурсы. – 2021. – Т. 23. – № 3. – С. 73–82.
 24. Козырев Н.Д. Оценка влияния параметров неопределенности на прогнозирование показателей разработки / Н.Д. Козырев, А.Ю. Вишняков, И.С. Путилов // Недропользование. – 2020. – Т. 20. – № 4. – С. 356–368.
 25. Mohammad Ali Ahmadi, Zhangxing Chen. Comparison of machine learning methods for estimating permeability and porosity of oil reservoirs via petro-physical logs // Petroleum. – 2019. – V. 5. – Iss. 3. – P. 271-284.
 26. Pan Y., Deng L., Lee W.J. A novel data-driven pressure/rate deconvolution algorithm to enhance production data analysis in unconventional reservoirs // Journal of Petroleum Science and Engineering. – 2020. – V. 192. – 107332.
 27. Anifowose F., Labadina J., Abdurraheem A. Improving the prediction of petroleum reservoir characterization with a stacked generalization ensemble model of support vector machines // Applied Soft Computing. – 2015. – V. 26. – P. 483–496.
 28. Al Khalifah H., Glover P.W.J., Lorinczi P. Permeability prediction and diagenesis in tight carbonates using machine learning techniques // Marine and Petroleum Geology. – 2020. – V. 112. – 104096.
 29. El-Amin M.F., Subasi A. Developing a generalized scaling-law for oil recovery using machine learning techniques // Procedia Computer Science. – 2019. – V. 163. – P. 237–247.
 30. A survey of methods for explaining black box models / R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, D. Pedreschi // ACM Computing Surveys. – 2018. – V. 51. – Iss. 5. – P. 1–42.

Поступила 22.09.2021 г.

Информация об авторах

Мартюшев Д.А., кандидат технических наук, доцент кафедры нефтегазовых технологий Пермского национального исследовательского политехнического университета.

Пономарева И.Н., доктор технических наук, доцент, профессор кафедры нефтегазовых технологий Пермского национального исследовательского политехнического университета.

Захаров Л.А., инженер отдела создания и мониторинга гидродинамических моделей филиала ООО «ЛУКОЙЛ-Инжиниринг» «ПермНИПИнефть» в г. Пермь.

Шадров Т.А., старший аналитик компании ООО «Диджитал фьюче системс».

UDC 622.276

APPLICATION OF MACHINE LEARNING FOR FORECASTING FORMATION PRESSURE IN OIL FIELD DEVELOPMENT

Dmitriy A. Martyushev¹,
martyushevdi@inbox.ru

Inna N. Ponomareva¹,
permpolitech@gmail.com

Lev A. Zakharov²,
lzakharov-ng@mail.ru

Timur A. Shadrov³,
shadrov.timur@yandex.ru

¹ Perm National Research Polytechnic University,
29, Komsomolskiy avenue, Perm, 614990, Russia.

² Branch of LLC «LUKOIL-Engineering» «PermNIPIneft» in Perm,
29, Sovetskoy Armii street, Perm, 614000, Russia.

³ Digital Futures Systems LLC,
111, Kosmonavtov street, Perm, 614066, Russia.

The relevance of the study is caused by the fact that the advent of artificial intelligence in the oil industry has led to an increase in its use in exploration, development, production, field design and management planning to speed up decision-making, reduce costs and time. Machine learning has gained immense popularity in correlating complex nonlinear datasets and has demonstrated its superiority over regression methods in petroleum engineering in terms of large data prediction errors, processing power and memory. This article discusses the use of machine learning to assess its effectiveness and potential for determining and predicting reservoir pressure values in oil field development, compared with conventional statistical models of oil and gas engineering.

The main aim of the study is to assess the possibilities of calculating and predicting reservoir pressure using the «random forest» machine learning method.

Object: dynamics of reservoir pressure during the development of terrigenous deposits of oil fields in the Perm Krai.

Methods: methods of probabilistic-statistical analysis and machine learning «random forest regression».

The results. The presented study proposes a new method for predicting reservoir pressure using machine learning, based on a nonparametric multidimensional model that links well performance over time. The proposed method takes into account the dynamics of indicators characterizing the operation of wells, and the predicted reservoir pressure is well correlated with the values measured using hydrodynamic studies. It was found that the «random forest» machine learning method provides better performance in terms of reservoir pressure prediction accuracy than the linear regression method. The prospects for further development are additional «training» of the «random forest» model and assessment of the possibility of using other machine learning methods to solve the problem, including expanding the set of factors for more accurate modeling of reservoir pressure.

Key words:

Artificial Intelligence, machine learning, multidimensional model, random forest, well operation indicators, current reservoir pressure, correlation.

REFERENCES

1. Molchanov A.A., Ageev P.G. Implementation of new technology is a reliable method of extracting reserves remaining in hydrocarbon deposits. *Journal of Mining Institute*, 2017, vol. 227, pp. 530–539. In Rus.
2. Shamsiev M.N. The results interpretation of thermogasdynamic studies of vertical gas wells incomplete in terms of the reservoir penetration degree. *Georesursy*, 2020, vol. 20, no. 1, pp. 39–43. In Rus.
3. Elesin A.V., Kadyrova A.S., Nikiforov A.I. Definition of the reservoir permeability field according to pressure measurements on wells with the use of spline function. *Georesursy*, 2018, vol. 20, no. 2, pp. 102–107. In Rus.
4. Bobb I.F. International experience of E&P software solutions development. *Georesursy*, 2018, vol. 20, no. 3, pp. 193–196. In Rus.
5. Belozerov I.P., Gubaidullin M.G. Concept of technology for determining the permeability and porosity properties of terrigenous reservoirs on a digital rock sample model. *Journal of Mining Institute*, 2020, vol. 244, pp. 402–407. In Rus.
6. Jamiu Oyekan Adegbite, Hadi Belhaj, Achinta Bera. Investigations on the relationship among the porosity, permeability and pore throat size of transition zone samples in carbonate reservoirs using multiple regression analysis, artificial neural network and adaptive neuro-fuzzy interface system. *Petroleum Research*, 2021. Available at: <https://doi.org/10.1016/j.ptlrs.2021.05.005> (accessed 15 August 2021).
7. Ponomareva I.N., Galkin V.I., Martyushev D.A. Operational method for determining bottom hole pressure in mechanized oil producing wells, based on the application of multivariate regression analysis. *Petroleum Research*, 2021. Available at: <https://doi.org/10.1016/j.ptlrs.2021.05.010> (accessed 15 August 2021).
8. Belhouchet H.E., Benzagouta M.S., Dobbi A., Alquraishic A., Duplay J. A new empirical model for enhancing well log permeability prediction, using nonlinear regression method: Case study from Hassi-Berkine oil field reservoir – Algeria. *Journal of King Saud University - Engineering Sciences*, 2021, vol. 33, Iss. 2, pp. 136–145.

9. Wong P.M., Jian F.X., Taggart I.J. A critical comparison of neural networks and discriminant analysis in lithofacies, porosity and permeability predictions. *Journal Petroleum Geology*, 1995, vol. 18, pp. 191–206.
10. Nikravesh M., Aminzadeh F. Ch. 1. Soft computing for intelligent reservoir characterization and modeling. *Developments in Petroleum Science*, 2003, vol. 51, pp. 3–32.
11. Chernikov A.D., Eremin N.A., Stolyarov V.E., Sboev A.G., Semenova-Chaschina O.K., Fitsner L.K. Application of artificial intelligence methods for identifying and predicting complications in the construction of oil and gas wells: problems and solutions. *Georesursy*, 2020, vol. 22, no. 3, pp. 87–96. In Rus.
12. Lalomov A.V., Boneva A.A., Chefranov R.M. The development of numerical forecasting systems of primary sources of gold on the results of placer sampling in the example Vagran placer cluster (North Urals). *Georesursy*, 2020, vol. 22, no. 2, pp. 67–76. In Rus.
13. Evsyutkin I.V., Markov N.G. Deep artificial neural networks for forecasting debit values for production wells. *Bulletin of the Tomsk Polytechnic University. Geo Assets Engineering*, 2020, vol. 331, no. 11, pp. 88–95. In Rus.
14. Popov N.A., Putilov I.S., Gulyaeva A.A., Vinokurova E.V. Application of deep learning technologies for studying thin sections on the example of Usinsk oil field. *Bulletin of the Tomsk Polytechnic University. Geo Assets Engineering*, 2020, vol. 331, no. 6, pp. 100–112. In Rus.
15. Asante Otchere D., Arbi Ganata T.O., Gholami R., Ridhaa S. Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models. *Journal of Petroleum Science and Engineering*, 2020, 108182.
16. Harpreet Singh. Machine learning for surveillance of fluid leakage from reservoir using only injection rates and bottomhole pressures. *Journal of Natural Gas Science and Engineering*, 2019, vol. 69, 102933.
17. Shahab D. Mohaghegh. Subsurface analytics: contribution of artificial intelligence and machine learning to reservoir engineering, reservoir modeling, and reservoir management. *Petroleum Exploration and Development*, 2020, vol. 47, Iss. 2, pp. 225–228.
18. Guoyi Yu, Feng Xu, Yingzhi Cui, Xiangling Li, Chujuan Kang, Cheng Lu, Siyu Li, Lin Bai, Shuheng Du. A new method of predicting the saturation pressure of oil reservoir and its application. *International Journal of Hydrogen Energy*, 2020, vol. 45, Iss. 55, pp. 30244–30253.
19. Guoyin Zhang, Zhizhang Wang, Shahab Mohaghegh, Chengyan Lin, Yanan Sun, Shengjie Pei. Pattern visualization and understanding of machine learning models for permeability prediction in tight sandstone reservoirs. *Journal of Petroleum Science and Engineering*, 2020, 108142.
20. Teixeira A.F., Secchi A.R. Machine learning models to support reservoir production optimization, *IFAC-PapersOnLine*, 2019, vol. 52, Iss. 1, pp. 498–501.
21. Zhong Zh., Sun A.Y., Wang Y., Ren B. Predicting field production rates for waterflooding using a machine learning-based proxy model. *Journal of Petroleum Science and Engineering*, 2020, vol. 194, 107574.
22. Khomehchi E., Bemani A. Prediction of pressure in different two-phase flow conditions: machine learning applications. *Measurement*, 2020, 108665.
23. Galkin V.I., Ponomareva I.N., Martyushev D.A. Prediction of reservoir pressure and study of its behavior in the development of oil fields based on the construction of multilevel multidimensional probabilistic-statistical models. *Georesursy*, 2020, vol. 23, no. 3, pp. 73–82. In Rus.
24. Kozlyev N.D., Vishnyakov A.Yu., Putilov I.S. Assessment of the uncertainty parameters influence on the development indicators forecasting. *Perm Journal of Petroleum and Mining Engineering*, 2020, vol. 20, no. 4, pp. 356–368. In Rus.
25. Ali Ahmadi M., Chen Zh. Comparison of machine learning methods for estimating permeability and porosity of oil reservoirs via petrophysical logs. *Petroleum*, 2019, vol. 5, Iss. 3, pp. 271–284.
26. Pan Y., Deng L., Lee W.J. A novel data-driven pressure/rate deconvolution algorithm to enhance production data analysis in unconventional reservoirs. *Journal of Petroleum Science and Engineering*, 2020, vol. 192, 107332.
27. Anifowosea F., Labadina J., Abdulraheem A. Improving the prediction of petroleum reservoir characterization with a stacked generalization ensemble model of support vector machines. *Applied Soft Computing*, 2015, vol. 26, pp. 483–496.
28. Al Khalifah H., Glover P.W.J., Lorinczi P. Permeability prediction and diagenesis in tight carbonates using machine learning techniques. *Marine and Petroleum Geology*, 2020, vol. 112, 104096.
29. El-Amin M.F., Subasi A. Developing a generalized scaling-law for oil recovery using machine learning techniques. *Procedia Computer Science*, 2019, vol. 163, pp. 237–247.
30. Guidotti R., Monreale A., Ruggieri S., Turini F., Giannotti F., Pedreschi D. A survey of methods for explaining black box models. *ACM Computing Surveys*, 2018, vol. 51, Iss. 5, pp. 1–42.

Received: 22 September 2021.

Information about the authors

Dmitriy A. Martyushev, Cand. Sc., assistant professor, Perm National Research Polytechnic University.

Inna N. Ponomareva, Cand. Sc., professor, Perm National Research Polytechnic University.

Lev A. Zakharov, engineer, branch of LLC «LUKOIL-Engineering» «PermNIPIneft» in Perm.

Timur A. Shadrov, senior analyst, Digital Futures Systems LLC.