

# Об увеличении продуктивного времени бурения нефтегазовых скважин с использованием методов машинного обучения

А.Н. Дмитриевский<sup>1</sup>, А.Г. Сбоев<sup>2</sup>, Н.А. Еремин<sup>1,3\*</sup>, А.Д. Черников<sup>1</sup>, А.В. Наумов<sup>1</sup>, А.В. Грязнов<sup>1</sup>,  
И.А. Молошников<sup>1</sup>, С.О. Бороздин<sup>3</sup>, Е.А. Сафарова<sup>1</sup>

<sup>1</sup>Институт проблем нефти и газа РАН, Москва, Россия

<sup>2</sup>Национальный исследовательский центр «Курчатовский институт», Москва, Россия

<sup>3</sup>Российский государственный университет нефти и газа (национальный исследовательский университет) им. И.М. Губкина, Москва, Россия

Статья посвящена разработке гибридного метода прогнозирования и предупреждения развития осложнений в процессе бурения скважин на базе методов машинного обучения и современных нейросетевых моделей. Осложнения в процессе бурения, такие как поглощения, газонефтеводопроявления и прихваты, приводят к росту непроизводительного времени, т.е. времени которое не является технически необходимым для строительства скважины и вызывается различными нарушениями производственного процесса. Рассмотрено несколько различных подходов, в т.ч. на основе регрессионной модели прогнозирования функции индикатора, которая отражает приближение к развивающемуся осложнению, а также моделей выделения аномалий, построенных как на базовых алгоритмах машинного обучения, так и с применением нейросетевой модели глубокого обучения. Показаны визуализированные примеры работы разработанных методов на симуляционных и реальных данных. Интеллектуальный анализ большого объема информации со станций геолого-технологических измерений основан на хорошо зарекомендовавших себя алгоритмах машинного обучения. На основе этих данных предложена нейросетевая модель для предотвращения осложнений и аварийных ситуаций в процессе строительства скважин. Применение данного метода позволит минимизировать непроизводительное время бурения.

**Ключевые слова:** машинное обучение, гибридное моделирование, нейронные сети, выявление аномалий, прогнозирование осложнений, бурение скважин, геолого-технологическая информация, предотвращение аварий и осложнений, искусственный интеллект, автоматизированная система, строительство скважин, нейросетевое моделирование

**Для цитирования:** Дмитриевский А.Н., Сбоев А.Г., Еремин Н.А., Черников А.Д., Наумов А.В., Грязнов А.В., Молошников И.А., Бороздин С.О., Сафарова Е.А. (2020). Об увеличении продуктивного времени бурения нефтегазовых скважин с использованием методов машинного обучения. *Георесурсы*, 22(4), с. 79–85. DOI: <https://doi.org/10.18599/grs.2020.4.79-85>

## 1. Введение

Современная мировая наука характеризуется значительным прогрессом в развитии современных методов анализа данных (Data Driven methods) и математических моделей, в том числе на основе технологий машинного обучения и нейронных сетей. Благодаря этим технологиям стали появляться современные (state-of-the-art) алгоритмы, позволяющие эффективно помогать в решении сложных задач в нефтегазовой области (Казначеев и др., 2016; Архипов и др., 2020; Borozdin et al., 2020; Дмитриевский и др., 2020).

Бурение нефтяных и газовых скважин является неотъемлемо важным аспектом нефтегазового производства. Повышение безопасности в ходе выполнения этого сложного технологического процесса, является актуальной и важной задачей. Одним из вариантов решения данной задачи является предотвращение осложнений и аварийных ситуаций путем своевременного предупреждения буровой бригады о начале их развития.

При наличии нескольких источников больших геоданных при бурении (система геонавигации, станция геолого-технологических измерений (ГТИ), буровой трена-

жер-симулятор) эффективно использование нового вида моделирования – гибридного. Гибридная модель – это комплекс моделей, состоящий из базовой 4D модели ствола строящейся скважины, вероятностной (или нечеткой) модели неопределенности и модели машинного обучения. Гибридная модель непрерывно уточняется в процессе бурения по мере поступления разнородных больших объемов геолого-технологических данных и используется при автоматизированном прогнозировании осложнений и аварийных ситуаций. В частности, в исследовании (Дьяконов, Головина, 2017) рассмотрена задача автоматического выявления поломок механизмов и определения их типов на основе собранных исторических данных, которая была сведена к классической задаче машинного обучения – обнаружению аномалий. В работе приведен обширный обзор методов решения данной задачи и результаты их апробации на реальных данных, где наилучший результат показал подход с «обучением без учителя», а именно модель «изолирующего леса» (Isolation Forest) (Liu et al., 2008). В работе (Gurina et al., 2020) задача обнаружения осложнений и определения их типов в процессе бурения решалась путем построения модели машинного обучения для выявления в данных аномалий. В отличие от предыдущей работы, в этом исследовании применялся подход «обучения с учителем». Данные каротажа, полученные в

\* Ответственный автор: Николай Александрович Еремин  
E-mail: [ermn@mail.ru](mailto:ermn@mail.ru)

© 2020 Коллектив авторов

реальном времени, сравнивались с аналогичными, ранее собранными в базе данных, в которых присутствовали различные типы осложнений. Результаты поиска ранжировались, и выбиралось наиболее подходящее осложнение. Для такого сравнения, ранжирования и определения осложнения была обучена классификационная модель градиентного бустинга (Chen, Guestrin, 2016), которая позволила добиться точности определения осложнений в 0.908 по метрике ROC AUC, т.е. вычисление площади под кривой эксплуатационных характеристик (ROC AUC – одна из популярных метрик, используемых в отрасли, где AUC (Area Under The Curve) – площадь под кривой, а кривая представлена изогнутой траекторией ROC (Receiver Operating Characteristic) эксплуатационных характеристик приемника). На практике в зависимости от значения AUC эффективность модели классифицируется следующим образом:  $0,8 \leq AUC \leq 1,0$  – модель работает превосходно;  $0,6 \leq AUC < 0,8$  – модель работает хорошо;  $0,5 < AUC < 0,6$  – модель работает удовлетворительно и  $AUC \leq 0,5$  – модель не работает. В исследовании (Кодиров, Шестаков, 2019) разработан метод определения прихвата колонны бурильных труб на основе нейронной сети. Авторами построена многослойная полносвязная нейронная сеть (MLP, multilayer perceptron), которая определяла возникновение прихвата и его тип с точностью 93% по базовой метрике Ассигасу (измеряет количество верно классифицированных объектов относительно общего количества всех объектов).

В рассмотренных выше работах описаны методики определения уже произошедших событий, которые не позволяют своевременно реагировать на стремительно развивающиеся предаварийные ситуации в бурении. Они часто успешно используются для предварительной разметки большого объема «сырой» информации со станций ГТИ.

Более сложной задачей является не определение типа осложнений, а прогнозирование вероятности их происхождения при бурении в будущем. В работе (Пичугин и др., 2013) показано, что с помощью обучения модели деревьев решений (Decision Tree) на различной геоинформации, полученной с пробуренных ранее скважин, можно оценить риск бурения и возникновения нежелательных ситуаций, а также повысить успешность ввода новых добывающих скважин в эксплуатацию на 15–25%. В другом исследовании (Линд и др., 2013), авторами решалась задача прогнозирования количества поглощений бурового раствора в процессе бурения новой скважины. Для этого была построена самообучающаяся нейросетевая модель, называемая карта Кохонена (Kohonen, 1990), которая обучалась на информации, собранной по ранее пробуренным скважинам. Полученная модель, по оценке авторов, позволит снизить стоимость бурения до 4%. Задачи прогнозирования значений различных параметров бурения в реальном времени является еще более сложными (Еремин, Столяров, 2020; Noshi, Schubert, 2018) и на данном этапе развития методов машинного обучения проработаны слабо. Самые успешные из этих методов построены на основе глубоких нейронных сетей с рекуррентными и конволюционными слоями (Kanfar et al., 2020; Li et al., 2019).

Рассмотренные выше модели прогнозирования позволяют оценить риски будущего бурения и заранее подготовиться к возможным осложнениям в процессе

бурения, но они не позволяют спрогнозировать возможные осложнения так, чтобы буровая бригада смогла в режиме реального времени совершить своевременные действия по их полному предотвращению или же минимизации возможных последствий. Для осуществления таких прогнозов необходимо использовать данные ГТИ, получаемые в реальном времени (Еремин и др., 2020а-2020е). В данной работе рассмотрены различные подходы для прогнозирования наступления предаварийных ситуаций трёх типов в процессе бурения, с использованием размеченных и неразмеченных наборов данных:

1. Газоводонефтепроявление (ГНВП);
2. Прихват буровой колонны;
3. Поглощение бурового раствора.

## 2. Данные

Данные, используемые в процессе разработки методов и для проведения запланированных экспериментов, предоставлены партнерами из РГУ нефти и газа (НИУ) им. И.М. Губкина в виде симуляционных данных, полученных с бурового тренажёра, а также данных со станций ГТИ при разбуривании скважин на реальных месторождениях. Оба набора данных состоят из показаний, снимаемых различными датчиками в процессе бурения скважины (или его симуляции), установленными на оборудовании. Количество отслеживаемых параметров, их полнота и частота записи в вышеуказанных наборах данных отличаются, что является дополнительной сложностью для анализа.

### Симуляционные данные

Симуляционный набор данных получен по итогам проведения экспериментов по моделированию процесса бурения на тренажёре DrillSim-5000. Получено 79 записей симуляций, среди которых 33 относятся к бурению с осложнением ГНВП, 27 со шламонакоплением (прихватом), 9 с поглощением бурового раствора и 10 симуляций безаварийного бурения. Пример данных одной симуляции изображен на рис. 1.

Данные записей симуляций с тренажёра представлены в виде таблиц с 16 параметрами (в т.ч. вес на крюке, скорость проходки, скорость вращения долота и др.). Для каждой симуляционной записи проставлена метка времени, которая означала начало развития каждого из моделируемых осложнений. Метка времени проставлялась вручную экспертами в области бурения.

### Данные с реального месторождения

В качестве данных с разбуриваемых месторождений представлены реальновременные записи отслеживаемых параметров в процессе бурения с 25 различных скважин, где для 23 скважин указаны записи штатного бурения; одна скважина содержит в себе записи с осложнением вида “Прихват”, вторая с осложнением вида “Поглощение”. Примеры некоторых отслеживаемых параметров с их диапазонами значений отражены в таблице 1.

Разметка данных проведена специалистами по бурению и включала указание времени начала осложнения для каждого из 2 типов осложнений (прихват и поглощение).

### Подготовка данных

Для того, чтобы данные можно было использовать в процессе обучения моделей искусственного интеллекта

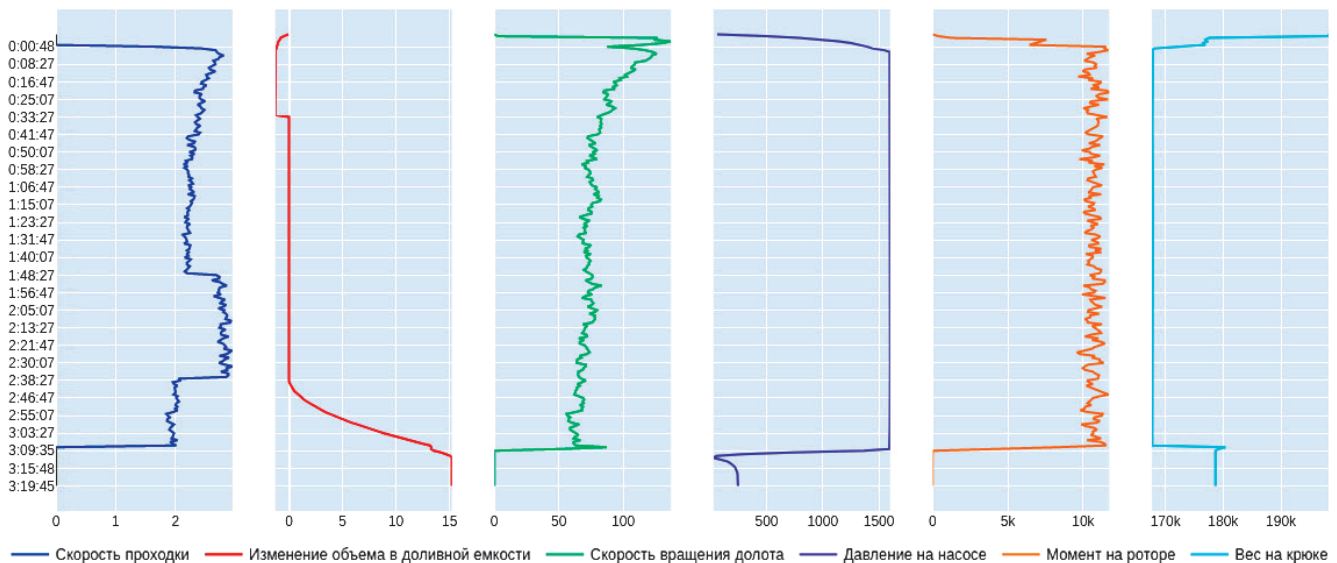


Рис. 1. Пример симуляционных данных с некоторыми параметрами, с осложнением ГНВП

и анализа качества входной информации, необходимо произвести их предварительную подготовку. Для анализа были отобраны отрезки непрерывного бурения (наблюдения с ненулевой скоростью проходки), в которых отбрасывались все значения параметров, соответствующих моментам времени после наступления начала осложнения заданного типа, так как такие наблюдения не представляют интереса для процесса прогнозирования его наступления. В связи с тем, что частота снятия параметров в симуляционных данных от эксперимента к эксперименту варьировалась, для их выравнивания использовалась линейная интерполяция по двум соседним точкам. Итоговый шаг по времени между точками – 2 секунды. Из всех параметров были отобраны основные, которые представляют наибольший интерес для определения рассматриваемых осложнений, а также присутствуют в обоих рассматриваемых наборах данных. В таблице 2 представлен финальный список используемых исходных параметров, которые применялись в обучении моделей машинного обучения и нейронной сети.

Для расширения пространства признаков и их нормализации был использован ряд дополнительных производных параметров, получаемых из показаний датчиков, отобранных после процедуры предварительной подготовки данных, в т.ч. разница между текущим и предыдущими значениями, расчет скользящего тренда и разложение значений параметров на угол наклона тренда и отклонение от него, расчет процентилей значений параметров и их нормирование в пределах процентилей. Данные были разделены на тренировочные и тестовые наборы данных. Так как для реальных данных отсутствовали размеченные примеры, кроме двух тестовых, то в тренировочный

Параметр	Единицы измерения	Тип данных
Нагрузка на крюк	фунты	числовой
Скорость механической проходки	фут/ч	числовой
Скорость вращения ротора	об/мин	числовой
Разница между значением потока БР на выходе и входе	баррель нефтиной США	числовой
Давление насоса	фунт/дюйм <sup>2</sup>	числовой

Табл. 2. Список используемых в моделях параметров

набор вошли все скважины, кроме двух, которые стали контрольными и также были добавлены в тестовый набор. В случае с симуляционными данными, для каждого из рассматриваемых осложнений был выделен свой тестовый набор скважин, включающий 20% от всех скважинных записей, содержащих рассматриваемое осложнение, а также 20% случайно выбранных безаварийных скважин.

### 3. Методы и подходы

Для решения задачи прогнозирования наступления различных предаварийных ситуаций (осложнений), в частности, рассматриваемых в этой работе прихватов, поглощений и ГНВП, реализованы следующие подходы на основе:

1. Выделения аномалий, с построением:
  - а. Одноклассовой модели машинного обучения;
  - б. Регрессионной нейросетевой модели.

2. Построения регрессионной функции индикатора, отражающей приближение к вероятному осложнению.

Подходы 1а и 2 апробированы на симуляционных данных, полученных с бурового тренажера, так как в них содержался набор размеченных примеров по различным

Наименование параметра	Минимальное значение	Среднее значение	Максимальное значение	Количество уникальных значений
Вес на крюке (10 кН)	0.0	3.84	46.82	4513
Давление (кПа)	0.0	1840.83	23781.44	87313
Тальблок (м)	0.0	10.83	25.27	2520
Забой (м)	0.0	625.52	849.01	20745
Нагрузка на долото (10 кН)	0.0	1.17	19.98	1865

Табл. 1. Некоторые параметры и их диапазон значений скважины №1

осложнениям. Подход 1б применялся к реальным данным (2 тестовых примера по осложнениям типа «Прихват» и «Поглощение»); размеченные тренировочные примеры отсутствовали.

### Модель выделения аномалий

Подход основан на задаче выделения аномальных ситуаций на множестве регистрируемых параметров. Основная идея заключалась в следующем: чем ближе к осложнению рассматриваются значения наблюдаемых параметров бурения, тем больше они отличны от тех, которые характерны для безаварийного штатного бурения в тех же условиях. Такой подход позволяет использовать большой объем неразмеченных данных, выделяя аномальные отклонения различных параметров бурения, а также необычную совокупность их значений. Для построения такой модели использовался метод изолирующего леса (Isolation Forest) с параметром  $n\_estimators=500$ . Данный метод взят из библиотеки с открытым исходным кодом `sklearn` и заключается в построении случайного бинарного решающего дерева, которое может распознавать аномалии различных видов: как изолированные точки с низкой локальной плотностью, так и кластеры аномалий малых размеров.

Результаты работы обученной модели на тестовом примере можно увидеть на рис. 2. Конец примера (правая граница графика) означает начало развития осложнения типа ГНВП. Из графика видно, что метод выделил аномальное поведение параметров ближе к концу примера. Недостатком такого подхода является то, что метод определяет не только приближение к осложнению, но и другие возможные отклонения параметров, такие как сбой датчиков, другие возможные осложнения, аномальное управление и т.п.

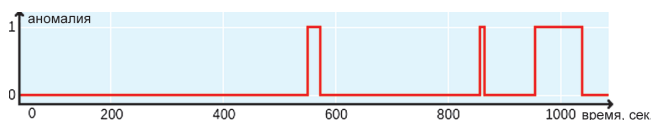


Рис. 2. Выделенные аномалии для тестового примера с осложнением типа «ГНВП»

### Модель прогнозирования функции-индикатора

Модель основана на введении функции индикатора, имеющей нулевое значение на временном интервале, достаточно удаленном от предаварийной ситуации, и возрастающая по мере её приближения (рис. 3). При обучении модели, функция индикатор задавалась в виде сигмоидальной функции, принимающей значение 0,5 за 7 минут до аварии и близкое к 1 в точке начала аварии. Сигнализировать о приближающемся осложнении модель начинает при превышении заданного порога, рассчитанного по тренировочной выборке примеров. В среднем, модель начинает превышать порог за 4–5 минут до момента возникновения события в виде осложнения.

В качестве основного алгоритма использовался метод на основе случайного леса (Random forest) со 100 деревьями. Случайный лес – это ансамблевый метод регрессии, задействующий ряд регрессионных деревьев для различных случайно выбранных подвыборок набора данных (параметров) и использующий усреднение для повышения точности прогнозирования и контроля переобучения.

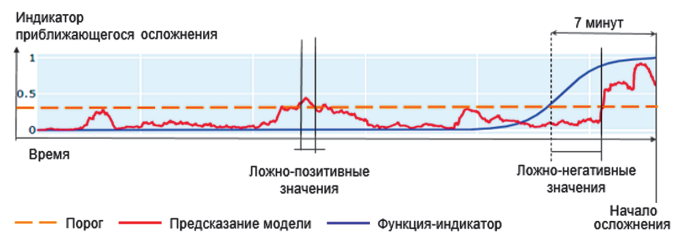


Рис. 3. Регрессионное прогнозирование функции индикатора

### Нейросетевая модель для реальных данных

Для проведения экспериментов с данными реального бурения не представляется возможным использовать подходы, основанные на моделях «обучения с учителем». Такие подходы нуждаются в достаточном наборе размеченных параметров, которые не представлены на рассматриваемом множестве реальных данных. Поэтому, было принято решение использовать метод, схожий с тем, что описан в подходе «выделение аномалий» и решать аналогичную задачу, с оценкой работы созданного метода на имеющихся размеченных примерах осложнений вида «Прихват» и «Поглощение». Так как неразмеченных данных в указанном источнике присутствует достаточно много, был опробован авторегрессионный подход с использованием сверточных, рекуррентных и полносвязных слоев нейронной сети. Для обработки реальных данных и подачи их в нейросетевую модель используется подход скользящего окна с перекрытием (Дмитриевский и др., 2020). Модель использует 5 отобранных параметров (табл. 2), для которых формируется окно в 1024 последовательных значения. Получившаяся матрица подается на вход модели, которая обучается прогнозировать пять параметров на следующем шаге. Для поиска аномальной динамики параметров оценивалась разница между предсказанными значениями параметров и наблюдаемыми. После обучения на 20 скважинах безаварийного бурения, модель была запущена на трёх безаварийных скважинах и двух примерах, в одном из которых присутствовала «потеря бурового раствора», а в другом – «прихват». Результаты показали, что при отсутствии осложнений, суммарная ошибка предсказанных значений не превышает 500 (например, рис. 4). Однако при приближении к осложнению типа «Прихват» и началу «Потерь бурового раствора» эта ошибка начинала расти. На рис. 5–6 последняя точка (правая граница графика) означает начало развития предаварийной ситуации. Видно, что в примере с осложнением типа «Прихват» (рис. 5), более чем за 5000 секунд (полтора часа), ошибка начинает возрастать и не опускается до самого начала осложнения типа «Прихват». В примере с осложнением «Потеря бурового раствора» (рис. 6), более чем за 6 часов начинаются сильные колебания ошибки в предсказаниях, при которых она превышает пороговое значение.

## 4. Заключение

В автоматизированной системе предупреждения основных видов осложнений в процессе бурения разработан и апробирован ряд современных гибридных методов интеллектуального анализа больших геоданных (Data Driven methods), в том числе на основе технологий машинного обучения и нейронных сетей Data Driven моделей, продемонстрировавших свою эффективность на

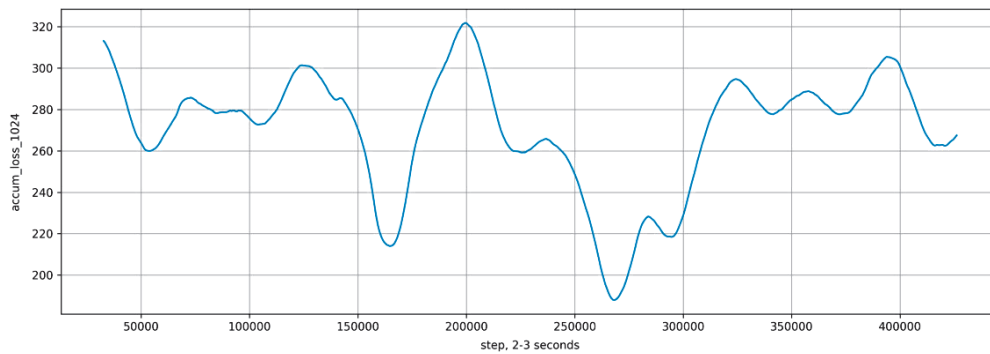


Рис. 4. Ошибка прогнозирования параметров в скважине без осложнений, не участвовавшей в процессе обучения модели

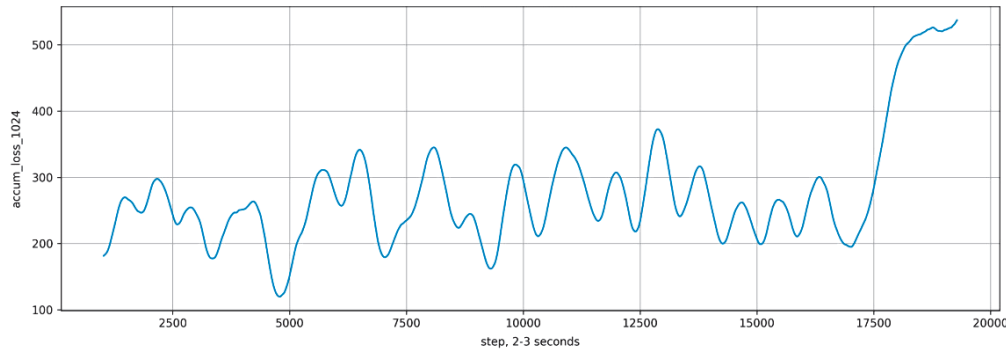


Рис. 5. Скважина с осложнением вида «Прихват»

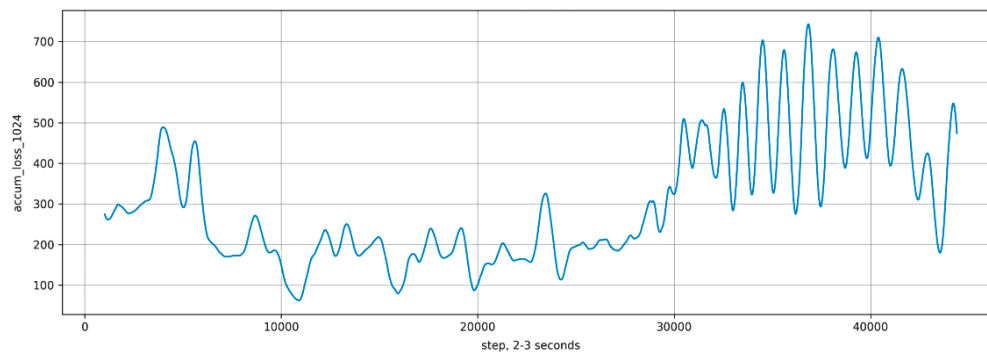


Рис. 6. Скважина с осложнением вида «Поглощение»

симуляционных данных с бурового тренажера и реальных данных бурения скважин (Еремин и др., 2020а-2020е). Дальнейшая работа будет направлена на расширение наборов больших геоданных систем геонавигации и станций геолого-технологических измерений, оценку на них точности предлагаемых гибридных моделей и их уточнение.

### Финансирование

Статья подготовлена по результатам работ, выполненных в рамках Программы государственных академий наук на 2013–2020 годы. Раздел 9 «Науки о Земле»; направление фундаментальных исследований: 132 «Комплексное освоение и сохранение недр Земли, инновационные процессы разработки месторождений полезных ископаемых и глубокой переработки минерального сырья», по теме государственного задания «Фундаментальный базис инновационных технологий нефтяной и газовой промышленности», №АААА-А 16-116031750016-3; № 0139-2019-0009 в Парусе и № АААА-А19-119013190038-2 в РОСРИДе.

### Литература

Архипов А.И., Дмитриевский А.Н., Еремин Н.А. и др. (2020). Анализ качества данных станции геолого-технологических исследований при распознавании поглощений и газонефтеводопроявлений для повышения точности прогнозирования нейросетевых алгоритмов.

*Нефтяное хозяйство*, 8, с. 63–67. DOI: 10.24887/0028-2448-2020-8-63-67

Дмитриевский А.Н., Еремин Н.А., Филиппова Д.С. и др. (2020). Качественный анализ геоданных временного ряда для предупреждения осложнений и аварийных ситуаций при бурении нефтяных и газовых скважин. *Socar Proceedings*, 3, с. 31–37. <http://dx.doi.org/10.5510/OGP20200300442>

Дмитриевский А.Н., Еремин Н.А., Филиппова Д.С., Сафарова Е.А. (2020). Цифровой нефтегазовый комплекс России. *Георесурсы, Спецвыпуск*, с. 32–35. <https://doi.org/10.18599/grs.2020.SI.32-35>

Дьяконов А.Г., Головина А.М. (2017). Выявление аномалий в работе механизмов методами машинного обучения. *Тр. XIX Межд. конф.: Аналитика и управление данными в областях с интенсивным использованием данных*. М., с. 469–476.

Еремин Н.А., Столяров В.Е. (2020). О цифровизации процессов газодобычи на поздних стадиях разработки месторождений. *Socar Proceedings*, 1, с. 59–69. DOI: 10.5510/ogp20200100424

Еремин Н.А., Водопьян А.О., Дуплякин В.О., Черников А.Д., Космос С.А. (2020а). Программный компонент «Нефтегазовый Блокчейн». Свид. о рег. программы для ЭВМ RU 2020614626. Заявка № 2020613699 от 27.03.2020.

Еремин Н.А., Дмитриевский А.Н., Чашина-Семенова О.К., Фицнер Л.К., Черников А.Д. (2020б). Программный компонент «Адаптация обобщенных нейросетевых моделей прогнозирования осложнений и аварийных ситуаций к геофизическим параметрам при бурении конкретной скважины». Свид. о рег. программы для ЭВМ RU 2020660890. Заявка № 2020660179 от 08.09.2020.

Еремин Н.А., Дмитриевский А.Н., Чашина-Семенова О.К., Фицнер Л.К., Черников А.Д. (2020с). Программный компонент «Оркестровка -

интеграция модулей системы прогнозирования осложнений и аварийных ситуаций при бурении и строительстве скважин». Свид. о рег. программы для ЭВМ RU 2020660891. Заявка № 2020660181 от 08.09.2020.

Еремин Н.А., Дмитриевский А.Н., Чашина-Семенова О. К., Фицнер Л.К., Черников А.Д. (2020d). Программный компонент «Нейросетевые расчеты - построение моделей прогноза осложнений и аварийных ситуаций при бурении и строительстве скважин» (ПКНР). Свид. о рег. программы для ЭВМ RU 2020660892. Заявка № 2020660182 от 08.09.2020.

Еремин Н.А., Дмитриевский А.Н., Чашина-Семенова О.К., Фицнер Л.К., Черников А.Д. (2020e). Программный компонент «Индикация прогноза осложнений и аварийных ситуаций при бурении и строительстве скважин» (ПК «Индикация»). Свид. о рег. программы для ЭВМ RU 2020661356. Заявка № 2020660450 от 14.09.2020.

Казначеев П.Ф., Самойлова Р.В., Курчиски Н.В. (2016). Применение методов искусственного интеллекта для повышения эффективности в нефтегазовой и других сырьевых отраслях. *Экономическая политика*, 11(5).

Кодиров Ш.Ш., Шестаков А.Л. (2019). Разработка искусственной нейронной сети для прогнозирования прихватов колонн бурильных труб. *Вестник Южно-Уральского государственного университета. Серия: Компьютерные технологии, управление, радиоэлектроника*, 19(3).

Линд Ю.Б., Мулюков Р.А., Кабирова А.Р., Мурзагалин А.Р. (2013). Оперативное прогнозирование осложнений при бурении. *Нефтяное Хозяйство*, 2, с. 55–57.

Пичугин О.Н., Прокофьева Ю.З., Александров Д.М. (2013). Деревья решений как эффективный метод анализа и прогнозирования. *Нефтепромысловое дело*, 11, с. 69–75.

Borozdin, S., Dmitrievsky, A., Eremin, N., Arkhipov, A., Sboev, A., Chashchina-Semenova, O., Fitzner L., Safarova, E. (2020). Drilling Problems Forecast System Based on Neural Network. *Society of Petroleum Engineers*. <https://doi.org/10.2118/202546-MS>

Chen T., Guestrin C. (2016). Xgboost: A scalable tree boosting system. *Proc. 22nd ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*. ACM, pp. 785–794. <https://doi.org/10.1145/2939672.2939785>

Gurina E. et al. (2020). Application of machine learning to accidents detection at directional drilling. *Journal of Petroleum Science and Engineering*, 184, 106519. <https://doi.org/10.1016/j.petrol.2019.106519>

Kanfar R. et al. (2020). Real-Time Well Log Prediction From Drilling Data Using Deep Learning. arXiv preprint arXiv: 2001.10156. <https://doi.org/10.2523/IPTC-19693-MS>

Kohonen T. (1990). The self-organizing map. *Proc. IEEE*, 78(9), pp. 1464–1480. <https://doi.org/10.1109/5.58325>

Li Y. et al. (2019). Deep learning for well data history analysis. *Society of Petroleum Engineers*. <https://doi.org/10.2118/196011-MS>

Liu F.T., Tony T.K.M., Zhou Z.H. (2008). Isolation forest. *Proc. 8th IEEE Int. Conf. on Data Mining*, pp. 413–422. <https://doi.org/10.1109/ICDM.2008.17>

Noshi C.I., & Schubert J.J. (2018). The Role of Machine Learning in Drilling Operations. A Review. *Society of Petroleum Engineers*. <https://doi.org/10.2118/191823-18ERM-MS>

Статья поступила в редакцию 01.09.2020;

Принята к публикации 19.09.2020; Опубликована 11.12.2020

IN ENGLISH

## On increasing the productive time of drilling oil and gas wells using machine learning methods

A.N. Dmitrievsky<sup>1</sup>, A.G. Sboev<sup>2</sup>, N.A. Eremin<sup>1,3\*</sup>, A.D. Chernikov<sup>1</sup>, A.V. Naumov<sup>1</sup>, A.V. Gryaznov<sup>1</sup>, I.A. Moloshnikov<sup>1</sup>, S.O. Borozdin<sup>3</sup>, E.A. Safarova<sup>1</sup>

<sup>1</sup>Oil and Gas Research Institute of the Russian Academy of Sciences, Moscow, Russian Federation

<sup>2</sup>National Research Center «Kurchatov Institute», Moscow, Russian Federation

<sup>3</sup>National University of Oil and Gas «Gubkin University», Moscow, Russian Federation

\*Corresponding author: Nikolai A. Eremin, e-mail: ermn@mail.ru

**Abstract.** The article is devoted to the development of a hybrid method for predicting and preventing the development of troubles in the process of drilling wells based on machine learning methods and modern neural network models. Troubles during the drilling process, such as filtrate leakoff, gas, oil and water shows and sticking, lead to an increase in unproductive time, i.e. time that is not technically necessary for well construction and is caused by various violations of the production process. Several different approaches have been

### Сведения об авторах

Анатолий Николаевич Дмитриевский – академик РАН, доктор тех. наук, профессор, научный руководитель Института проблем нефти и газа РАН

Россия, 119333, Москва, ул. Губкина, д. 3

Александр Георгиевич Сбоев – ведущий научный сотрудник, канд. тех. наук, Национальный исследовательский центр «Курчатовский институт»

Россия, 123098, Москва, пл. Академика Курчатова, д. 1

Николай Александрович Еремин – доктор тех. наук, главный научный сотрудник, заместитель директора по инновационной работе, Институт проблем нефти и газа РАН; профессор кафедры разработки и эксплуатации нефтяных месторождений, Российский государственный университет нефти и газа (национальный исследовательский университет) им. И.М. Губкина

Россия, 119333, Москва, ул. Губкина, д. 3

Александр Дмитриевич Черников – ведущий научный сотрудник, канд. тех. наук, Институт проблем нефти и газа РАН

Россия, 119333, Москва, ул. Губкина, д. 3

Александр Владимирович Наумов – инженер, Институт проблем нефти и газа РАН

Россия, 119333, Москва, ул. Губкина, д. 3

Артём Викторович Грязнов – инженер, Институт проблем нефти и газа РАН

Россия, 119333, Москва, ул. Губкина, д. 3

Иван Александрович Молошников – инженер, Институт проблем нефти и газа РАН

Россия, 119333, Москва, ул. Губкина, д. 3

Сергей Олегович Бороздин – старший преподаватель кафедры бурения нефтяных и газовых скважин, Российский государственный университет нефти и газа (национальный исследовательский университет) им. И.М. Губкина

Россия, 119991, Москва, Ленинский просп., д. 65

Елизавета Александровна Сафарова – младший научный сотрудник, Институт проблем нефти и газа РАН

Россия, 119333, Москва, ул. Губкина, д. 3

considered, including based on the regression model for predicting the indicator function, which reflects an approach to a developing trouble, as well as anomaly extraction models built both on basic machine learning algorithms and using the neural network model of deep learning. Showing visualized examples of the work of the developed methods on simulation and real data. Intelligent analysis of Big Geodata from geological and technological measurement stations is based on well-proven machine learning algorithms. Based on these

data, a neural network model was proposed to prevent troubles and emergencies during the construction of wells. The use of this method will minimize unproductive drilling time.

**Keywords:** machine learning, neural networks, detection of anomalies, prediction of troubles, hybrid simulation, drilling of oil and gas wells, geological and technological information, prevention of accidents and complications, artificial intelligence, automated system, neural network modeling

**Recommended citation:** Dmitrievsky A.N., Sboev A.G., Eremin N.A., Chernikov A.D., Naumov A.V., Gryaznov A.V., Moloshnikov I.A., Borozdin S.O., Safarova E.A. (2020). On increasing the productive time of drilling oil and gas wells using machine learning methods. *Georesursy = Georesources*, 22(4), pp. 79–85. DOI: <https://doi.org/10.18599/grs.2020.4.79-85>

#### Acknowledgments

The article was prepared based on the results of work carried out within the framework of the Program of State Academies of Sciences for 2013–2020. Section 9 «Earth Sciences»; direction of fundamental research: 132 «Integrated development and conservation of the Earth's interior; innovative processes for the development of mineral deposits and deep processing of mineral raw materials», on the topic of the state assignment «Fundamental basis of innovative technologies in the oil and gas industry», No. AAAA-A 16-116031750016-3; No. 0139-2019-0009 in Parus and No. AAAA-A19-119013190038-2 in ROSRID.

#### References

- Arkhipov A.I., Dmitrievsky A.N., Eremin N.A. et al. (2020). Data quality analysis of the station of geological and technological researches in recognizing losses and kicks to improve the prediction accuracy of neural network algorithms. *Neftyanoe khozyaystvo = Oil industry*, 8, pp. 63–67. DOI: 10.24887/0028-2448-2020-8-63-67
- Borozdin, S., Dmitrievsky, A., Eremin, N., Arkhipov, A., Sboev, A., Chashchina-Semenova, O., Fitzner L., Safarova, E. (2020). Drilling Problems Forecast System Based on Neural Network. *Society of Petroleum Engineers*. <https://doi.org/10.2118/202546-MS>
- Chen T., Guestrin C. (2016). Xgboost: A scalable tree boosting system. *Proc. 22nd ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*. ACM, pp. 785–794. <https://doi.org/10.1145/2939672.2939785>
- Diakonov A.G., Golovina A.M. (2017). Detection of anomalies in the work of mechanisms by machine learning methods. *Proc. XIX Int. Conf.: Analytics and data management in areas with intensive data use*. Moscow, pp. 469–476. (In Russ.)
- Dmitrievsky A.N., Eremin N.A., Filippova D.S. et al. (2020). Qualitative Analysis of Time Series GeoData to Prevent Complications and Emergencies During Drilling of Oil and Gas Wells. *Socar Proceedings*, 3, pp. 31–37. <http://dx.doi.org/10.5510/OGP20200300442>
- Dmitrievsky A.N., Eremin N.A., Filippova D.S., Safarova E.A. (2020). Digital oil and gas complex of Russia. *Georesursy = Georesources, Special issue*, pp. 32–35. DOI: <https://doi.org/10.18599/grs.2020.SI.32-35>
- Eremin N.A., Stolyarov V.E. (2020). On the digitalization of gas production processes at the late stages of field development. *Socar Proceedings*, 1, pp. 59–69. (In Russ.). DOI: 10.5510/ogp20200100424
- Eremin N.A., Vodopyan A.O., Duplyakin V.O., Chernikov A.D., Kosmos S.A. (2020a). Software component «Oil and Gas Blockchain». Software Registration Certificate RU 2020614626. Application form 2020613699.
- Eremin N.A., Dmitrievsky A.N., Chashchina-Semenova O.K., Fitsner L.K., Chernikov A.D. (2020b). Software component «Adaptation of generalized neural network models for predicting complications and emergencies to geophysical parameters when drilling a specific well». Software Registration Certificate RU 2020660890. Application form 2020660179.
- Eremin N.A., Dmitrievsky A.N., Chashchina-Semenova O.K., Fitsner L.K., Chernikov A.D. (2020c). Software component «Orchestration - integration of modules of the system for predicting complications and emergencies during drilling and well construction». Software Registration Certificate RU 2020660891. Application form 2020660181.
- Eremin N.A., Dmitrievsky A.N., Chashchina-Semenova O. K., Fitsner L.K., Chernikov A.D. (2020d). Software component «Neural network calculations - construction of models for forecasting complications and emergencies during drilling and well construction». Software Registration Certificate RU 2020660892. Application form 2020660182.

Eremin N.A., Dmitrievsky A.N., Chashchina-Semenova O.K., Fitsner L.K., Chernikov A.D. (2020e). Software component «Indication of forecast of complications and emergencies during drilling and well construction» (SC «Indication»). Software Registration Certificate RU 2020661356. Application form 2020660450.

Surina E. et al. (2020). Application of machine learning to accidents detection at directional drilling. *Journal of Petroleum Science and Engineering*, 184, 106519. <https://doi.org/10.1016/j.petrol.2019.106519>

Kanfar R. et al. (2020). Real-Time Well Log Prediction From Drilling Data Using Deep Learning. arXiv preprint arXiv: 2001.10156. <https://doi.org/10.2523/IPTC-19693-MS>

Kaznacheev P.F., Samoilova R.V., Kjurichinski N.V. (2016). Improving Efficiency of the Oil and Gas Sector and Other Extractive Industries by Applying Methods of Artificial Intelligence. *Ekonomicheskaya Politika = Economic Policy*, 11(5). (In Russ.)

Kodirov S.S., Shestakov A.L. (2019). The development of an artificial neural network for predicting stuck pipe columns of drill pipes. *Vestnik Yuzhno-Ural'skogo gosudarstvennogo universiteta. Seriya: Komp'yuternye tekhnologii, upravlenie, radioelektronika = Bulletin of the South Ural State University. Series Computer Technology, Automatic Control, Radio Electronics*, 19(3). (In Russ.)

Kohonen T. (1990). The self-organizing map. *Proc. IEEE*, 78(9), pp. 1464–1480. <https://doi.org/10.1109/5.58325>

Li Y. et al. (2019). Deep learning for well data history analysis. *Society of Petroleum Engineers*. <https://doi.org/10.2118/196011-MS>

Liu F.T., Tony T.K.M., Zhou Z.H. (2008). Isolation forest. *Proc. 8th IEEE Int. Conf. on Data Mining*, pp. 413–422. <https://doi.org/10.1109/ICDM.2008.17>

Noshi C.I., Schubert J.J. (2018). The Role of Machine Learning in Drilling Operations. A Review. *Society of Petroleum Engineers*. <https://doi.org/10.2118/191823-18ERM-MS>

Pichugin O.N., Prokofiev Y.Z., Alexandrov D.M. (2013). Decision Trees as an effective method for analysis and forecasting. *Neftepromyslovoe delo*, 11, pp. 69–75. (In Russ.)

#### About the Authors

*Anatoliy N. Dmitrievsky* – Dr. Sci. (Engineering), Professor, member of the Russian Academy of Sciences, Research advisor, Oil and Gas Research Institute of the Russian Academy of Sciences 3, Gubkin st., Moscow, 119333, Russian Federation

*Alexander G. Sboev* – Leading Researcher, Cand. Sci. (Engineering), National Research Center «Kurchatov Institute» 1, Ak. Kurchatov pl., Moscow, 123098, Russian Federation

*Nikolai A. Eremin* – Dr. Sci. (Engineering), Professor, Deputy Director for Innovation, Oil and Gas Research Institute of the Russian Academy of Sciences; Professor, National University of Oil and Gas «Gubkin University» 3, Gubkin st., Moscow, 119333, Russian Federation

*Alexander D. Chernikov* – Leading Researcher, Cand. Sci. (Engineering), Oil and Gas Research Institute of the Russian Academy of Sciences 3, Gubkin st., Moscow, 119333, Russian Federation

*Aleksandr V. Naumov* – Engineer, Oil and Gas Research Institute of the Russian Academy of Sciences 3, Gubkin st., Moscow, 119333, Russian Federation

*Artem V. Gryaznov* – Engineer, Oil and Gas Research Institute of the Russian Academy of Sciences 3, Gubkin st., Moscow, 119333, Russian Federation

*Ivan A. Moloshnikov* – Engineer, Oil and Gas Research Institute of the Russian Academy of Sciences 3, Gubkin st., Moscow, 119333, Russian Federation

*Sergei O. Borozdin* – Senior Lecturer, Department of Oil and Gas Wells Drilling, National University of Oil and Gas «Gubkin University» 65, Leninsky ave, Moscow, 119991, Russian Federation

*Elizaveta A. Safarova* – Junior Researcher, Oil and Gas Research Institute of the Russian Academy of Sciences 3, Gubkin st., Moscow, 119333, Russian Federation

Manuscript received 1 September 2020;

Accepted 19 September 2020; Published 11 December 2020