

ВЫЯВЛЕНИЕ ПЕРСПЕКТИВНЫХ НЕФТЯНЫХ МЕСТОРОЖДЕНИЙ НА ОСНОВЕ ПОСТРОЕНИЯ ПРОГНОЗНЫХ МОДЕЛЕЙ

Еременко М.С.¹, Губин Е.И.²

¹ НИ ТПУ, ИШИТР, 8ПМ1И, e-mail: eremenko@tpu.ru

² НИ ТПУ, ИШИТР, доцент ОИТ, e-mail: gubine@tpu.ru

Введение

На сегодняшний день нефтяная промышленность мира представляет собой одну из важнейших составляющих мирового хозяйства, а также влияет на развитие других отрасли. Одной из главных проблем современного развития нефтегазового комплекса России является проблема нерационального недропользования (низкий уровень извлечения запасов нефти) [1]. Есть потребность в оптимизации процессов добычи. Выявление зависимостей между качественными характеристиками нефти и физическими свойствами пласта дает возможность сэкономить на дорогостоящих лабораторных исследованиях и помогает в создании новых технологий для разработки трудноизвлекаемых запасов нефти.

Одной из главных задач, стоящих перед нефтегазовым комплексом, является оценка эффективности и значимости измеряемых параметров при добыче углеводородов. Хорошо известно насколько этот процесс сложен технически и финансово затратный. Поэтому критически важно выделить наиболее значимые параметры, которые влияют на величину добычи нефти.

Целью данной работы является оценить важность и влияние измеренных входных параметров на величину добычи нефти.

Основная часть

В данной работе сделана попытка на основе экспериментальных данных 185 нефтяных скважин, имеющих трещиноватую структуру грунта, оценить важность и влияние измеренных входных параметров на величину добычи нефти.

В качестве входных параметров (переменных) были взяты следующие данные в формате .xlsx (таблица 1).

Таблица 1

Исходные переменные и их размерность

Имя переменной	Расшифровка
R	Радиус зоны дренирования, м
<i>Q_{grp}</i>	Количество стадий ГРП, шт
W	Мощность продуктивного пласта, м
P_{plast}	Пластовое давление, атм
K	Проницаемость, мД
Noil	Вязкость нефти, сП
<i>Azimuth</i>	Азимут распространения трещины, градусы
M_p	Масса проппанта, тонн
St	Полудлина трещины (длина одного крыла), м
H_t	Высота трещины, м
L_t	Ширина трещины, мм
<i>P_z</i>	Забойное давление, атм
DeltaP	Депрессия (разница пластового и забойного давлений), атм
<i>D_f</i>	Дебит жидкости скважины, м³/сут
Doil	Дебит нефти скважины, м³/сут
<i>D_w</i>	Дебит воды скважины, м³/сут
Y	Целевая функция/ Дебит нефти скважины

Для решения поставленной задачи в качестве технического инструмента был использован метод логистической регрессии с использованием программного лицензионного обеспечения SAS.

Эффективность и предсказательная сила будущей модели во многом зависит от качества исходных данных. На основании методологии подготовки исходных, описанной в [2],[3] были проведены следующие шаги:

1. перевод исходного файла из формата .xlsx в формат SAS
2. проверку исходных данных на ошибки (*описки*);
- 2) проверку исходных данных на пропущенные значения («*missing*»);
- 3) проверку исходных данных на выбросы данных («*outliers*»);
- 4) проверку исходных данных на наличие дублирующих строк (*наблюдений*);
- 5) проверку исходных данных объясняющих переменных (*атрибутов*) на мультиколлинеарность;
- 6) трансформация исходных данных в цифровой формат («*цифровизация*»);
- 7) выбор целевой переменной;
- 8) разбиение исходных данных на тренировочную и тестовую выборки.

В таблице 1 жирным цветом выделены входные переменные, являющиеся наиболее значащими, полученными в результате предварительного статистического анализа [4]. Целевой переменной выбрана Doil, которая определяет дебит нефти скважины ($\text{м}^3/\text{сут}$). На рисунке 1 показана величина добычи Doil в зависимости от номера скважины.

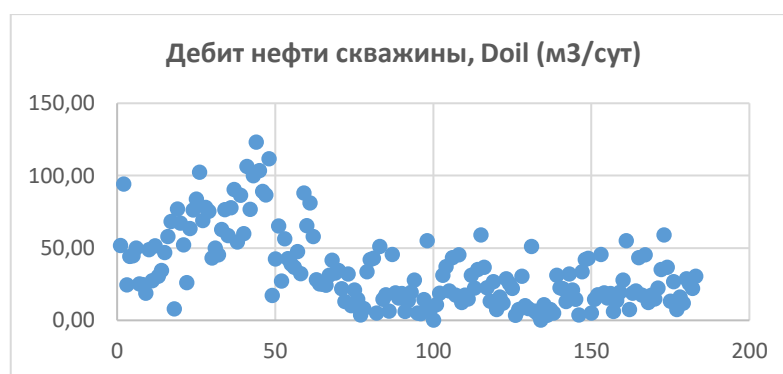


Рис. 1. Doil в зависимости от номера скважины

Для использования Doil в качестве безразмерной целевой функции в логистической регрессии возьмем «Y». «Y» принимает значение 0, если Дебит нефти скважины меньше среднего значения, и равен 1, если Дебит нефти скважины, больше или равен среднему значению ($35 \text{ м}^3/\text{сут}$). Поскольку количество исходных данных невелико (185 наблюдений) и используется только одна предсказательная модель, исходная выборка разделена на две – тренировочную (139) и тестовую (46) в соотношении 75% к 25% соответственно. Полученные выборки стратифицированы по целевой переменной.

В итоге проведенных процедур по подготовки исходных данных для использования логистической регрессии для оценки вероятности события (p) от входных параметров (x), мы получили 10 входных (вместо 13) безразмерных переменных и целевую переменную «Y». Ниже приведен график изменения вероятности дебита нефти скважины «Y» от входных параметров скважин для тренировочной выборки (рисунок 2). Результаты проведенного анализа показали, что наибольшее влияние на целевую функцию Y оказывают переменные: R (Радиус зоны дренирования), W (Мощность продуктивного пласта), K (Проницаемость), Noil (Вязкость нефти). Остальные переменные, в соответствии с выбранной моделью, существенного влияния на Y не оказывают.



Рис. 2. Изменения вероятности дебита нефти скважины «Y» от входных параметров скважин.

На рисунке 3 показан численный расчет вероятности Y от значимых параметров логистической регрессии R , W , K , $Noil$, включая константу для двух предельных случаев: а) $Y=1$ ($P=0,75$) – перспективная скважина и б) не перспективная скважина $Y=0$ ($P=0,40$).

$P_1/P_{1_}$	Int	R	W	K	$Noil$	Y
	-1,9	2,39	7,66	8,48	-5,5	
		0,95	0,074	0,12	0,22	
0,75/0,72	-1,9	2,2705	0,56684	1,0176	-1,21	$P=1$
		0,54	0,037	0,12	0,175	
0,4/0,5	-1,9	2,0076	0,28342	1,0176	-0,9625	$P=0$

Рис. 3. Расчёты вероятности дебита добычи от входных параметров

Достоверность полученной модели (ROC) для тренировочной выборки составляет 0,91, а для тестовой 0,81, что говорит о удовлетворительном качестве модели.

Заключение

В ходе выполнения работы создана методика подготовки данных и построена численная прогнозная модель. Разработанный подход в дальнейшем будет использоваться для создания алгоритма анализа эффективности работы эксплуатационных скважин.

Полученные результаты могут быть использованы для оптимизации процессов добычи нефти и газа, улучшения бизнес-процессов нефтегазовых компаний.

Список использованных источников

1. Саенко В.В. Нефтяная промышленность России – сценарии сбалансированного развития / Саенко В.В., Крюков В.А., Шмат В.В., Силкин В.Ю., Шафраник Ю.К., Бушуев В.В., Токарев В.А. – М.: Энергия, 2010. – 160 с.
2. Губин Е.И. Методика подготовки больших данных для прогнозного анализа «Наука и бизнес: пути развития». Выпуск № 3(105). 2020, 2020. – С. 33-35.
3. Губин Е.И. Методология подготовки больших данных для прогнозного анализа. Современные технологии, экономика и образование: Сборник трудов Всероссийской научно-методической конференции. / Томский политехнический университет. – Томск: Изд-во Томского политехнического университета, 2019. – 139с. – С. 25-28.
4. Кобзарь А.И. Прикладная математическая статистика. – М.: Физматлит, 2006. – 238 с.