ЛУКОЙЛ: Мониторинг разработки месторождения с использованием статистических методов анализа на примере перфорационных работ

LUKOIL: How Data Mining Enhances Oilfield Development

А.А. Анкудинов, к.т.н., Н.С. Полякова, Ю.Е. Радевич
Филиал ООО «ЛУКОЙЛ-Инжиниринг» «КогалымНИПИнефть» в г. Тюмени

A.A. Ankudinov, N.S. Polyakova,Y.E. Radevich,
KogalymNIPIneft Branch of LUKOIL-Engineering LLC Tuymen, RF, Tyumen

В условиях длительной разработки месторождения, с целью увеличения продуктивности скважин применяют различные геолого-технические мероприятия для интенсификации притока. Одни из них методов являются перфорационные работы (дострел не перфорированного ранее интервала и перестрел уже работающего). Особенностью данного метода является его невысокая стоимость, а прогнозирование эффекта может проводится по небольшому числу исходных параметров. Часто полученный эффект от мероприятия не соответствуют ожидаемому, в связи с чем возникает необходимость прогнозирования результатов с более высокой точностью.

В представленной работе в качестве объекта исследования выбран пласт БВ8 Повховского месторождения. Данный объект выбран в связи с достаточным для анализа объемом фактически проведенных дострелов: 1376 мероприятий. Выполнен поиск зависимостей между влияющими параметрами и эффективностью мероприятий для

Long-term field development plans dictate that a variety of well services are carried out in order to increase well productivity. Perforation jobs (perforation of previously unperforated intervals and reperforation) are one of these methods. Perforation is relatively low cost, and its effect can be predicted based on a small amount of initial data. In some cases, the outcome of an operation is not what is expected, and on this basis, it is necessary to try to predict the results as accurately as possible. The BV8 formation in the Povkhovskoye field was selected as a research target for this paper. The target was selected due to the volume (1376) of reperforated jobs. The correspondence between the influencing variables and the efficiency of the activities was explored to build a mathematical model with the use of standard “Data Evaluation” packages of MS Excel and the “Data Mining” methods (regression of the neural network, the linear regression) making it possible to select candidate wells for certain types of jobs. The approach applied in this research work can be employed at other production facilities and fields, as well as in other types of wells.

построения математических моделей с применением стандартного пакета «Анализ данных» MS Excel и методов Data Mining (регрессия нейронной сети, линейная регрессия), позволяющих находить скважины-кандидаты для определенных работ. Подход, используемый в работе, может быть распространен на другие эксплуатационные объекты и месторождения, а также на другие виды геолого-технических мероприятий.

Введение

В условиях длительной разработки месторождения с целью увеличения продуктивности скважин проводят различные геолого- технические мероприятия (ГТМ) по интенсификации притока. Одними из них являются перфорационные работы (дострел не перфорированного ранее интервала и перестрел уже работающего). Особенностью данного метода является его невысокая стоимость, а эффект может прогнозироваться по небольшому числу исходных параметров. В настоящее время накоплен значительный объем информации по мероприятиям данного типа. Целью работы является поиск зависимостей между влияющими параметрами и эффективностью мероприятий для построения математических моделей, позволяющих находить скважины-кандидаты для определенного вида работ на примере объекта БВ8 Повховского месторождения. Данный объект выбран в связи с достаточным для анализа объемом фактически проведенных мероприятий: 1376 дострелов.

Повховское месторождение открыто в 1972 г., введено в разработку в 1978 г. Месторождение находится в Сургутском районе Ханты- Мансийского автономного округа Тюменской области, в 80 км от г. Когалыма. В разработке находятся объекты: БВ8, Ач и ЮВ1. Месторождение многопластовое со сложным геологическим строением. Нефтеносность установлена в отложениях мегионской, баженовской, васюганской и тюменской свит. По величине запасов месторождение относится к категории крупных. Наибольший объем за- пасов приходится на продуктивный пласт БВ8, который характеризуется самой высокой проницаемостью. Объект БВ8 является основным, определяющим добычу месторождения. Основная площадь объекта разрабатывается с 1978 г. В настоящее время отбор составляет 85,2 % начальных извлекаемых запасов (НИЗ) при обводненности продукции 89,5 %. Обработка собранной информации по фактически выполненным дострелам включала: – построение моделей для оценки прироста дебита жидкости и изменения обводненности с применением пакета «Анализ данных» в MS Excel; – построение моделей с использованием методов Data Mining

Introduction

When a field is in long-term development, various well services are carried out to increase well stimulation. Perforation jobs are one of these methods, it is low cost and it's performance can be predicted with limited data. The purpose of this work has been to search for correlations between the influencing variables and the efficiency of the activities, and to build up mathematical models that make it possible to easily find candidate wells for certain types of jobs, using the BV8 target of the Povkhosvksoye field as an example.

The given target was selected due to the fact that it has a sufficient volume of perf jobs for the analysis. The Povkhovskoye field was discovered in 1972, and brought into development in 1978. The field is located in the Surgut district of the Khanty-Mansi Autonomous Area of Tyumen region, 80 km away from the town of Kogalym. The targets of BV8, Achimov and JV1 are currently under development. The field is multilayered, with a complex subsurface geology and oil-bearing capacity was confirmed in the deposits of the Megion, Bazhenov, Vasyugan and Tyumen Suites. Judging by the size of its' reserves, it can be referred to as a "major" field. The largest reserves are located on the BV8 productive reservoir which is characterized with the highest degree of permeability. The BV8 target is the core formation that determines the total field production volume and the target area has been developed since 1978. Total recovery amounts to 85.2% of the initial recoverable reserves, with water cut amounting to 89.5%. Processing the data on the previously performed perforation jobs comprised of the following steps: - model building to evaluate the liquid rate value and changes in water cut using the MS Excel "Data Evaluation" package; - model building with the use of the Data Mining methods (regression of the neural network, the linear regression) of MS Azur MLS. The first stage comprised of the collection and processing of the field geologic and engineering data. 5810 perforation jobs were performed. The activities that matched the following specific criteria were used for the analysis:

- the perforation jobs were not primary for a selected well, although they were primary for a selected interlayer (1653);
- the database contained additional perforation jobs performed since the year 2000 (1256);.
- additional perforation jobs were carried out without hydraulic fracturing operations (541);
- well shut-down period prior to additional perforation jobs was not exceeding 1 year (420);
- additional perforation jobs were performed at the producing well stock (213);
- additional perforation jobs were carried out with no additional activities following them (195);

(регрессия нейронной сети, линейная регрессия) в MS Azur MLS. На первом этапе исследования были проведены сбор и обработка геолого-промысловых и технологических данных по фактически проведенным дострелам. Общее число перфорационных работ составило 5810. Для анализа выбраны мероприятия, удовлетворяющие следующим критериям:

- перфорационные работы не являются первичными для конкретной скважины, но являются первичными для определенного прослоя (1653);
- в базу включены дострелы, осуществленные с 2000 г. (1256);
- дострелы проводились без гидроразрыва пласта (541);
- период простоя скважины перед проведением дострела не превышает 1 года (420);
- дострелы выполнены на добывающем фонде скважин (213);
- дострелы осуществлялись без проведения дополнительных мероприятий (195);

Регрессионные модели, полученнын в результате выполнения работы, легко реализуются и показывают достаточно высокие коэффициенты детерминации. Подход, используемый в работе, может быть распространен на другие эксплуатационные объекты и месторождения, а также на другие виды ГТМ

Исходя из формулы Дюпюи, основными параметрами, влияющими на дебит жидкости, являются: толщина, проницаемость пласта, депрессия, контур питания, вязкость и скин-фактор. Для построения математической модели оценки дебита жидкости приняты первые три параметра ввиду высокой точности их определения. В уравнение также добавлен параметр – максимальный дебит жидкости окружающих скважин, для оценки фактической продуктивности пласта. Для определения обводненности после проведения мероприятия приняты следующие параметры: величина текущих извлекаемых запасов нефти на 1 м толщины пласта – учитывает текущее насыщение пласта; обводненность до мероприятия – учитывает влияние уже подключенных прослоев в скважине на значение обводненности после мероприятия; прирост дебита жидкости после мероприятия – учитывает взаимовлияние фаз воды и нефти. В первую очередь в данной работе рассматривается построение моделей методом линейной регрессии с использованием встроенного пакета «Анализ данных» в MS Excel. Суть регресссии заключается в определении степени влияния каждого из предполагаемых влияющих параметров X1, … , Xn на зависимую переменную Y. Далее находят

The regression models obtained can be easily be implemented and they demonstrate sufficiently high coefficients of determination. The approach applied in this work can be employed at other production facilities and fields, as well as in other types of well stock activities.

According to the Dupuis formula, the basic data values having impact on the liquid rate level are as follows - formation thickness, permeability, drawdown, external reservoir boundary, viscosity and skin-factor. To build up a mathematical model for evaluation of the liquid rate level, the first three parameters were employed due to high accuracy of their determination. One more parameter, namely the maximum liquid rate from the offset wells, was included into the equation in order to evaluate the actual reservoir productivity. To determine the water cut obtained after performed activities, the following parameters were employed:

- The size of remaining recoverable reserves of oil per 1m of formation thickness – it takes into account the current stratal water saturation;
- The water cut before the performed job – it takes into account the impact the already connected interlayers provide for the value of water cut after the activities have been performed;
- The liquid rate increment after the performed activities – it takes into account the mutual interaction of the water and oil phases.

First of all, model building with the use of the linear regression methods has been considered in the given research work, with the help of the imbedded “Data Evaluation” package of MS Excel. The essence of the regression is in determination of the degree of impact of each of asserted influencing variables X1, … , Xn on the dependent variable Y. After that, the regression coefficients Ki are to be searched, for the best presentation of the liquid rate and water cut obtained after performed additional perforation jobs [1-3]

$$Y = \sum_{i=1}^{n} X_i K_i + b. \qquad (1)$$

Building a Model to Evaluate the Liquid Rate and Changes in Water Cut Using the MS Excel “Data Evaluation” Package

The first stage included model building to determine the liquid rate increment after additional perforation jobs. The regression coefficients, the constant error values, the Student’s t-distribution, and the P-values are presented in Table 1. The formula used to determine the liquid rate increment Yq, obtained after performed activities has the following view in the linear regression model:

N-G-K.ru
МОСКОВСКИЕ НЕФТЕГАЗОВЫЕ КОНФЕРЕНЦИИ

Встречи заказчиков и подрядчиков топливно-энергетического комплекса

Москва, улица Тверская, 22, отель InterContinental

05 Декабря 2019 | НЕФТЕГАЗШЕЛЬФ

Подряды на нефтегазовом шельфе

Заказчиками оборудования выступают “Газпром нефть”, “Роснефть”, “ЛУКОЙЛ”, “Газфлот” и другие крупные компании. В условиях введения экономических санкций необходимо быстро освоить производство жизненно важного оборудования, в первую очередь запасных частей

- **ТЭК-R.RU** — Выявление предприятий, способных работать для шельфа по итогам ежегодного опроса нефтегазовых компаний
- **ТЭК-ALL.RU** — База оборудования для нефтегазового шельфа
- **ТЭК-MAP.RU** — Настенная нефтегазовая карта

20 Февраля 2020 | ИНВЕСТЭНЕРГО

Инвестиционные проекты в электроэнергетике

Обзор инвестиционных проектов и модернизация российской электроэнергетики, вопросы материально-технического обеспечения в отрасли, практика закупочной деятельности в крупнейших российских компаниях

- **ТЭК-R.RU** — Награждение лучших поставщиков оборудования и услуг в электроэнергетике
- **ТЭК-MAP.RU** — Настенная карта по электроэнергетике

17 Марта 2020 | НЕФТЕГАЗСНАБ

Снабжение в нефтегазовом комплексе

Конференция собирает руководителей служб материально-технического обеспечения нефтегазовых компаний. Обсуждается организация закупочной деятельности, практика импортозамещения, оплата и приемка поставленной продукции, информационное обеспечение рынка

- **ТЭК-R.RU** — Награждение лучших производителей нефтегазового оборудования по итогам ежегодного опроса нефтегазовых компаний
- **ТЭК-ALL.RU** — База поставщиков нефтегазового комплекса
- **ТЭК-MAP.RU** — Настенная нефтегазовая карта

28 Мая 2020 | НЕФТЕГАЗСТРОЙ

Строительство в нефтегазовом комплексе

Формирование цивилизованного рынка в нефтегазовом строительстве, практика выбора строительных подрядчиков, создание российских EPC-фирм, увеличение доли российских компаний на нефтегазостроительном рынке, расценки и порядок оплаты проводимых работ

- **ТЭК-R.RU** — Награждение лучших строительных подрядчиков по итогам ежегодного опроса нефтегазовых компаний
- **ТЭК-ALL.RU** — База поставщиков нефтегазостроительных компаний
- **ТЭК-MAP.RU** — Настенная нефтегазовая карта

10 Сентября 2020 | НЕФТЕГАЗОПЕРЕРАБОТКА

Модернизация производств для переработки нефти и газа

Вопросы модернизации нефтеперерабатывающих и нефтехимических мощностей, проблемы взаимодействия с лицензиарами, практика импортозамещения, современные модели управления инвестиционными проектами, стандарты и требования безопасности

- **ТЭК-R.RU** — Награждение лучших производителей оборудования для модернизации нефтегазоперерабатывающих предприятий по итогам ежегодного опроса нефтегазовых компаний
- **ТЭК-ALL.RU** — База подрядчиков для модернизации НПЗ
- **ТЭК-MAP.RU** — Настенная нефтегазовая карта

29 Октября 2020 | НЕФТЕГАЗСЕРВИС

Нефтегазовый сервис в России

Традиционная площадка для встреч руководителей геофизических, буровых предприятий, а также компаний, занятых ремонтом скважин. Подрядчики в неформальной обстановке обсуждают актуальные вопросы со своими заказчиками – нефтегазовыми компаниями

- **ТЭК-R.RU** — Награждение лучших нефтесервисных компаний по итогам ежегодного опроса нефтегазовых компаний
- **ТЭК-ALL.RU** — База поставщиков нефтесервисных компаний
- **ТЭК-MAP.RU** — Настенная нефтегазовая карта

Телефоны: (495) 514-58-56, 514-44-68; факс: (495) 788-72-79; info@n-g-k.ru; n-g-k.ru

Download on the App Store
GET IT ON Google play

коэффициенты регрессии Ki для наилучшего представления дебита жидкости и обводненности после проведения дострела [1–3]

$$Y = \sum_{i=1}^{n} X_i K_i + b. \quad (1)$$

Построение моделей для оценки прироста дебита жидкости и изменения обводненности с применением пакета «Анализ Данных» В MS Excel

На первом этапе строилась модель для прироста жидкости после дострелов. Коэффициенты регрессии, значения стандартных ошибок, критерия Стьюдента t, P-значения представлены в табл. 1. Формула для определения прироста дебита жидкости Yq после проведения мероприятия в модели линейной регрессии имеет следующий вид:
Yq = -0,1965X1 + 0,0936X2 + 0,1861X3 + 0,0013X4 - 5,5673. (1)

Сопоставление расчетного дебита жидкости с фактическим после дострела показывает высокую сходимость полученных результатов, коэффициент детерминации R2 = 0,712 (рис. 1). По нефтенасыщенной толщине и депрессии Рзначение составляет более 0,05, что согласно правилам

Yq = -0,1965X1 + 0,0936X2 + 0,1861X3 + 0,0013X4 - 5,5673. (1)

The correlation of the estimated and actual liquid rate value, obtained after additional perforation jobs demonstrates the high precision of the data, the coefficient of determination R2 being 0.1712 (Fig.1). The P-value on the oil net pay and drawdown amounts to over 0.05, which, according to the rules of descriptive statistics, points to low impact of the factor on the resultant value or just means no effect. The results can be explained in the following way. First, there appears to be some uncertainty while selecting the version for interpretation of geological data. For instance, according to the actual version of the “Reinterpretation_2001”, the reservoir with oil net pay of 1.5m was selected, while the relevant parameter according to the version of the “CSRM-Neft_2011” amounts to 12.1m. In this connection, the results of the stream gauging agree with the second version of the interpretation. Secondly, the formation pressure values are taken off from the isobar map, the maps, in this regard, are prepared with the periodicity of once per quarter, based on a certain set of values obtained directly from wells, while the remaining part of the values can be obtained through calculations. The bottom-hole pressure is generally calculated through dynamic level which in many cases also downgrades the quality of initial data. Hereafter, a model was built to determine the water cut after performed additional

Показатели Data values		Коэффициенты регрессии Regression constants		Стандартная ошибка Constant error	Критерии Стьюдента t Student's t-distribution	P-значение P-value
Y-пересечение Y-crosscut	Const	b	-5,5673	1,630	-3,417	0,001
X_1	Эффективная нефтенасыщенная толщина / Oil net pay	K_1	-0,1965	0,151	-1,301	0,195
X_2	Проницаемость / Permeability	K_2	-0,0936	0,012	7,606	0,03
X_3	Максимальный дебит по окружающим скважинам / Max. flow rate from offset wells	K_3	-0,1861	0,014	12,843	0,0000057
X_4	Депрессия / Drawdown	K_4	-0,0013	0,014	0,096	0,924

Таблица 1 - Table 1

Показатели Data values		Коэффициенты регрессии Regression constants		Стандартная ошибка Constant error	Критерии Стьюдента t Student's t-distribution	P-значение P-value
Y-пересечение Y-crosscut	Const	b	19,0754	3,083	6,188	$4,3 \cdot 10^{-9}$
X_1	Трудноизвлекаемые запасы / начальная нефтенасыщенность / Hard-to-recover reserves / initial oil saturation	K_1	-0,1293	0,059	-2,176	$3,1 \cdot 10^{-2}$
X_2	Прирост дебита жидкости Liquid rate increment	K_2	0,4107	0,088	4,685	$5,7 \cdot 10^{-6}$
X_3	Обводненность до дострела Water cut before perforation of additional intervals	K_3	0,7088	0,037	19,260	$6,8 \cdot 10^{-45}$

Таблица 2 - Table 2

EAGE
EUROPEAN ASSOCIATION OF GEOSCIENTISTS & ENGINEERS

2020 Санкт-Петербург

9-я международная конференция
Геонауки: трансформируем знания в ресурсы

Важные даты

Окончание регистрации по льготному тарифу
25 декабря 2019 г.

Окончание предварительной регистрации
25 марта 2020 г.

Окончание приема заявок на коммерческие презентации и рекламное место компании
1 марта 2020 г.

Темы научных сессий

1 Региональные, тематические и научные геологические исследования
2 Геологическое изучение и поэтапное освоение Арктической зоны и шельфов
3 Практические примеры при ГРР
4 Разведка и разработка месторождений
5 Круглый стол: Перспективы внедрения и развития интеллектуальных систем в процессах разведки и добычи

Научная программа опубликована на сайте www.eage.ru

РЕГИСТРАЦИЯ ОТКРЫТА!

6–9 АПРЕЛЯ 2020 ГОДА I САНКТ-ПЕТЕРБУРГ, РОССИЯ

WWW.EAGE.RU | WWW.EAGE.ORG

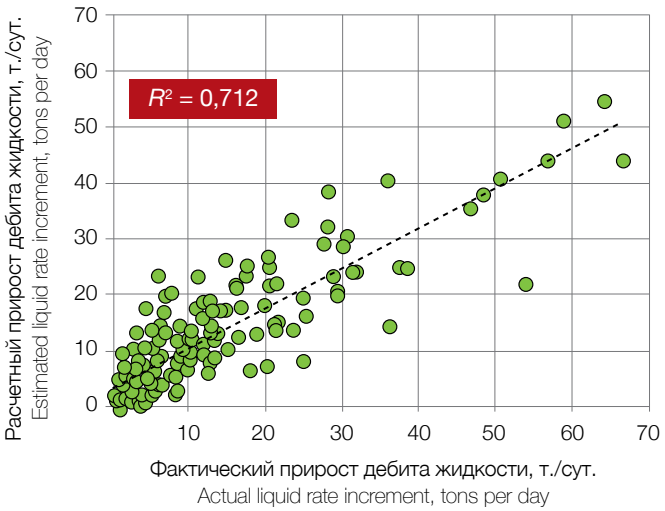

Рис.1: Сопоставление рассчитанных и фактических значений прироста дебита жидкости после проведения дострела
Fig.1: Comparison of the estimated and actual values of liquid rate increment after perforation of additional intervals

описательной статистики означает слабое влияние фактора на результирующий параметр или отсутствие влияния. Полученные результаты можно объяснить следующим образом. Во-первых, возникает неопределенность в выборе версии интерпретации геологофизических параметров. Например, в скв. 2598 по актуальной версии «Переинтерпретация_2011 год» выделен коллектор с нефтенасыщенной толщиной, равной 1,5 м, а по версии «ООО_ ЦСРМнефть_2011 год» соответствующий параметр составляет 12,1 м, при этом результаты потокометрии согласуются со второй версией интерпретации. Во-вторых, пластовое давление снимается с карты изобар, при этом карты строятся с периодичностью один раз в квартал по некоторому числу значений, полученных непосредственно в скважинах, остальная часть значений рассчитывается.

Забойное давление в большинстве случаев пересчитывается через динамический уровень, что также снижает качество исходной информации. Далее проводилось построение модели для определения обводненности после проведения дострелов. Коэффициенты регрессии, значения стандартных ошибок, критерия Стьюдента t, P-значения представлены в табл. 2. Формула для определения обводненности Yf после проведения мероприятия в модели линейной регрессии выглядит следующим образом:

Yf = -0,1293X1 + 0,4107X2 + 0,7088X3 + 19,0754. (2)

Сопоставление расчетной и фактической обводненности после проведения дострела показывает высокую сходимость полученных результатов, R2 = 0,8185 (рис. 2).

perforation jobs. The coefficients of regress, the constant error values, the Student's t-distribution, and the P-values are presented in Table 2. The formula determining the water cut Yf after the performed activities looks, in the linear regression model, as follows:

Yf = -0,1293X1 + 0,4107X2 + 0,7088X3 + 19,0754. (2)

Correlation of the estimate and actual water cut values, after the additional perforation jobs were performed, demonstrates high precision of data, R2 being 0.8185 (Fig.2).

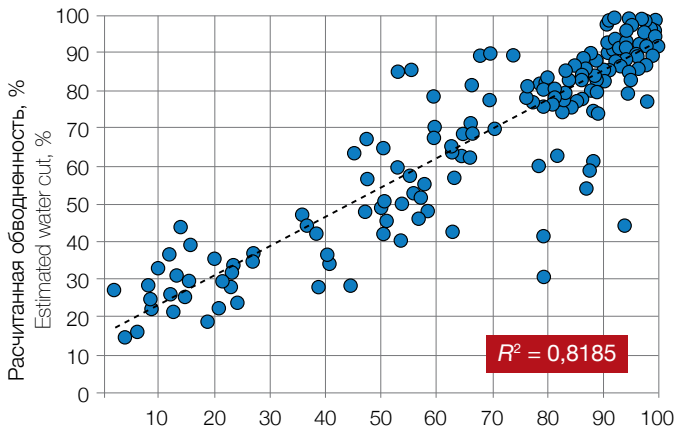

Рис.2: Сопоставление рассчитанных и фактических значений обводненности после проведения дострела
Fig.2: Correlation of the estimated and actual values of water cut after perforation of additional intervals

Data Mining is a catch all term used when using complex of methods to discover previously unknown interpretations that decision making.[4]. Data Mining is used in such disciplines as statistics, the information theory and machine learning. Therefore, the major part of its algorithms was elaborated on the basis of different methods from those disciplines. As a concept, Data Mining comprises: the theory of databases, image identification, descriptive statistics, machine learning, visualization, algorithms and artificial intelligence. In this regard, you can use a variety of software products to implement your Data Mining objectives, such as R-Studio, Microsoft Azure MLS, Orange and many others.

Model Building with the Use of the Microsoft Azure MLS Software

The presented work considers two methods: the linear regression method and the method of regression of the neural network implemented in the Microsoft Azure MLS software (Fig.3). This software is sufficiently easy to use, free to download and install, and has the imbedded modules where various algorithms are embodied. That

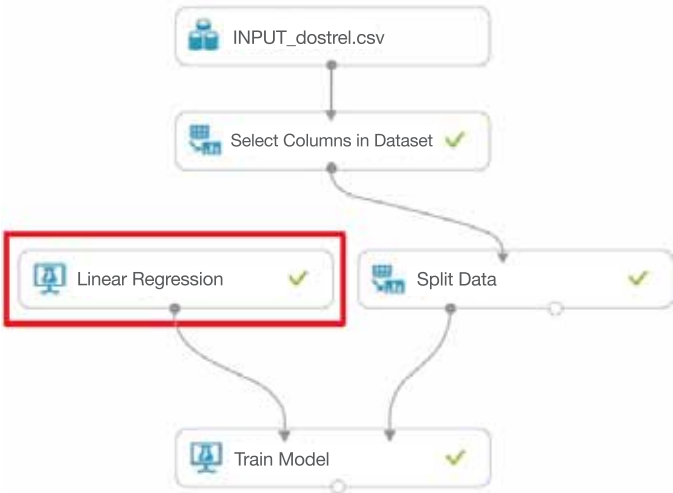

Рис.3: Построение модели линейной регрессии в MS Azure MLS
Fig.3: Build-up of the MS Azure MLS linear regression model

Data Mining – собирательное название, используемое для обозначения совокупности методов обнаружения ранее неизвестных, нетривиальных, практически полезных и доступных интерпретации знаний, необходимых для принятия решений в различных сферах человеческой деятельности [4]. Data Mining развивалась на стыке таких дисциплин, как статистика, теория информации, машинное обучение, поэтому большинство алгоритмов было разработано на основе различных методов из этих дисциплин. Data Mining включает: теорию баз данных, распознавание образов, описательную статистику, машинное обучение, визуализацию, алгоритмизацию, искусственный интеллект. При этом существует множество программ для реализации задач Data Mining, такие как R-Studio, Microsoft Azure MLS, Orange и др.

Построение моделей с применением программного продукта Microsoft Azure MLS

В представленной работе рассмотрены два метода: метод линейной регрессии и метод регрессии нейронной сети, реализованные в программном продукте Microsoft Azure MLS (рис. 3). Данный софт достаточно прост в понимании, бесплатный, имеет встроенные модули, в которых реализованы различные алгоритмы, при этом соблюдается конфиденциальность загружаемых данных в связи с их обезличенностью. Создание эксперимента в студии машинного обучения включает следующие шаги: – передача набора данных в виде CSV-файла (табличные данные в текстовом формате) в студию машинного обучения; – создание эксперимента и использование модуля Select Columns in Dataset (выбор столбцов в наборе данных) для выбора тех же признаков данных, что используются в MS Excel;

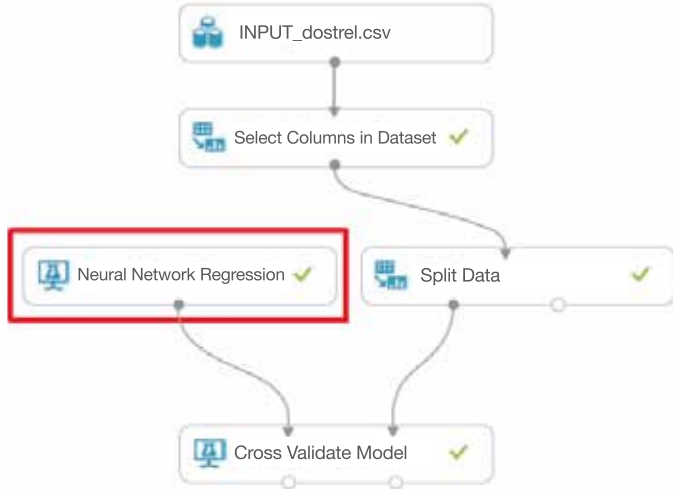

Рис.4: Построение модели «Регрессия нейронной сети» в MS Azure MLS
Fig.4: Build-up of the MS Azure MLS “Regression of the neural network” model

said, confidentiality and protection of the uploaded data is also preserved. Creating an experiment in Azure Machine Learning Studio comprises the following steps: the transmission of a data set in the format of CSV-file (tabular data in the format of text) to the Machine Learning Studio;

- Creating an experiment and using the Module “Select Columns in Dataset” (selection of columns in the data set) to select the same data features as they are used in MS Excel;

- Using the “Split Data” Module in the mode of relative expression to split data into the same datasets for learning as they are applied in MS Excel. When using the “Linear regression” Module in Azur Machine Learning Studio, there are two methods available to apply: The Gradient Descent method in the online mode (used for large-scale problems), the method of Least Squares (the mostly spread method of the linear regression). This would be an optimal selection for a small data amount. Table 3 displays correlation of the coefficients obtained in the models of MS Excel and MS Azure MLS for estimation of the liquid rate increment. It is obvious that the obtained models have similar data values.

After that, the method of Data Mining was considered in the research work, and namely, the “Regression of the Neural Network” (Fig. 4). The algorithm of the neural network (Microsoft) is useful for the evaluation of large datasets. The algorithm of the neural network in the SQL Server Analysis Services (Microsoft) combines every possible state of the initial data value with every probable state of the predicted data value, applying it to calculate the probabilities.

• использование модуля Split Data (разделение данных) в режиме относительного выражения для разделения данных на точно такие же наборы данных для обучения, как в MS Excel. При использовании модуля «Линейная регрессия» в студии машинного обучения доступны два метода: метод градиентного спуска в режиме онлайн (для крупномасштабных проблем); метод наименьших квадратов (наиболее распространенный метод линейной регрессии). Это оптимальный выбор для небольшого объема данных. В табл. 3 представлено сравнение коэффициентов, полученных в моделях MS Excel и MS Azure MLS для расчета прироста дебита жидкости. Видно, что полученные модели имеют схожие показатели.

Показатели Data values		Коэффициенты модели Coefficients of the model equation	
		MS Excel	MS Azure MLS
Y-пересечение Y-crosscut	Const	-5,5673	-5,6703
X_1	Эффективная нефтенасыщенная толщина / Oil net pay	-0,1965	-0,1965
X_2	Проницаемость / Permeability	0,0936	0,0936
X_3	Максимальный дебит по окружающим скважинам / Max. flow rate from offset wells	0,1861	0,1861
X_4	Депрессия / Drawdown	0,0013	0,0013

Таблица 3 - Table 3

Далее в работе был рассмотрен метод Data Mining – «Регрессия нейронной сети» (рис. 4). Алгоритм нейронной сети (Microsoft) полезен при анализе большого объема данных. В службах SQL Server Analysis Services алгоритм нейронной сети (Microsoft) сочетает каждое возможное состояние входного показателя с каждым возможным состоянием прогнозируемого показателя и использует обучающие данные для вычисления вероятностей.

Далее эти вероятности можно применять для прогнозирования рассматриваемого параметра на основе входных данных. В модели интеллектуального анализа данных, которая создается при помощи алгоритма нейронной сети Microsoft, может содержаться несколько сетей. Их число определяется числом столбцов, используемых для входа. Алгоритм создает сеть, состоящую из двух или трех слоев нейронов. К таким слоями относятся входной, необязательный скрытый и выходной слои. Модель нейронной сети должна содержать ключевой столбец, один или несколько входных и прогнозируемых столбцов.

На рис. 5 приведены результаты расчета прироста дебита жидкости и изменения обводненности, с помощью моделей, построенных в MS Azure MLS методом «Регрессия нейронной сети». Из рис. 1, 2 и 5 видно, что модель, построенная в MS Azure MLS методом «Регрессия нейронной сети», характеризуется

Further, these probabilities can be applied to predict a considered parameter, based on the input data. The model of data mining, created with the use of the Microsoft neural network algorithm, can contain several networks. Their number is determined with the number of columns used for data input. The algorithm creates a network consisting of two or three layers of neurons.

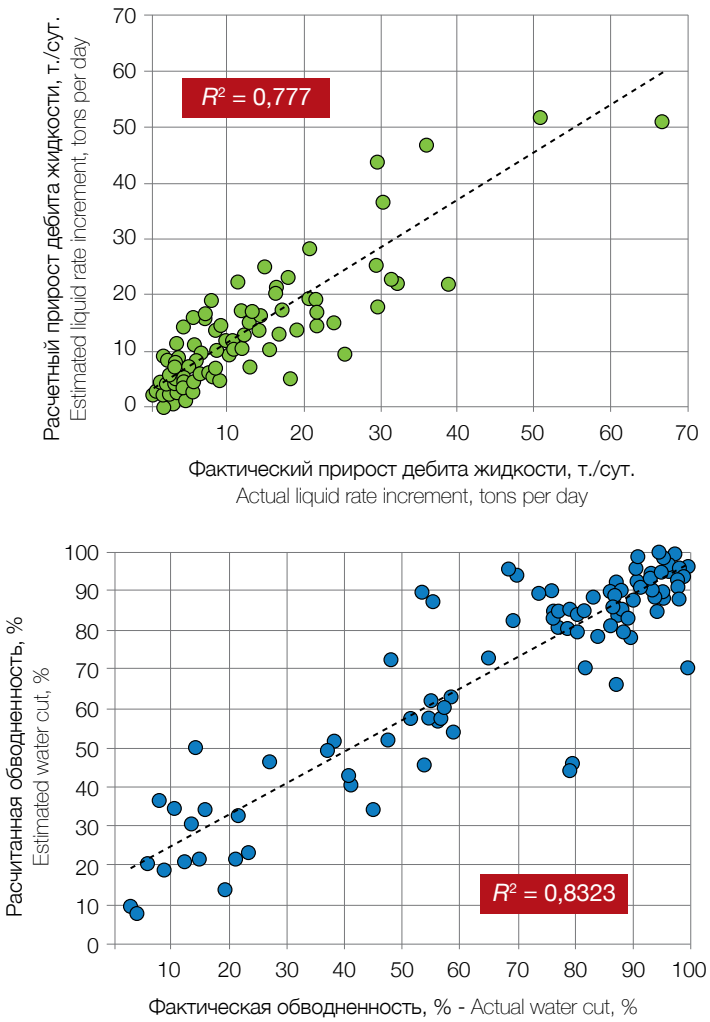

Рис.5: Результаты построения модели методом «Регрессия нейронной сети» в MS Azure
Fig.5: Simulation results of the MS Azure "Regression of the neural network" model

несколько лучшей сходимостью с фактическими показателями.

Заключение

Внедрение Big Data технологий в процесс разработки месторождений является неоспоримо перспективным направлением. Выполненная работа указывает на возможность применения методов Data Mining для анализа исходных данных по фактически проведенным операциям дострелов, а также для прогноза эффективности планируемых. Регрессионные модели, полученные в результате выполнения работы, легко реализуются и показывают достаточно высокие коэффициенты детерминации. Подход, используемый в работе, может быть распространен на другие эксплуатационные объекты и месторождения, а также на другие виды ГТМ.

Список литературы

1. Колемаев В.А., Староверов О.В., Турундаевский В.Б. Теория вероятностей и математическая статистика. – М.: Высшая школа, 1991. – 400 с.

2. Макарова Н.В., Трофимец В.Я. Статистика в Excel. – М.: Финансы и статистика, 2002. – 368 с.

3. Савельев В. Статистика и котики. – М.: Издательские решения, 2017. – 126 с.

4. http://pandia.ru/text/78/271/8297.php

Авторы статьи:

А.А. Анкудинов, к.т.н.,

Н.С. Полякова,

Ю.Е. Радевич

Филиал ООО «ЛУКОЙЛ-Инжиниринг» «КогалымНИПИнефть» в г. Тюмени

Материал любезно предоставлен компанией ПАО «Газпром нефть» и журналом «PROнефть»

The layers refer to the input data layer, optional hidden data layer, and the output data layer. The model of the neural network must contain a key column, one or several input data columns and the predicted data value columns.

Fig.5 presents the calculation data of the liquid rate increment and changes in water cut, obtained with the use of the models built in MS Azure MLS by the “Regression of the neural network” method. Comparing Figures 1, 2 and 5, we can see that the model built in MS Azure MLS, using the “Regression of the neural network” method, is characterized with somewhat better precision in relation to the actual data values.

Conclusion

The introduction of “Big Data” technologies into the oil and gas field development process, from our perspective, has been a success. The presented research work points at the possibility of using the Data Mining methods to evaluate the initial data based on previously conducted perforation jobs, as well as for predicting the efficiency of scheduled jobs. The regression models obtained as a result of this work can be easily implemented and demonstrate sufficiently high coefficients of determination. The approach applied in this work can be employed at other production facilities and fields, as well as in other well stock activities.

References

1. Kolemaev V.A., Staroverov O.V., Turundaevskiy V.B., Teoriya veroyatnostey i matematicheskaya statistika (Theory of probability and mathematical statistics), Moscow: Vysshaya shkola Publ., 1991, 400 p.

2. Makarova N.V., Trofimets V.Ya., Statistika v Excel (Excel statistics), Moscow: Finansy i statisti-ka Publ., 2002, 368 p.

3. Savel’ev V., Statistika i kotiki (Statistics and kitties), Moscow: AST Publ., 2017, 126 p.

4. http://pandia.ru/text/78/271/8297.php

Article Authored By:

A.A. Ankudinov,
Ph.D. Engineering,

N.S. Polyakova,

Y.E. Radevich

Branch of “LUKOIL Engineering” LC “KogalymNIPINeft” in Tyumen

Published with thanks to Gazprom Neft & PROneft Magazine