АЛГОРИТМ ОПРЕДЕЛЕНИЯ ЦЕЛЕСООБРАЗНОСТИ ПРОВЕДЕНИЯ ГЕОЛОГО-ТЕХНИЧЕСКИХ МЕРОПРИЯТИЙ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ DATA MINING

DETERMINATION OF OPTIMAL WELL INTERVENTIONS USING DATA MINING

М. Н. Харисов
Mansur N. Kharisov

Уфимский государственный
авиационный технический
университет,
г. Уфа, Российская Федерация

Ufa State Aviation
Technical University,
Ufa, Russian Federation

Э. А. Юнусова
Eleonora A. Yunusova

Уфимский государственный
нефтяной технический
университет,
г. Уфа, Российская Федерация

Ufa State Petroleum
Technological University,
Ufa, Russian Federation

А. М. Вагизов
Azat M. Vagizov

Уфимский государственный
нефтяной технический
университет,
г. Уфа, Российская Федерация

Ufa State Petroleum
Technological University,
Ufa, Russian Federation

Р. А. Майский
Ravil A. Maiski

Уфимский государственный
нефтяной технический
университет,
г. Уфа, Российская Федерация

Ufa State Petroleum
Technological University,
Ufa, Russian Federation

Э. Р. Валиев
Emil R. Valiev

Уфимский государственный
авиационный технический
университет,
г. Уфа, Российская Федерация

Ufa State Aviation
Technical University,
Ufa, Russian Federation

Проведение геолого-технических мероприятий (ГТМ) является важным этапом процесса рациональной разработки зрелых месторождений углеводородного сырья. Применение методов Data Mining позволяет повысить эффективность определения целесообразности проведения ГТМ в условиях несовершенства имеющихся геолого-промысловых данных, необходимых для математического описания сложных фильтрационных процессов. С целью повышения качества оперативного определения целесообразности проведения геолого-технических мероприятий авторами разработан алгоритм с использованием метода Fuzzy Logic, позволяющий повысить точность оценки технологического эффекта от планируемых ГТМ и, соответственно, снизить неопределенности и риски при разработке зрелых нефтяных месторождений. Предложенный алгоритм предполагает применение трапецидальных функций принадлежности базового терм-множества лингвистических переменных, нечеткой базы знаний, включающей в себя продукционные правила, состоящие из предпосылок и заключений вида: ЕСЛИ предпосылка И предпосылка И ..., ТО заключение, импликацию нечетких высказываний и агрегацию функций принадлежности заключений в соответствии с формулами Мамдани, а также дефазификацию итогового нечеткого подмножества с целью определения четкого значения функции целесообразности проведения ГТМ методом центра тяжести.

Ключевые слова

нефтяное месторождение,
планирование
геолого-технических
мероприятий,
Data Mining,
нечеткая логика

Performing of well interventions is an important stage of the optimal brownfield development. The implementation of Data Mining techniques allows increasing the efficiency of the well workover planning in case of insufficient data quantity and quality for the mathematical modeling of the fluid flow. In order to improve the efficiency of the determination of optimal well interventions, a fuzzy logic algorithm was developed. The proposed algorithm can increase the accuracy of the prediction of well interventions production rates, and reduce uncertainties and risks related to the brownfield development. It involves using trapezoidal membership functions for the terms of linguistic variables, a set of IF-THEN fuzzy rules with one or several conditions and a conclusion, Mamdani implication and aggregation functions, and defuzzification based on the Center of Gravity method.

Key words

oil field,
well intervention planning,
Data Mining, Fuzzy Logic

Значительная часть эксплуатируемых нефтяных месторождений Российской Федерации находится на поздней стадии разработки, характеризующейся высокой степенью выработанности запасов углеводородного сырья (УВС) [1]. Зрелые месторождения являются основным источником добычи нефти, и для ее поддержания предприятия проводят большой объем геолого-технических мероприятий (ГТМ).

Основной проблемой при проведении ГТМ является недостижение планового прироста дебита нефти фактическим значением данного показателя (рисунок 1) [2], что подтверждает недостаточно высокую эффективность имеющихся решений по определению целесообразности проведения геолого-технических мероприятий.

В условиях недостаточного для геолого-гидродинамического моделирования качества

данных по фильтрационно-емкостным свойствам пород, а также физико-химическим свойствам флюидов месторождений УВС широкое применение получили статистические методы прогнозирования технологических показателей разработки, а также методы Data Mining, позволяющие определять скрытые зависимости между успешностью геолого-технических мероприятий и факторами, влияющими на целесообразность проведения ГТМ [3].

Эффективным методом решения задач в системах управления производственными процессами является нечеткая логика (Fuzzy Logic), позволяющая использовать результаты практического наблюдения за системой, сложно поддающейся математическому описанию [4]. К основным преимуществам нечеткой логики относятся [4]:

— возможность использования несовершенных данных и нечетких знаний при со-

Рисунок 1. Успешность проведенных ГТМ

хранении высокого качества получаемых решений;

— решение сложных комплексных задач без предварительного моделирования системы с использованием качественных оценок, основанных на практическом опыте со снижением временных и ресурсных затрат;

— возможность саморегулирования системы управления, основываясь на данных, которые не могут быть учтены в случае применения традиционных методов.

Таким образом, проблема повышения качества оперативного определения целесообразности проведения ГТМ путем использования метода Fuzzy Logic является актуальной.

Целью настоящей работы является разработка алгоритма определения целесообразности проведения геолого-технических мероприятий с использованием методов Data Mining.

В ходе данного исследования были использованы методы системного анализа, интерполяционного кригинга, корреляционного анализа и нечеткой логики.

Системный анализ научной литературы предметной области с использованием реферативных баз данных рецензируемых публикаций Web of Science, Scopus, а также научной поисковой системы Google Scholar, подтвердил научную новизну предложенного алгоритма, позволяющего, в отличие от существующих, используя несовершенные геолого-промысловые данные методом нечеткой логики с саморегулированием оперативно определять целесообразность проведения геолого-технических мероприятий.

Корреляционный анализ геолого-промысловых данных

С целью снижения размерности геолого-промысловых данных в настоящей работе применяется корреляционный анализ, позволяющий определить геолого-промысловые параметры со слабой связью по шкале Чеддока [5] с параметром целесообразности проведения ГТМ (SR), описываемым следующим уравнением:

$$SR_i = \frac{1}{2} \cdot \left(\frac{q_{0i}}{\max_i q_0} + \frac{Q_{0i}}{\max_i Q_0} \right), \quad (1)$$

где q_0 — прирост дебита нефти, т/сут; Q_0 — дополнительная добыча нефти за анализируемый период, тыс т; i — номер фактически проведенного геолого-технического мероприятия.

Сила корреляционной связи анализируемых параметров X и Y определяется следующим образом:

$$r = \max(r_P, r_S), \quad (2)$$

где r_P — коэффициент корреляции Пирсона, определяемый следующим образом [6]:

$$r_P = \frac{\sum_{i=1}^n ((x_i - \bar{x})(y_i - \bar{y}))}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (3)$$

где x_i — i -е значение параметра X ; $\bar{x}$ — среднее арифметическое значение параметра X ; y_i — i -е значение параметра Y ; $\bar{y}$ — среднее арифметическое значение параметра Y ;

r_S — коэффициент корреляции Спирмена, описываемый следующим уравнением [6]:

$$r_S = \frac{\sum_{i=1}^n ((R_{x_i} - \bar{R}_x)(R_{y_i} - \bar{R}_y))}{\sqrt{\sum_{i=1}^n (R_{x_i} - \bar{R}_x)^2} \sqrt{\sum_{i=1}^n (R_{y_i} - \bar{R}_y)^2}}, \quad (4)$$

где R_{x_i} — ранг i -го значения параметра X ; $\bar{R}_x$ — средний ранг значений параметра X ; R_{y_i} — ранг i -го значения параметра Y ; $\bar{R}_y$ — средний ранг значения параметра Y .

Использование выражения (2) позволяет эффективно определять силу корреляционной связи как для случая линейной зависимости параметров, так и для случая нелинейной монотонной зависимости [6].

Алгоритм определения целесообразности проведения ГТМ с использованием нечеткой логики

Блок-схема предложенного в рамках настоящей работы алгоритма определения целесообразности проведения ГТМ с использованием методов Data Mining представлена на рисунке 2.

Построение карт геолого-промысловых параметров осуществлялось методом интерполяционного кригинга [7].

При фаззификации геолого-промысловых параметров формирование базового термножества лингвистических переменных производилось с использованием трапециевидных функций принадлежности (рисунок 3).

Сформированная нечеткая база знаний включает в себя продукционные правила, состоящие из предпосылок и заключения (таблица 1). Продукционные правила имеют вид: ЕСЛИ предпосылка И предпосылка И..., ТО заключение.

Импликация нечетких высказываний и агрегация функций принадлежности заключе-

Рисунок 2. Алгоритм определения целесообразности проведения геолого-технических мероприятий

Рисунок 3. Функции принадлежности базового терм-множества лингвистических переменных

Таблица 1. Пример продукционных правил нечеткой базы знаний

Если	Лингвистическая переменная № 1	И	Лингвистическая переменная № 2	То	Лингвистическая переменная № 3
	Предпосылка 1.1		Предпосылка 2.1		Заключение 3.1
	Предпосылка 1.2		Предпосылка 2.2		Заключение 3.1
	Предпосылка 1.3		Предпосылка 2.3		Заключение 3.1

ний осуществлялись в соответствии с алгоритмом Мамдани [8].

Дефаззификация итогового нечеткого подмножества с целью определения четкого значения функции целесообразности проведения ГТМ проводилась методом центра тяжести [9]:

$$y = \frac{\int_{\min}^{\max} x \cdot \mu(x) dx}{\int_{\min}^{\max} \mu(x) dx}, \quad (5)$$

где x — анализируемый параметр; $\mu(x)$ — функция принадлежности лингвистической переменной анализируемого параметра.

Выводы

С целью анализа вопроса оперативного определения целесообразности проведения геолого-технических мероприятий авторами был проведен системный анализ научной литературы из реферативных баз данных Web of Science, Scopus, а также с использованием поисковой системы научной литературы Google Scholar. Результаты анализа показали актуальность проблемы повышения эффективности рациональной разработки зрелых нефтяных месторождений с использованием ГТМ путем применения методов Fuzzy Logic для определения целесообразности проведения геолого-

технических мероприятий и подтвердили научную новизну предложенного в настоящей работе алгоритма.

Разработанный алгоритм определения целесообразности проведения ГТМ с использованием методов Data Mining способен результативно решать задачу повышения эффективности рациональной разработки месторождений УВС в условиях недостаточного для геолого-гидродинамического моделирования качества данных по фильтрационно-емкостным свойствам пород, а также физико-химическим свойствам флюидов.

СПИСОК ИСПОЛЪЗУЕМЫХ ИСТОЧНИКОВ

1. Бобылев Ю.Н., Расенко О.А. Нефтяной сектор экономики России: основные тенденции М.: Издательский дом «Дело» РАНХиГС, 2016. 68 с.

2. Илюшин П.Ю., Рахимзянов Р.М., Соловьев Д.Ю., Колычев И.Ю. Анализ проведения геолого-технических мероприятий по увеличению продуктивности добывающих скважин на нефтяных месторождениях Пермского края // Вестник Пермского национального исследовательского политехнического университета. Геология, нефтегазовое и горное дело. 2015. № 15. С. 81–89.

3. Харисов М.Н., Карпов А.А., Петров С.В., Дарий С.Д. Алгоритм определения оптимальных характеристик вытеснения // Нефтяное хозяйство. 2018. № 5. С. 56–59.

4. Шеври Ф., Гели Ф. Нечеткая логика // Техническая коллекция Schneider Electric. 2009. № 31. С. 1–31.

5. Chaddock R.E. Principles and Methods of Statistics. Houghton Mifflin Company, 1925.

6. Chok N.S. Pearson's versus Spearman's and Kendall's correlation coefficients for continuous data: Master's of Science Thesis. University of Pittsburgh, 2010.

7. Lichtenstern A. Kriging Methods in Spatial Statistics: Bachelor's Thesis. München, Technische Universität, 2013.

8. Рогозин О.В. Метод нечеткого вывода решения в задаче подбора программного обеспечения на основе качественных характеристик этого обеспечения как объекта инвестиций // Качество. Инновации. Образование. 2009. № 2. С. 45–54.

9. Bai Y., Wang D. Fundamentals of Fuzzy Logic Control — Fuzzy Sets, Fuzzy Rules and Defuzzifications // Advanced Fuzzy Logic Technologies in Industrial Applications. London: Springer, 2006. P. 17–36.

REFERENCES

1. Bobylev Yu.N., Rasenko O.A. *Neftyanoi sektor ekonomiki Rossii: osnovnye tendentsii* [Oil Sector of the Russian Economy: Main Trend]. Moscow, Delo Publ., 2016. 68 p. [in Russian].

2. Ilyushin P.Yu., Rakhimzianov R.M., Solov'ev D.Iu., Kolychev I.Iu. Analiz provedeniya geologo-tekhnicheskikh meropriyatii po uvelicheniyu produktivnosti dobyvayushchikh skvazhin na neftyanykh mestorozhdeniyakh Permskogo kraya [Analysis of Well Intervention Aimed at Oil Production Enhancement in the Perm kraj's Fields]. Vestnik Permskogo natsional'nogo issledovatel'skogo politekhnicheskogo universiteta. *Geologiya, neftegazovoe i gornoe delo* — Perm Journal of Petroleum and Mining Engineering, 2015, No. 15, pp. 81–89. [in Russian].

3. Kharisov M.N., Karpov A.A., Petrov S.V., Darii S.D. Algoritm opredeleniya optimal'nykh kharakteristik vytesneniya [Algorithm for the Determination of Displacement Characteristics]. *Neftyanoe khozyaistvo* — Oil Industry, 2018, No. 5, pp. 56–59. [in Russian].

4. Shevri F., Geli F. Nechetkaya logika [Fuzzy Logic]. *Tekhnicheskaya kolleksiya Schneider Electric* — Technical Collection Schneider Electric, 2009, No. 31, pp. 1–31. [in Russian].

5. Chaddock R.E. [Principles and Methods of Statistics]. Houghton Mifflin Company, 1925.

6. Chok N.S. *Pearson's Versus Spearman's and Kendall's Correlation Coefficients for Continuous Data*. Master's of Science Thesis, University of Pittsburgh, 2010.

7. Lichtenstern A. *Kriging Methods in Spatial Statistics*. Bachelor's Thesis. Technical University of Munich, 2013.

8. Rogozin O.V. Metod nechetkogo vyvoda resheniya v zadache podbora programmnogo obespecheniya na osnove kachestvennykh kharakteristik etogo obespecheniya kak ob"ekta investitsii [Method of Fuzzy Inference Solutions in the Problem of Selection of Software Based on the Qualitative Characteristics of the Software as an Object of Investment]. *Kachestvo. Innovatsii. Obrazovanie* — Quality. Innovations. Education, 2009, No. 2, pp. 45–54. [in Russian].

9. Bai Y., Wang D. Fundamentals of Fuzzy Logic Control — Fuzzy Sets, Fuzzy Rules and Defuzzifications. *Advanced Fuzzy Logic Technologies in Industrial Applications*, Springer, London, 2006. pp. 17–36.

СВЕДЕНИЯ ОБ АВТОРАХ
ABOUT THE AUTHORS

Харисов Мансур Нагимович, канд. техн. наук, доцент кафедры «Вычислительная математика и кибернетика», УГАТУ, г. Уфа, Российская Федерация

Mansur N. Kharisov, Candidate of Engineering Sciences, Assistant Professor of Computational Mathematics and Cybernetics Department, USATU, Ufa, Russian Federation

e-mail: kharisov-usatu@yandex.ru

Юнусова Элеонора Аликовна, магистрант кафедры «Разработка и эксплуатация нефтяных и газонефтяных месторождений», УГНТУ, г. Уфа, Российская Федерация

Eleonora A. Yunusova, Undergraduate Student of Development and Exploitation of Oil, and Oil and Gas Fields Department, USPTU, Ufa, Russian Federation

e-mail: elenalikovna@gmail.ru

Вагизов Азат Минзякиевич, аспирант кафедры «Разработка и эксплуатация нефтяных и газонефтяных месторождений», УГНТУ, г. Уфа, Российская Федерация

Azat M. Vagizov, Post-graduate Student of Development and Exploitation of Oil, and Oil and Gas Fields Department, USPTU, Ufa, Russian Federation

e-mail: vagizov-7@mail.ru

Майский Равиль Анварович, канд. техн. наук, доцент кафедры «Математика», УГНТУ, г. Уфа, Российская Федерация

Ravil A. Maiski, Candidate of Engineering Sciences, Assistant Professor of Mathematics Department, USPTU, Ufa, Russian Federation

e-mail: ravanmay@yandex.ru

Валиев Эмиль Рафаилович, студент кафедры «Вычислительная математика и кибернетика», УГАТУ, г. Уфа, Российская Федерация

Emil R. Valiev, Student of Computational Mathematics and Cybernetics Department, USATU, Ufa, Russian Federation

e-mail: vemiliev@yandex.ru