ISSN 2411-3336; e-ISSN 2541-9404

ЗАПИСКИ ГОРНОГО ИНСТИТУТА
Journal of Mining Institute
Сайт журнала: pmi.spmi.ru

Научная статья
Геотехнология и инженерная геология

Прогнозирование температуры образования газовых гидратов с использованием гибридного подхода на основе искусственного интеллекта и эмпирических данных

М.Риази[1], **Х.Мехраджу**[1], **М.-С.Дабири**[1], **Дж.Исмаилова**[2], **М.Риази**[3]✉
[1] *Университет им. Шахида Бахонара в Кермане, Керман, Иран*
[2] *Казахский национальный исследовательский технический университет им. К.И.Сатпаева, Алматы, Республика Казахстан*
[3] *Назарбаев Университет, Астана, Республика Казахстан*

***Как цитировать эту статью*:** Риази М., Мехраджу Х., Дабири М.-С., Исмаилова Дж., Риази М. Прогнозирование температуры образования газовых гидратов с использованием гибридного подхода на основе искусственного интеллекта и эмпирических данных // Записки Горного института. 2026. Т. 280. С. 145-162. EDN QIWCEK

Аннотация
Точное прогнозирование температуры образования газовых гидратов необходимо для обеспечения бесперебойной подачи при добыче природного газа и его подводной транспортировке. В данном исследовании использовали 459 экспериментальных точек данных, собранных из литературных источников и охватывающих широкий диапазон давлений – 330-68600 кПа, а также диапазон удельной плотности газа – 0,552-1,03. Эти данные послужили основой для разработки надежных инструментов прогнозирования с использованием всего двух практических входных переменных – давления и удельной плотности газа. Были реализованы два взаимодополняющих подхода к моделированию. Во-первых, проведен систематический сравнительный анализ нескольких передовых методов машинного обучения, включая многослойный перцептрон (MLP), обученный с применением алгоритмов Левенберга – Марквардта (LM), байесовской регуляризации (BR) и алгоритма масштабируемых сопряженных градиентов (SCG); сеть радиально базисных функций (RBF); CatBoost; XGBoost; метод опорных векторов с использованием метода наименьших квадратов (LSSVM). Во-вторых, сформулирована оптимизированная эмпирическая корреляция, коэффициенты которой определены с помощью алгоритма нелинейной оптимизации обобщенного приведенного градиента (GRG) для минимизации общей погрешности прогнозирования на всем экспериментальном наборе данных. Среди моделей машинного обучения CatBoost продемонстрировал лучшую прогностическую производительность, достигнув значения AAPRE 0,094 % и коэффициента R^2 0,996 на тестовом наборе данных. Разработанная корреляция на основе GRG также показала хорошее соответствие экспериментальным данным, одновременно представляя собой прозрачную и вычислительно эффективную альтернативу, подходящую для быстрых инженерных расчетов. Анализ чувствительности и значения рычага подтвердили надежность, статистическую достоверность и широкую применимость предложенных моделей в исследуемой области.

Ключевые слова
температура образования газовых гидратов; метод опорных векторов с использованием метода наименьших квадратов; машинное обучение; многослойный перцептрон; сеть радиально базисных функций; обобщенный приведенный градиент

Финансирование
Отсутствует.

Поступила: 09.02.2025 ***Принята: 25.06.2026*** ***Онлайн: 28.08.2026*** ***Опубликована: 10.09.2026***

Введение

В последние годы природный газ стал широко распространенным, ценным и экологически чистым источником энергии [1-4]. Предпринимаются значительные усилия по продвижению и расширению его использования в различных областях, что обусловлено не только его техническими и экономическими преимуществами, но и низким уровнем выбросов парниковых газов [2, 5]. Однако добыча природного газа часто ведется в нетрадиционных условиях, что создает многочисленные эксплуатационные проблемы. Добыча, а также сопутствующие промысловые операции, такие как глубоководное бурение, транспортировка по трубопроводам, хранение и сепарация, сопряжены с серьезными проблемами обеспечения бесперебойной подачи [6-9]. Среди них наиболее

Статья опубликована в открытом доступе по лицензии CC BY 4.0

критичным признано образование гидратов, приводящее к тяжелым последствиям, в том числе закупоркам трубопроводов и противовыбросовых превенторов, нестабильности работы, потенциальным нарушениям безопасности и целостности промысловых установок [10-13]. Газовые гидраты – это кристаллические твердые соединения с клатратной структурой. Они образуются, когда небольшие молекулы газа (так называемые гости), включая сероводород, диоксид углерода, метан, этан, пропан и азот, оказываются заключены в полостях, созданных молекулами воды («хозяином»). Обычно гидраты образуются в условиях высокого давления и низкой температуры [14-19]. Образование гидратов часто наблюдается из-за постоянного присутствия воды при добыче газа либо в результате неправильного проектирования и строительства трубопроводов.

В многофазных магистральных трубопроводах, поверхностных сооружениях и устьях скважин гидраты создают серьезную проблему для бесперебойной подачи. Эффективное и своевременное внедрение стратегий предотвращения образования гидратов имеет решающее значение – закупорки, вызванные образованием гидратов, могут остановить добычу на дни или даже месяцы. Такие перебои могут привести к существенным экономическим потерям, исчисляемым миллионами долларов, потраченным на системы сбора газа или магистральные газопроводы высокого давления. E.Hammerschmidt [20] впервые сообщил о закупорке трубопроводов гидратами в 1934 г. С тех пор отложение гидратов при транспортировке и добыче природного газа стало важной областью исследований. Значительные достижения в изучении гидратов привлекли к себе большое внимание, подчеркнув важность понимания процесса их образования, определения ключевых параметров и разработки эффективных стратегий снижения рисков как для научных исследований, так и для промышленной практики [17, 21-26]. Учитывая междисциплинарную значимость образования гидратов, растет интерес к точному определению условий их формирования [27], среди которых температура образования – один из наиболее важных параметров [28, 29]. Хотя экспериментальные методы дают надежные прогнозы температуры образования гидратов, они зачастую дорогостоящи и требуют много времени. Поэтому исследователи ищут более экономичные прогнозирующие подходы, включая эмпирические корреляции, методы на основе диаграмм, а также теоретические модели, основанные на различных уравнениях состояния [27-29].

Хотя самая ранняя модель для прогнозирования температуры образования гидратов – корреляция, предложенная E.Hammerschmidt в 1934 г. [20], – связывала температуру образования гидратов исключительно с давлением, последующие и более совершенные корреляции включали в качестве важнейших входных параметров давление и состав газа [30-35]. Опираясь на эти разработки, ряд исследователей также применил методы машинного обучения для прогнозирования температуры образования газовых гидратов, что обеспечило повышенную точность и эффективность прогнозов по сравнению с традиционными методами.

A.A.Elgibaly и A.M.Elkamel [36] разработали сложную модель искусственной нейросети (ANN) для прогнозирования условий образования гидратов для различных чистых газов, газовых смесей и ингибиторов. В исследовании были рассмотрены несколько архитектур нейросетей для выявления наиболее точной конфигурации. Согласно выводам, сеть с одним скрытым слоем, содержащим 25 нейронов, дала достаточно точные результаты для модели, основанной на удельной плотности. Для композиционных моделей сети с 40 или 50 нейронами в скрытом слое обеспечили наименьшие погрешности прогнозирования. Кроме того, они представили новую модель нейросети, основанную на удельной плотности газа, с использованием обновленных экспериментальных данных об условиях образования гидратов. Эта модель продемонстрировала более высокую точность по сравнению с традиционными графическими методами изучения удельной плотности и эмпирическими корреляциями. Однако для удельных плотностей газа в диапазоне 0,8-1,0 ее применимость была ограничена нехваткой экспериментальных данных. Для удельных плотностей свыше 1,0 модель давала неоднозначные результаты – вероятно, из-за недостаточного объема данных и ограничений, связанных с использованием только удельной плотности газа в качестве входного параметра.

G.Zahedi et al. [37] применили подход на основе нейросети для прогнозирования температуры образования гидратов и сообщили о высокой прогностической эффективности модели. Результаты показали, что модель нейросети обеспечивает более высокую точность по сравнению с обычными методами прогнозирования.

Статья опубликована в открытом доступе по лицензии CC BY 4.0

S.H.Riazi et al. [38], используя экспериментальные данные E.D.Sloan и C.A.Koh [15], разработали новую корреляцию с помощью методов подбора кривых в MATLAB. Корреляция была дополнительно оптимизирована с применением генетического алгоритма (GA), оптимизации роем частиц (PSO) и алгоритма империалистической конкуренции (ICA) для повышения ее точности. По сравнению с предыдущими моделями предложенная корреляция показала наименьшую погрешность и наивысшую общую точность.

S.M.Hesami et al. [39] представили два усовершенствованных метода на основе нейросети, оптимизированных с помощью ICA, для прогнозирования условий образования газовых гидратов. Модель ICA-ANN была разработана с использованием эмпирических данных из литературы и оценена в сравнении с обычной моделью нейросети, а также с точной термодинамической моделью в различных рабочих условиях. Неожиданно модель ICA-ANN показала худшие результаты по сравнению с традиционным подходом нейросети.

M.N.Amar [40] предложил надежную и удобную корреляцию для оценки температуры образования гидратов различных типов природного газа, включая кислый, кислотный и сладкий газы. Эта корреляция была разработана с применением программирования с экспрессией генов (GEP) – метода машинного обучения. Результаты показали, что модель на основе GEP прогнозирует температуру образования гидратов с высокой точностью – средняя абсолютная ошибка (AAPRE) – 0,1397 %.

M.Mesbah et al. [41] разработали универсальную и удобную для пользователя корреляцию на основе гибридного метода группового учета аргументов (GMDH) для прогнозирования температуры образования гидратов, предназначенную для широкого спектра природных газовых смесей, включая сладкие и кислые газы. По сравнению с традиционными корреляциями эта модель продемонстрировала превосходную статистическую эффективность. Примечательно, что она показала многообещающие результаты для кислых газов с высокими концентрациями H_2S, превосходя существующие термодинамические модели и эмпирические корреляции. Для выявления сомнительных точек данных и оценки надежности модели был применен статистический метод рычага. Анализ подтвердил, что почти все точки данных попадают в допустимый диапазон применимости, что подтверждает статистическую устойчивость предложенной гибридной модели GMDH.

M.Hosseini и Y.Leonenko [42] предложили модели на основе машинного обучения для прогнозирования температуры образования гидратов метана в широком диапазоне концентраций рассола. Их результаты показали, что особо случайные решающие деревья демонстрируют высокую прогностическую эффективность с корнем квадратным из среднеквадратичной ошибки (RMSE) на тестовой выборке 0,6248.

В 2024 г. M.Karaköse и Ö.Yücel [43] разработали усовершенствованные модели регрессии в машинном обучении для оценки температуры образования гидратов на основе химического состава сладких газовых смесей. Применяемые методы включали линейную регрессию, регрессию с использованием дерева решений (DT), регрессию случайного леса (RF), обобщенные аддитивные модели и нейросети. Ансамблевые методы, особенно метод случайного леса, показали превосходную эффективность, достигнув значений R^2 и скорректированного $R^2 = 0{,}998$. Эти результаты подчеркивают потенциал ансамблевых методов обучения для повышения безопасности и эффективности систем транспортировки и переработки газа.

N.Khan et al. [44] оценили три алгоритма машинного обучения – регрессию опорных векторов (SVR), нейросети и DT – для прогнозирования температуры образования гидратов. Их результаты показали следующие значения R^2/RMSE: (0,9999; 0,0631) для SVR, (0,9986; 0,5011) для нейросети и (0,9278; 3,5606) для DT. На основании этих данных видно, что SVR продемонстрировала наивысшую прогностическую точность, за ней следовала нейросеть, а метод DT показал сравнительно низкую эффективность.

W.Aleem et al. [45] представили всесторонний и новый сравнительный анализ девяти различных моделей машинного обучения для точного прогнозирования температуры образования гидратов метана в присутствии рассолов. В исследованиях применялись основные алгоритмы машинного обучения, включая множественную линейную регрессию (MLR), долговременную краткосрочную память (LSTM), сеть радиально-базисных функций (RBF), метод опорных векторов (SVM), нейросети, регрессию гауссовского процесса (GPR), метод случайного леса (RF) и метод K-ближайших соседей (KNN).

Статья опубликована в открытом доступе по лицензии CC BY 4.0

Точность модели была проверена на большом наборе данных, включающем 1000 точек с широким диапазоном концентраций солей. Экспериментальные результаты показали, что алгоритм KNN быстро сходится, точен и стабилен на всем диапазоне точек данных, достигая значения $R^2 = 0{,}975$ и AAPRE = 0,385 %.

M.Khafaie et al. [46] в своем исследовании разработали и оценили две модели машинного обучения, включая нейросети и гибридную оптимизацию роем частиц-нейросеть (PSO-ANN). Подход был сосредоточен на прогнозировании температуры образования гидратов с использованием ключевых входных параметров – мольной доли компонентов газа и давления в системе. Был использован набор данных 1841 экспериментального образца. Модель нейросети выдала среднеквадратичную ошибку СКО = 0,0003533 и R^2 = 0,9976. Модель PSO-ANN выдала СКО = 0,0003298 и R^2 = 0,9899. Обе модели продемонстрировали высокую прогностическую точность, причем PSO-ANN имела небольшое преимущество по СКО. Это исследование подчеркнуло эффективность оптимизированного подхода машинного обучения для точного и обобщаемого прогнозирования температуры газовых гидратов.

M.R.Atta et al. [47] провели систематический обзор в соответствии с рекомендациями PRISMA и библиометрическое картирование для анализа исследований по искусственному интеллекту и машинному обучению для прогнозирования термодинамических условий газовых гидратов, опубликованных в период с 2016 по 2025 гг. Анализ показал преобладающую ориентацию на бинарную систему вода – газ (87,5 %) и сильную зависимость от экспериментальных наборов данных, взятых из литературы (73,9 %). Наиболее распространенными алгоритмами были нейросети, SVR и RF, а основными метриками оценки – R^2 и СКО. Значительным ограничением стало отсутствие количественной оценки неопределенности и недостаточная представленность наборов данных, полученных в полевых условиях. Обзор выявил возможности для предстоящей работы по оценке устойчивости моделей, расширению выходных данных моделей для включения кинетики и морфологии гидратов. Несмотря на обширные экспериментальные и численные исследования температуры образования гидратов сохраняется потребность в комплексных исследованиях, при которых одновременно разрабатываются точные, устойчивые и практически применимые модели машинного обучения и упрощенные корреляции.

В данном исследовании для разработки передовых прогностических моделей оценки температуры образования газовых гидратов использовался комплексный набор данных, включающий 459 точек. В рамках первого подхода были реализованы несколько надежных методов машинного обучения, в том числе CatBoost, XGBoost, метод опорных векторов с использованием метода наименьших квадратов (LSSVM), сеть радиально-базисных функций (RBF), многослойный перцептрон (MLP). Модели обучались с использованием алгоритмов Левенберга – Марквардта (LM), байесовской регуляризации (BR) и алгоритма масштабируемых сопряженных градиентов (SCG). Во втором подходе была разработана эмпирическая корреляция с применением метода оптимизации обобщенного приведенного градиента (GRG), учитывающего изменения давления и удельной плотности газа. Цель исследования состояла в том, чтобы вывести новую корреляцию, способную обеспечить большую точность и производительность по сравнению с существующими моделями. Для оценки влияния входных параметров на температуру образования газовых гидратов был проведен анализ коэффициентов релевантности. Кроме того, для статистического анализа применялся метод рычага – для выявления потенциальных выбросов и определения области применимости разработанных моделей.

Данные

Обобщающая способность и надежность разработанных моделей в значительной степени зависят от разнообразия и объема входных данных, использованных при разработке моделей. В настоящем исследовании для разработки точных моделей прогноза температуры образования газовых гидратов был составлен комплексный банк данных, включающий 459 экспериментальных точек данных и охватывающий широкий диапазон рабочих условий [48-55]. В табл.1 представлена статистическая сводка входных переменных, в том числе удельной плотности газа и давления, а также выходной переменной (экспериментальной температуры). Полный набор данных, использованный в данном исследовании, приведен в приложении https://pmi.spmi.ru/pmi/article/download/16663/16705.

Статья опубликована в открытом доступе по лицензии CC BY 4.0

Таблица 1

Статистическое описание используемой базы данных [48-56]

Характеристики	Удельная плотность газа	Давление, кПа	Экспериментальная температура, К
Максимум	1,03	68600	304,483
Минимум	0,552	330	272,04
Асимметрия	1,555	3,474	0,265
Эксцесс	2,338	14,396	–0,809
Мода	0,629	20684,272	283,15
Медиана	0,651	3164,999	284,349
Среднее значение	0,685	6722,435	284,956
Стандартное отклонение	0,097	10 136,045	7,52

Как показано в табл.1, удельная плотность газа варьируется от 0,552 до 1,03. Давление изменяется в диапазоне от 330 до 68600 кПа при среднем значении 6722,435 кПа. Экспериментальные температуры в банке данных находятся в пределах от 272,04 до 304,483 К со средним значением 284,956 К.

Методы

В данном исследовании для прогнозирования температуры образования газовых гидратов были использованы несколько методов машинного обучения.

Метод опорных векторов с использованием метода наименьших квадратов (LSSVM). Алгоритм LSSVM широко применяется для регрессионного анализа, распознавания образов и классификации данных. Он сохраняет основные характеристики традиционного метода SVM, предлагая при этом упрощенную формулировку, которая снижает вычислительную сложность и время обучения. Подробную информацию о теории SVM можно найти в источнике [57]. Модель LSSVM была впервые представлена J.A.Suykens и J.Vandewalle [58, 59]. Функция затрат этого метода аналогична функции затрат SVM и выражается следующим образом:

$$g(y) = O^T \varphi(y_i) + bi,$$

где bi – члены смещения; O^T – транспонированный вектор выходного слоя; $\varphi(y_i)$ – функция ядра; y_i – входной вектор функции.

Входная матрица имеет размеры $M \times m$, где m и M – количество входных параметров и общее количество точек данных соответственно. Путем минимизации функции затрат можно определить член смещения и вектор выходного слоя [60]:

$$\mathrm{CF} = \frac{1}{2} O^T O + \frac{1}{2} \gamma \sum_{i=1}^{num} s_i^2.$$

Функция затрат [60]:

$$x_i = O^T \varphi(y_i) + bi + s_i^2, \quad i = 1,2,3,...,M,$$

где s_i – ошибка регрессии; γ – регулируемый параметр в LSSVM.

Лагранжева форма задачи [60]:

$$K(O, s, q, bi) = \frac{1}{2} O^T O + \frac{1}{2} \gamma \sum_{i=1}^{num} s_i^2 - \sum_{i=1}^{num} q_i \left(O^T \varphi(y_i) + bi + s_i - x_k \right),$$

где q_i – множитель Лагранжа.

Взяв частные производные Лагранжа по bi, q_i, s_i и O, можно получить неизвестные параметры [60]:

Статья опубликована в открытом доступе по лицензии CC BY 4.0

$$\begin{cases} \dfrac{\partial K}{\partial O} = 0 \to O = \sum_{i=1}^{num} q_i \varphi(y_i); \\ \dfrac{\partial K}{\partial bi} = 0 \to \sum_{i=1}^{num} q_i = 0; \\ \dfrac{\partial K}{\partial s_i} = 0 \to q_i = \gamma s_i, \ i = 1,2,3,...,M; \\ \dfrac{\partial K}{\partial q_i} = 0 \to O^T \varphi(y_i) + bi + s_i - x_k = 0, \ i = 1,2,3,...,M. \end{cases}$$

Решение этой системы уравнений дает необходимые переменные в рамках LSSVM. Следующие уравнения представляют три традиционно используемые ядерные функции в литературе:

• ядро сети радиально-базисной функции – $\varphi(y, y_i) = \exp\left(\dfrac{-\|y_i - y\|^2}{\sigma^2}\right)$;

• линейное ядро – $\varphi(y, y_i) = y_i^T y$;

• полиномиальное ядро – $\varphi(y, y_i) = \left(\tau + y_i^T y\right)^n$.

В LSSVM σ^2 является регулируемым параметром, который можно оптимизировать путем минимизации целевой функции, такой как среднеквадратичная ошибка. При использовании следующего уравнения разница между прогнозируемыми выходными данными разработанной модели и соответствующими экспериментальными значениями минимизируется следующим образом:

$$\text{СКО} = \frac{\sum_{j=1}^{m}\left(R_{exp_j} - R_{\text{rep/pred}_j}\right)^2}{m}.$$

Сеть радиально-базисных функций (RBF). Сеть RBF – это хорошо зарекомендовавший себя тип прямой искусственной нейросети, широко используемой в качестве надежного подхода к моделированию для решения сложных задач [61, 62]. Нейросеть RBF обычно состоит из трех слоев – входного, скрытого и выходного. В скрытом слое каждый нейрон связан с центром и заранее определенным радиусом, задающим конкретную область во входном пространстве. В каждом нейроне вычисляется расстояние между входным вектором и соответствующим центром [63]. Для этого обычно применяются различные функции активации, такие как аппроксимация сплайном типа «тонких пластин», гауссова функция, кубическая функция, логистическая функция, квадратичная функция [64]. Евклидово расстояние между входным вектором и центром каждого нейрона вычисляется с помощью уравнения [65]

$$r_i = \left[\sum_{j=1}^{l}\left(y_i - Ce_{ji}\right)^2\right]^{0,5}, \ i = 1,...,M,$$

где индексы i, l – количество векторов центра и количество входов соответственно; y_i, r_i – входной вектор и евклидово расстояние от центра нейрона; Ce_{ji} – центр кластера.

Выход модели RBF [66], более подробная информация об этой модели нейронной сети доступна в источниках [67-69]:

$$x_c(y) = \sum_{i=1}^{M} cw_i \varphi_{ci}\left(\|y_c - Ce_i\|\right), \quad i = 1,...,M, \ c = 1,...,N,$$

где cw_i – вес связи; M – количество нейронов скрытого слоя; Ce_i – вектор центра; φ – функция активации.

Статья опубликована в открытом доступе по лицензии CC BY 4.0

Многослойный перцептрон (MLP) – это широко используемая прямая искусственная нейронная сеть [65], состоящая из входного слоя, одного или нескольких скрытых слоев и выходного слоя. Входные переменные подаются на входной слой, а скрытые слои, обычно использующие нелинейные функции активации, такие как tansig и logsig, преобразуют эти входные данные в представления более высокого уровня [70, 71]. Выходной слой, как правило, использует линейную функцию активации purelin для формирования окончательных прогнозов. Обучение MLP выполняется с помощью алгоритма обратного распространения ошибки, в ходе которого веса сети и смещения итеративно обновляются для минимизации ошибки между прогнозируемыми и целевыми выходными значениями. Хотя для выявления взаимосвязей между входными и выходными данными зачастую достаточно одного скрытого слоя, при моделировании более сложных систем может потребоваться добавление дополнительных слоев.

Метод обобщенного приведенного градиента (GRG) широко используется для оптимизации с множеством переменных. Этот подход основан на концепции приведенных градиентов и способен работать как с линейными, так и с нелинейными ограничениями, обеспечивая выполнение активных ограничений на протяжении всего процесса решения. Путем аппроксимации градиента в заданной точке z метод GRG одновременно учитывает градиенты целевой функции и уравнения ограничений. В результате ограничения включаются в градиенты целевой функции, сужая поиск допустимыми направлениями и эффективно сокращая пространство решений. В данном исследовании $f(z)$ представляет целевую функцию для ограничения $h(z)$ [72].

Минимизирует $f(z) = z;$ подвергается минимизации $h_k(z) = 0.$ Формула GRG выражается следующим образом:

$$\frac{df}{dz_k} = \nabla z_k^t f - \nabla z_i^t f \left(\frac{dh}{dz_i} \right) \frac{dh}{dz_k}.$$

Необходимым условием минимизации $f(y)$ является то, что ее первая производная равна нулю, т.е. $df(y) = 0$.

Алгоритм оптимизации. Эффективность моделей нейронных сетей во многом зависит от выбора подходящих алгоритмов обучения, используемых для оптимизации их весов и смещений. В данном исследовании для обучения модели MLP были применены три алгоритма оптимизации. Алгоритм Левенберга – Марквардта широко известен благодаря своей способности повышать производительность MLP за счет эффективного решения нелинейных задач наименьших квадратов [73, 74]. Алгоритм байесовской регуляризации (BR) предоставляет надежную альтернативу – он корректирует веса и смещения сети в рамках обучающей схемы LM, применяя при этом теорему Байеса для обновления их распределений [75, 76]. Алгоритм масштабируемых сопряженных градиентов (SCG) устраняет необходимость в вычислительно затратных линейных поисках, тем самым обеспечивая более быстрый и эффективный процесс обучения по сравнению с другими методами, основанными на градиенте.

Результаты и обсуждение

В данном исследовании моделирование температуры образования газовых гидратов проводилось в широком диапазоне рабочих условий. В интеллектуальных моделях в качестве входных переменных использовались давление и удельная плотность газа, а выходной переменной служила температура образования газовых гидратов. Комплексный банк данных, содержащий 459 экспериментальных точек, был случайным образом разделен на два подмножества для разработки моделей:

• Обучающая выборка использовалась для разработки модели, ее оптимизации и повышения производительности. В процессе обучения модель устанавливает взаимосвязь между входными параметрами и выходной переменной путем итеративной корректировки смещений и весов для минимизации отклонения между экспериментальными и прогнозируемыми значениями.

• Тестовая выборка. Это подмножество было зарезервировано для оценки прогностических возможностей модели и проверки ее точности.

Для всех моделей исследования 80 % данных были выделены для этапа обучения, а оставшиеся 20 % использовались для тестирования. Важно отметить, что высокая производительность на

Статья опубликована в открытом доступе по лицензии CC BY 4.0

этапе обучения не гарантирует надежности модели. Поэтому для подтверждения точности и устойчивости разработанных моделей необходимо тестирование на зарезервированных данных.

Для моделирования температуры образования газовых гидратов были использованы передовые алгоритмы машинного обучения, включая CatBoost, XGBoost, MLP, RBF и LSSVM. Дополнительно для обучения сети MLP были применены три метода оптимизации – LM, SCG и BR. Для оценки производительности разработанных моделей проведены различные статистические и графические анализы ошибок. Также выполнен анализ чувствительности – чтобы определить влияние входных параметров на выходные данные моделей.

Перед разработкой моделей все входные и выходные переменные были нормализованы с помощью масштабирования min-max до диапазона от −1 до 1, что позволило повысить численную устойчивость и скорость сходимости. Прогнозируемые выходные данные впоследствии были преобразованы обратно в исходный масштаб. На начальном этапе была разработана корреляция на основе входных параметров для оценки температуры образования газовых гидратов.

Разработка корреляции. Для алгоритма GRG была установлена корреляция, продемонстрировавшая его способность прогнозировать целевые параметры с высокой точностью. В результате было разработано понятное уравнение, включающее два входных параметра и пять дополнительных коэффициентов:

$$T = aSG^{b}\left(\frac{\log(P)}{1{,}5}\right)^{C} - dSG\sqrt{\frac{\log(P)}{3{,}25}} + e, \qquad (1)$$

где T – температура, К; P – давление, кПа; SG – удельная плотность газа; a, b, c, d, e – константы корреляции, a = 162,4292; b = 0,244852; c = 0,36583; d = 50,60196; e = 116,777.

Корреляция, представленная в уравнении (1), может быть использована для оценки температуры образования газовых гидратов как функции давления и удельной плотности газа.

Оценка моделей. В данном исследовании предложенные модели оценивались с помощью статистического анализа ошибок, а их производительность – с использованием нескольких метрик, в том числе коэффициента детерминации R^2, RMSE и AAPRE:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\left(T_{exp,i} - T_{\mathrm{pred},i}\right)^2}{N}};$$

$$\mathrm{AAPRE} = \frac{100}{N}\sum_{i=1}^{N}\frac{\left|T_{exp,i} - T_{\mathrm{pred},i}\right|}{T_{exp,i}};$$

$$\mathrm{R}^2 = 1 - \frac{\sum_{i=1}^{N}\left(T_{exp,i} - T_{\mathrm{pred},i}\right)^2}{\sum_{i=1}^{N}\left(T_{exp,i} - \bar{T}\right)^2},$$

где N − общее количество точек данных; $T_{exp,i}$ − экспериментально измеренная температура; $T_{\mathrm{pred},i}$ − прогнозируемая температура.

Статистическая оценка. Для выявления наиболее точного варианта разработанные модели были систематически оценены и сопоставлены с использованием статистических метрик, в том числе R^2, AAPRE и RMSE. Согласно статистическим принципам, точность и достоверность модели повышаются по мере приближения значения R^2 к единице. При этом более низкие значения RMSE и AAPRE указывают на более высокую точность процесса моделирования. Значения параметров представлены в табл.2.

Результаты показали, что среди протестированных моделей наивысшую точность по уровню прогностической эффективности продемонстрировала CatBoost, за ней следуют XGBoost, MLP-LM, MLP-BR, MLP-SCG и LSSVM. Как показано в табл.2, значения AAPRE 0,126 % для тестовой

Статья опубликована в открытом доступе по лицензии CC BY 4.0

выборки, 0,086 % для обучающей выборки и 0,094 % для всего набора данных указывают на то, что подход CatBoost обеспечивает высокоточные прогнозы температуры образования газовых гидратов. Кроме того, этот метод выдает заметное общее значение R^2, равное 0,996. Результаты также позволили установить следующую иерархию точности прогнозирования температуры образования газовых гидратов: CatBoost превосходит XGBoost, за которым следуют MLP-LM, MLP-BR, MLP-SCG, RBF и, наконец, LSSVM. Переобучение оценивалось путем сравнения показателей на обучающей и тестовой выборках; минимальная разница в значениях R^2 и AAPRE демонстрирует, что разработанные модели хорошо обобщают данные на всем наборе.

Таблица 2

Статистические показатели разработанной модели для экспериментальной температуры

Ошибка	MLP			RBF	LSSVM	CatBoost	XGBoost	Разработанная корреляция
	LM	BR	SCG					
Обучение								
AAPRE, %	0,187	0,21	0,271	0,308	0,311	0,086	0,103	0,35
RMSE	0,753	0,86	1,037	1,12	1,14	0,414	0,405	1,44
R^2	0,99	0,987	0,981	0,978	0,977	0,997	0,997	0,96
Тестирование								
AAPRE, %	0,18	0,208	0,277	0,309	0,31	0,126	0,249	0,35
RMSE	0,756	0,837	1,102	1,155	1,13	0,494	1,198	1,44
R^2	0,99	0,987	0,978	0,975	0,978	0,996	0,974	0,96
Итого								
AAPRE, %	0,186	0,209	0,272	0,308	0,311	0,094	0,132	0,35
RMSE	0,753	0,855	1,05	1,127	1,138	0,431	0,647	1,43
R^2	0,989	0,987	0,98	0,977	0,977	0,996	0,992	0,96

Гиперпараметры всех моделей машинного обучения, использованных в данном исследовании, были тщательно подобраны для обеспечения оптимальной производительности:

• MLP: скрытые слои – [5,5]; функция переноса – tansig, logsig, purelin; функция обучения – LM, BR, SCG;

• RBF: степень разброса – 2,78; максимальное число нейронов – 100;

• LSSVM: σ^2 – 1,3130; тип ядра – ядро RBF; γ – 589,5852;

• CatBoost: colsample bytree – 0,9; гамма – 0,01; скорость обучения – 0,05; максимальная глубина – 5; минимальный вес дочернего узла – 5; n_estimator – 1000; reg_lambda – 10; subsample – 0,9;

• XGBoost: скорость обучения – 0,05; l2 leaf reg – 5; итерации – 1500; глубина – 6.

В дополнение к моделям машинного обучения, разработанным в рамках настоящего исследования, была предложена корреляция для оценки температуры образования газовых гидратов. Эффективность предложенной корреляции в сравнении с существующими моделями представлена в табл.3.

Таблица 3

Статистический анализ предложенной корреляции относительно существующих моделей

Модель	AAPRE, %	RMSE	R^2
Новая корреляция	0,3536	1,4287	0,9645
B.Towler, S.Mokhatab [33]	71,4943	203,552	0,9565
M.Safamirzaei et al. [35]	0,4071	1,5506	0,9607
E.Hammerschmidt et al. [20]	66,6774	190,9806	0,8660
S.Chavoshi et al. [77]	0,5292	1,9424	0,9455
M.Motiee et al. [31]	74,2416	211,3726	0,9546

Статья опубликована в открытом доступе по лицензии CC BY 4.0

Как показано в табл.3, разработанная в настоящем исследовании корреляция продемонстрировала более высокую эффективность по сравнению с существующими моделями: AAPRE – 0,3536 %, RMSE – 1,4287; R^2 – 0,9645.

Помимо статистической оценки, для изучения производительности моделей был проведен визуальный анализ с помощью графиков. На рис.1 представлены сводные графики (сопоставление прогнозируемых и экспериментальных значений) температуры образования газовых гидратов для обучающей и тестовой выборок. Эти графики позволяют визуально оценить точность моделей.

Сводный график – это тип диаграммы, иллюстрирующий взаимосвязь между прогнозируемыми и экспериментальными значениями. Как показано на рис.1, плотная группировка прогнозируемых точек данных вдоль линии равенства (под углом 45°) указывает на сильную корреляцию между наблюдаемыми и прогнозируемыми значениями, подчеркивая впечатляющую точность модели CatBoost. Кроме того, наибольшая дисперсия данных наблюдается на сводном графике модели LSSVM; эта модель прогнозирует более низкую температуру образования газовых гидратов по сравнению с другими разработанными моделями. Это наблюдение согласуется со значениями ошибок, приведенными в табл.2.

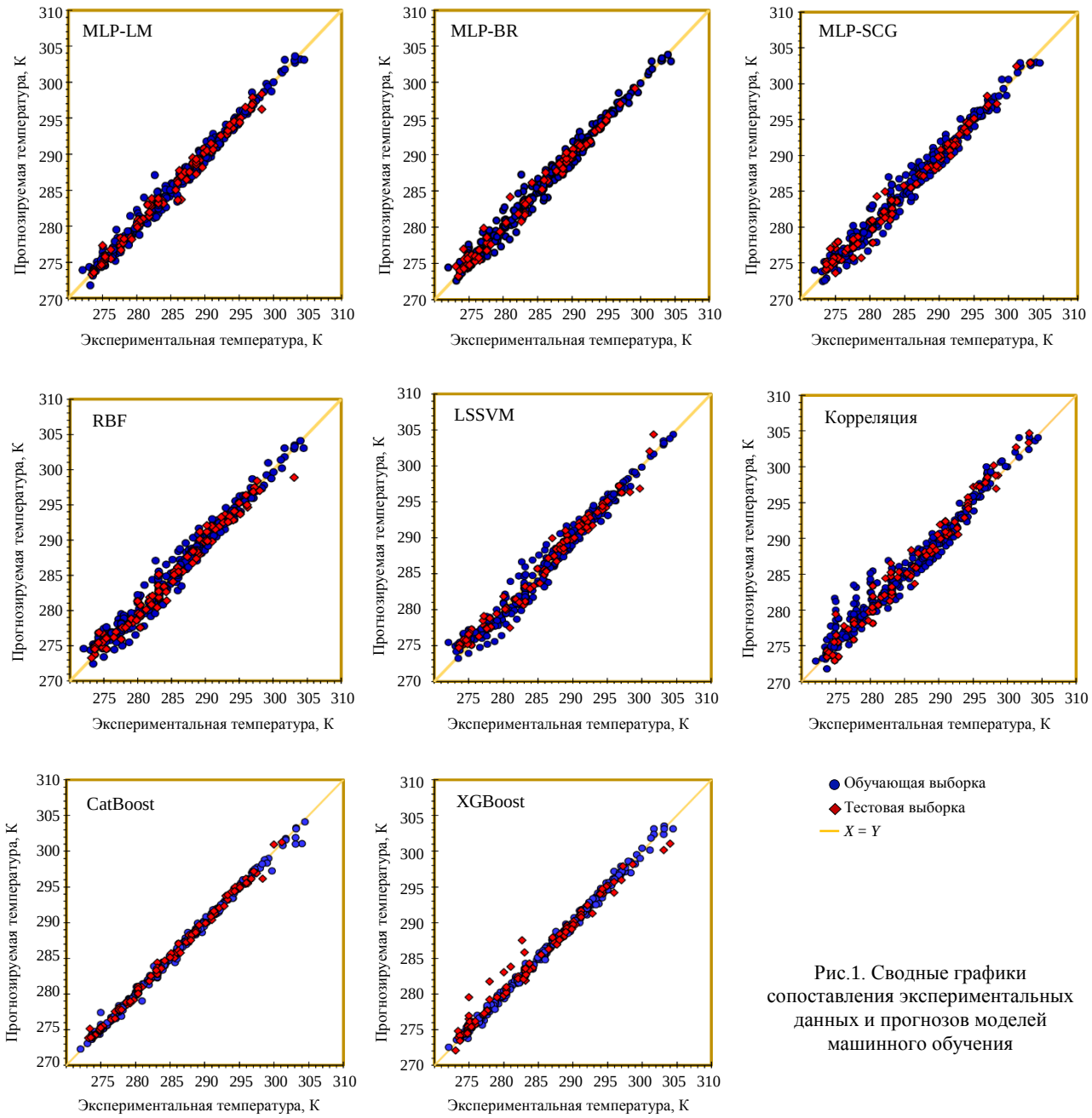


Рис.1. Сводные графики сопоставления экспериментальных данных и прогнозов моделей машинного обучения

Статья опубликована в открытом доступе по лицензии CC BY 4.0

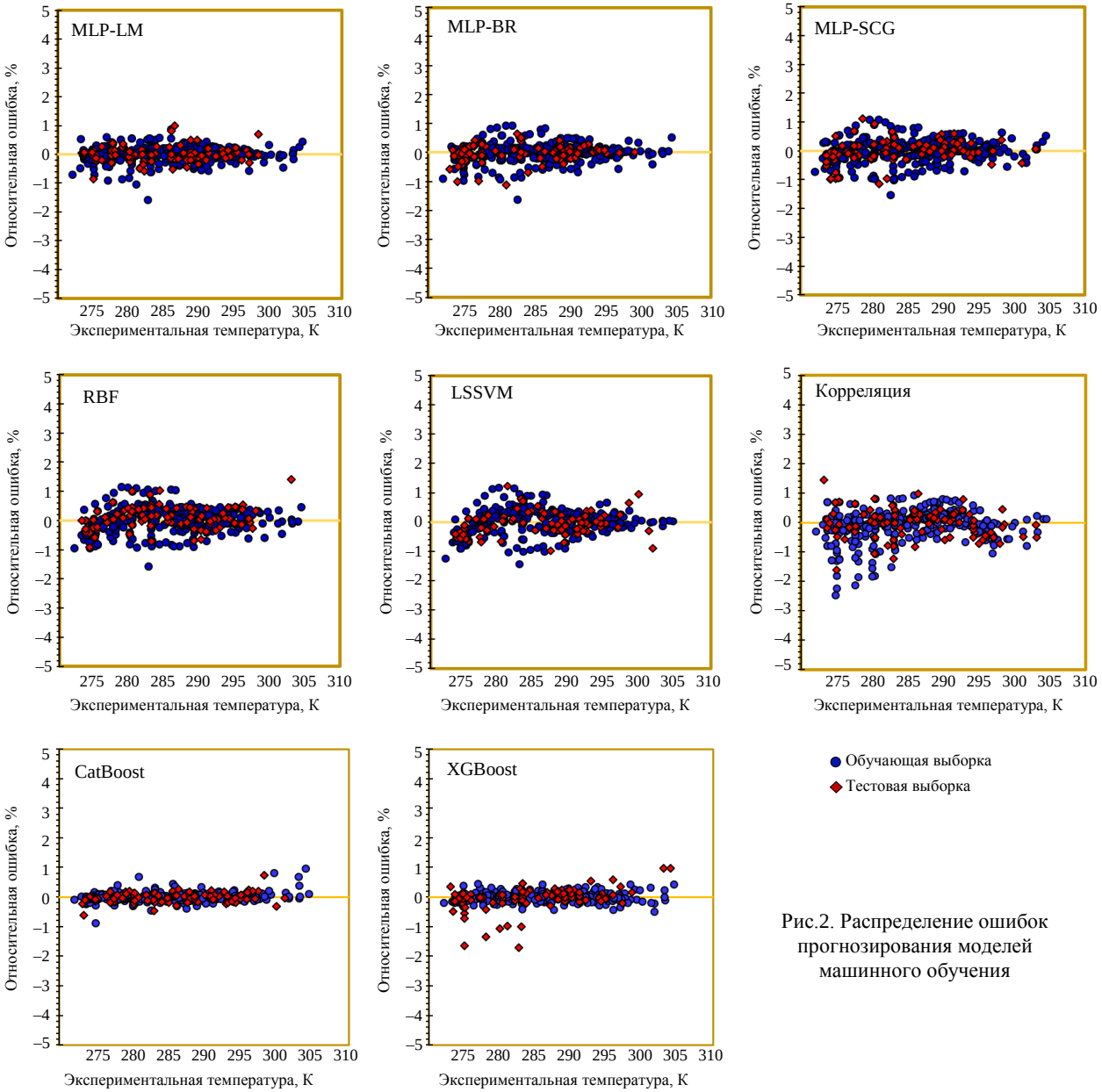


Рис.2. Распределение ошибок прогнозирования моделей машинного обучения

На рис.2 представлено распределение процентных значений относительной ошибки в зависимости от экспериментально измеренных температур образования газовых гидратов для всех моделей. При такой оценке ошибок наблюдаемые выходные данные сравниваются со значениями AAPRE.

Модель прогнозирования демонстрирует превосходную производительность при оценке температуры газовых гидратов, когда большинство точек данных имеют ошибки, близкие к нулю. Ошибки, близкие к нулю, у модели CatBoost подтвердили ее точность и надежность. Однако другие модели, в частности XGBoost, также могут обеспечивать точные прогнозы с минимальными ошибками при оценке температуры образования газовых гидратов.

Производительность разработанных моделей, оцененная с использованием параметра R^2, показана на рис.3.

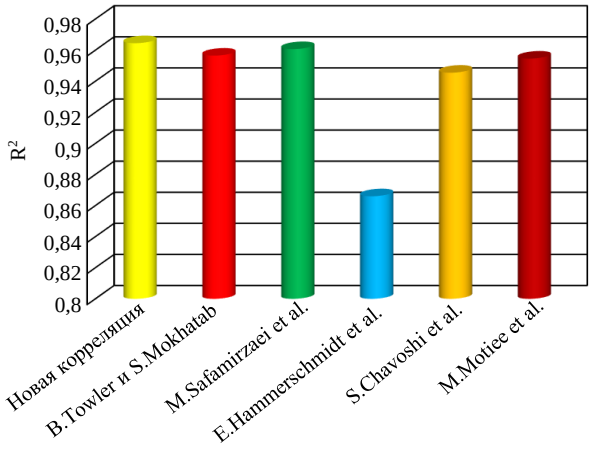


Рис.3. Сравнение значений R^2 предложенных корреляционных моделей

Статья опубликована в открытом доступе по лицензии CC BY 4.0

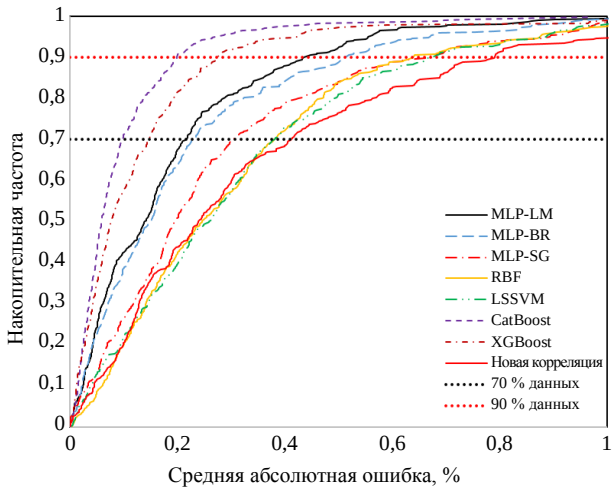


Рис.4. График накопленной частоты для нескольких моделей машинного обучения и новой корреляционной модели, предложенной в настоящем исследовании

Новая разработанная корреляционная модель продемонстрировала наивысшую производительность среди предложенных моделей, что показано на рис.3.

Следующим этапом визуальной оценки предложенных моделей машинного обучения для прогнозирования температуры образования газовых гидратов является анализ частоты накопленной ошибки. Этот анализ позволяет выявить большую долю точек данных с минимальными ошибками. На рис.4 представлены накопленные распределения частоты средней абсолютной ошибки для различных моделей, примененных ко всему набору данных.

На рис.4 темная горизонтальная пунктирная линия, соответствующая 70 % данных, указывает на значения средней абсолютной оценки, %: 0,1 – для модели CatBoost, 0,15 – для модели XGBoost, 0,2 – для модели MLP-LM, 0,22 – для модели MLP-BR, 0,29 – для модели MLP-SCG, 0,37 – для модели RBF, 0,38 – для модели LSSVM, 0,42 – для новой корреляционной модели. Эти результаты демонстрируют, что для 70 % набора данных ошибки прогнозирования находятся ниже соответствующих значений. Кроме того, примерно 90 % прогнозируемых значений модели CatBoost имеют среднюю абсолютную ошибку в процентах ниже 0,2 %, тогда как у остальных моделей этот порог превышен. Полученные результаты, подкрепленные статистическим и графическим анализом, демонстрируют высокую точность модели CatBoost в прогнозировании температуры образования газовых гидратов.

Анализ тренда для модели CatBoost и разработанной корреляции. Проведен анализ тренда, чтобы проверить, точно ли оптимальные модели, предложенные в настоящем исследовании (а именно – модель CatBoost и разработанная корреляция), отражают ожидаемые изменения температуры образования газовых гидратов в зависимости от выбранных независимых переменных. Рис.5 иллюстрирует взаимосвязь между температурой образования газовых гидратов и давлением при постоянной удельной плотности газа.

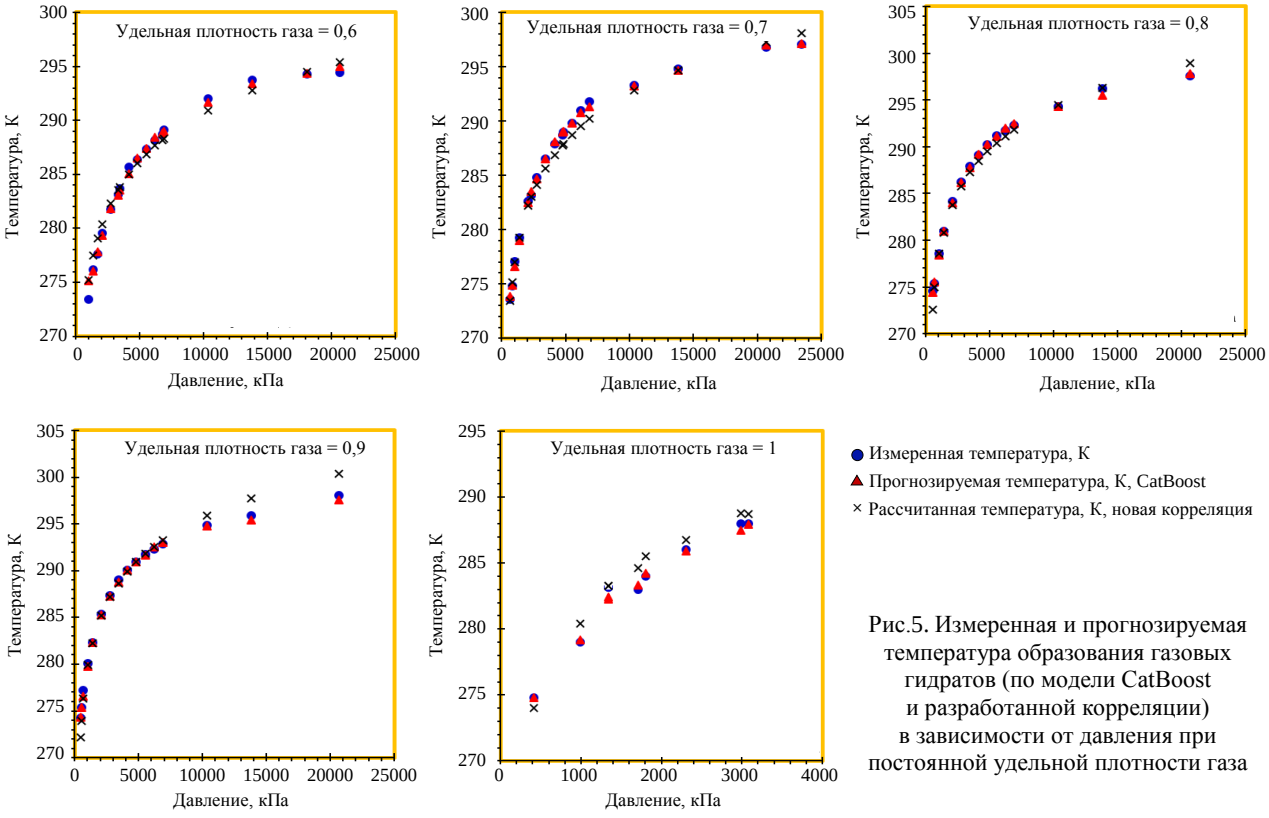


Рис.5. Измеренная и прогнозируемая температура образования газовых гидратов (по модели CatBoost и разработанной корреляции) в зависимости от давления при постоянной удельной плотности газа

Статья опубликована в открытом доступе по лицензии CC BY 4.0

Рис.5 дополнительно подтверждает точность разработанных моделей: для значений удельной плотности газа 0,6; 0,7; 0,8; 0,9 модель CatBoost и разработанная корреляция хорошо согласуются с экспериментальными данными. Однако при удельной плотности газа, равной единице, наблюдается небольшое расхождение между прогнозными моделями и экспериментальными данными. Кроме того, при высоких давлениях для всех значений удельной плотности газа точность прогнозирования моделей несколько снижается. Основная причина этих незначительных отклонений связана с ограниченным объемом данных для удельной плотности газа, равной единице, и для высоких давлений в рамках собранного набора данных. Кроме того, на расхождения в прогнозах разработанной корреляции могут влиять и экспериментальные погрешности. В целом, как видно из рис.5, прогнозы модели CatBoost и разработанной корреляции демонстрируют отличное соответствие экспериментальным данным в различных рабочих условиях. Также рис.5 показывает, что при постоянной удельной плотности газа повышение давления приводит к увеличению температуры образования газовых гидратов – эта тенденция точно отражена как моделью, так и корреляцией.

Анализ чувствительности. Проведен анализ чувствительности с использованием абсолютного коэффициента релевантности *r* для оценки влияния входных параметров на температуру образования газовых гидратов [67, 78, 79]. Входной параметр считается значимым, если его абсолютное значение релевантности велико. Данный коэффициент определяется следующим образом:

$$r = \frac{abs\left(\sum_{i=1}^{n}\left(I_{k,i} - \bar{I}_k\right)\left(O_i - \bar{O}\right)\right)}{\sqrt{\sum_{i=1}^{n}\left(I_{k,i} - \bar{I}_k\right)^2 \sum_{i=1}^{n}\left(O_i - \bar{O}\right)^2}},$$

где $I_{k,i}$, $\bar{I}_k$ – *k*-й входной параметр и среднее значение I_k соответственно; *i* относится либо к удельному давлению, либо к давлению; O_i, $\bar{O}$ – выходной параметр и *i*-е среднее значение O_i.

Параметр *r*, принимающий значения в диапазоне от –1 до 1, отражает корреляцию между зависимыми и независимыми переменными [80]. Положительное значение *r* указывает на прямую зависимость – при увеличении независимой переменной зависимая переменная также возрастает.

Результат анализа чувствительности входных переменных, влияющих на температуру образования газовых гидратов, выполненного с применением метода корреляции Пирсона: удельная плотность газа – 0,067; давление – 0,752. Входные переменные, такие как давление и удельная плотность газа, оказывают положительное влияние на целевое значение – увеличение этих параметров приводит к повышению температуры образования газовых гидратов. Среди них наибольшее влияние оказывает давление – его вклад в оценку температуры гидратообразования составляет 75,2 %. Температура газового гидрата также повышается с увеличением давления.

Подход с использованием метода рычага. Для оценки надежности предложенной модели и набора данных был применен метод рычага. Цель этого метода – выявить записи, демонстрирующие значительные расхождения в данных [81]. Одним из ключевых параметров настоящего метода являются стандартизованные остатки:

$$\mathrm{SR} = \frac{Z_i}{\left[\mathrm{MSE}\left(1 - H_{ii}\right)\right]^{0,5}},$$

где H_{ii} – значения «шляпы»; Z_i – значения ошибок; MSE – среднеквадратическая ошибка.

Еще одним ключевым параметром данного подхода является порог значения рычага

$$H^* = \frac{3\left(I+1\right)}{N},$$

где *N* – общее количество точек данных.

Последний важный параметр в этом методе – матрица «шляпы»

$$H = X\left(X^t X\right)^{-1} X^t,$$

где *X* – двумерная матрица; X^t – транспонированная матрица [81-83].

Статья опубликована в открытом доступе по лицензии CC BY 4.0

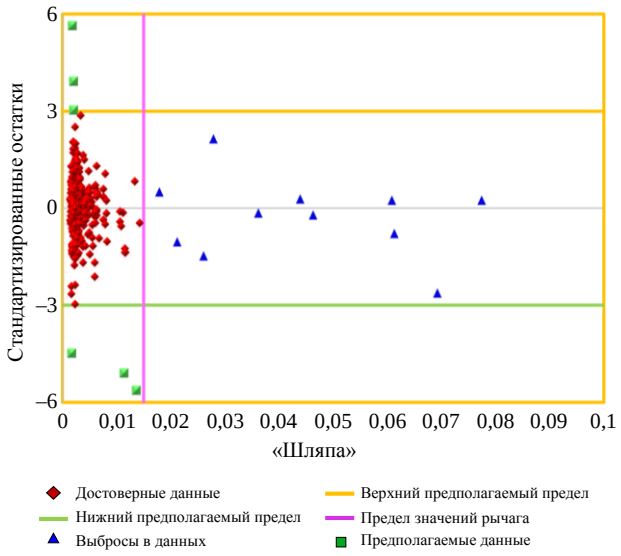


Рис.6. График Вильямса для полного набора данных, использованных в разработанной в представленном исследовании модели

График Вильямса демонстрирует корреляцию между стандартизованными остатками SR и значениями «шляпы», разделяя данные на четыре категории: точки плохого значения рычага, допустимые точки, точки хорошего значения рычага, подозрительные значения. Критерии надежной модели: большинство точек данных попадают в категорию допустимых точек; значения SR лежат в диапазоне от −3 до 3; значения «шляпы» не превышают порог значения рычага. Точки данных считаются точками с хорошим значением рычага, если их значения «шляпы» превышают порог, а значения SR остаются в диапазоне от −3 до 3, такие точки способствуют точным прогнозам [71]. И наоборот, точки со значением рычага, выходящие за пределы этого диапазона, считаются плохими средствами прогнозирования. Подозрительные значения – точки, у которых значения SR выходят за пределы диапазона −3...3, но значения «шляпы» остаются ниже порогового значения рычага. Для модели CatBoost был построен график зависимости стандартизованных остатков от значений «шляпы» с целью оценки температуры образования газовых гидратов (рис.6).

График Вильямса используется для оценки области применимости новой корреляции и выявления потенциальных выбросов. Как показано на рис.6, большинство точек данных попадают в диапазоны $0 \le H \le 0{,}015$ и $-3 \le \mathrm{SR} \le 3$, что указывает на высокое качество экспериментальных данных и высокую надежность разработанной модели. Точки данных с $\mathrm{SR} > 3$ или $\mathrm{SR} < -3$ считаются подозрительными, поскольку они предполагают высокий уровень неопределенности. Выбросы – это точки данных с $H > 0{,}015$ и $-3 < \mathrm{SR} < 3$, что означает, что они выходят за пределы допустимого диапазона предложенной модели. Всего в качестве подозрительных выбросов было выявлено шесть точек (1,3 % от общего объема набора данных), при этом в наборе данных по температуре присутствует 11 выбросов, что подтверждает надежность экспериментального набора данных, использованного для моделирования.

Ограничения модели, область применимости и направления дальнейших исследований. Настоящие модели разработаны с использованием экспериментальных данных, охватывающих диапазон давлений 330-68600 кПа, удельную плотность газа от 0,552 до 1,03, лабораторные температуры от 272,04 до 304 К. Поэтому предложенные модели рекомендуется использовать строго в указанных диапазонах. За пределами этих границ, в частности при давлениях ниже 330 кПа или выше 68600 кПа, температурах вне диапазона 272-305 К либо для значений удельной плотности газа, не представленных в обучающих данных, точность прогнозов может снижаться.

На рис.5 наблюдается небольшое расхождение между прогнозными моделями и экспериментальными данными при удельной плотности газа, равной единице. Кроме того, при высоких давлениях для всех значений удельной плотности газа точность прогнозирования моделей несколько снижается. Основная причина этих незначительных отклонений связана с ограниченным объемом данных для удельной плотности газа, равной единице, и для высоких давлений в рамках собранного набора данных. Кроме того, корреляция носит эмпирический характер и напрямую не учитывает такие факторы, как соленость, примеси или кинетические эффекты неравновесия. В подобных условиях эти факторы могут вносить дополнительную неопределенность в результаты.

Вывод

В настоящем исследовании для прогнозирования температуры образования газовых гидратов были использованы две различные методологии. Первый подход, основанный на данных, задействовал несколько передовых алгоритмов машинного обучения: многослойный перцептрон (MLP),

Статья опубликована в открытом доступе по лицензии CC BY 4.0

оптимизированный методами Левенберга – Маркуардта (LM), байесовской регуляризации (BR) и алгоритмом масштабируемых сопряженных градиентов (SCG); сеть радиально-базисных функций (RBF) и метод опорных векторов с применением метода наименьших квадратов (LSSVM). Второй метод заключался в разработке корреляции с применением техники оптимизации обобщенного приведенного градиента GRG, учитывающей изменения давления и удельной плотности газа. Модели были разработаны и проверены с использованием обширного набора данных, включающего 459 экспериментальных значений температуры, полученных из существующих исследований. Эффективность моделей тщательно оценивалась с помощью статистических метрик, в том числе средней абсолютной ошибки в процентах (AAPRE), корня квадратного из среднеквадратической ошибки (RMSE), коэффициента детерминации (R^2), а также посредством визуального анализа – с построением сводных графиков, графиков распределения ошибок и кумулятивных распределений частот.

Среди всех моделей наивысшую точность прогнозирования продемонстрировал алгоритм CatBoost: он достиг значения AAPRE 0,094 % и R^2 0,996. Остальные модели – XGBoost, MLP-LM, MLP-BR, MLP-SCG, RBF и LSSVM – показали значения AAPRE 0,103; 0,187; 0,209; 0,272; 0,308 и 0,311 % соответственно, что свидетельствует о высокой эффективности. Корреляция, полученная с помощью метода GRG, также превзошла существующие модели по точности прогнозирования.

Хотя разработанные модели демонстрируют высокую точность прогнозирования в рамках пространства параметров обучающих данных, их применимость по своей природе ограничена диапазоном входных параметров, использованных при разработке. Прогнозы за пределами этих установленных границ могут быть менее точными – это критически важно учитывать при практическом внедрении. Кроме того, на достоверность прогнозов могут влиять неустранимые неопределенности во входных данных, особенно для наиболее значимых параметров. Для более критической оценки точности помимо стандартных метрик мы проанализировали выбросы в прогнозах и отметили, что они преимущественно возникают на крайних значениях входных диапазонов или для менее распространенных точек данных. Это дает более детальное понимание границ эффективности модели. Сравнительный анализ с существующими моделями был усилен: четко обозначены сценарии, в которых наши модели показывают лучшие результаты, и обозначены отдельные случаи, где альтернативные подходы могут оказаться более подходящими. Таким образом обеспечивается сбалансированный взгляд относительно текущего состояния исследований в этой области. Анализ чувствительности показал, что давление оказывает прямое и положительное влияние на температуру образования газовых гидратов и является наиболее значимой входной переменной.

В будущих исследованиях можно расширить область применения, включив такие значимые факторы, как соленость, состав газа и наличие ингибиторов, которые не были учтены в данной работе из-за ограничений в данных. Такое расширение позволит дополнительно повысить обобщающую способность и надежность прогнозных моделей.

Доступ к данным

Полный набор данных, использованный для прогнозирования температуры образования газовых нитратов, доступен по ссылке:

Приложение, https://pmi.spmi.ru/pmi/article/download/16663/16705.

ЛИТЕРАТУРА

1. *Safamirzaei M.P.* Predict gas hydrate formation temperature with a simple correlation // Gas Processing Journal. 2015. July/August. P. 51-54.

2. *Esfahani S., Baselizadeh S., Hemmati-Sarapardeh A.* On determination of natural gas density: Least square support vector machine modeling approach // Journal of Natural Gas Science and Engineering. 2015. Vol. 22. P. 348-358. DOI: 10.1016/j.jngse.2014.12.003

3. *Inkong K., Rangsunvigit P., Kulprathipanja S., Linga P.* Effects of temperature and pressure on the methane hydrate formation with the presence of tetrahydrofuran (THF) as a promoter in an unstirred tank reactor // Fuel. 2019. Vol. 255. № 115705. DOI: 10.1016/j.fuel.2019.115705

4. *Tatar A., Barati-Harooni A., Najafi-Marghmaleki A. et al.* Prediction of moisture content of natural gases using a GA-RBF model // Journal of Molecular Liquids. 2016. Vol. 223. P. 994-999. DOI: 10.1016/j.molliq.2016.08.108

Статья опубликована в открытом доступе по лицензии CC BY 4.0

5. *Jingzhou Wang, Lu Zhang, Ran Li et al.* A simple generalized viscosity model for natural gas components and their mixtures based on free-volume theory // Journal of Molecular Liquids. 2020. Vol. 320. Part B. № 114483. DOI: 10.1016/j.molliq.2020.114483

6. *Jianbo Zhang, Zhiyuan Wang, Shun Liu et al.* Prediction of hydrate deposition in pipelines to improve gas transportation efficiency and safety // Applied Energy. 2019. Vol. 253. № 113521. DOI: 10.1016/j.apenergy.2019.113521

7. *Yan Wang, Koh C.A., White J. et al.* Hydrate formation management simulations with anti-agglomerants and thermodynamic inhibitors in a subsea tieback // Fuel. 2019. Vol. 252. P. 458-468. DOI: 10.1016/j.fuel.2019.04.146

8. *Zhiyuan Wang, Shikun Tong, Chao Wang et al.* Hydrate deposition prediction model for deep-water gas wells under shut-in conditions // Fuel. 2020. Vol. 275. № 117944. DOI: 10.1016/j.fuel.2020.117944

9. *Olajire A.A.* Flow assurance issues in deep-water gas well testing and mitigation strategies with respect to gas hydrates deposition in flowlines – A review // Journal of Molecular Liquids. 2020. Vol. 318. № 114203. DOI: 10.1016/j.molliq.2020.114203

10. *Chuanliang Yan, Yang Li, Yuanfang Cheng et al.* Sand production evaluation during gas production from natural gas hydrates // Journal of Natural Gas Science and Engineering. 2018. Vol. 57. P. 77-88. DOI: 10.1016/j.jngse.2018.07.006

11. *Kelland M.A., Iversen J.E.* Kinetic Hydrate Inhibition at Pressures up to 760 Bar in Deep Water Drilling Fluids // Energy & Fuels. 2010. Vol. 24. Iss. 5. P. 3003-3013. DOI: 10.1021/ef9016152

12. *Xin Zhao, Zhengsong Qiu, Zhen Zhang, Yubin Zhang.* Relationship between the gas hydrate suppression temperature and water activity in the presence of thermodynamic hydrate inhibitor // Fuel. 2020. Vol. 264. № 116776. DOI: 10.1016/j.fuel.2019.116776

13. *Yarveicy H., Ghiasi M.M.* Modeling of gas hydrate phase equilibria: Extremely randomized trees and LSSVM approaches // Journal of Molecular Liquids. 2017. Vol. 243. P. 533-541. DOI: 10.1016/j.molliq.2017.08.053

14. *Minwei Sun, Firoozabadi A.* Gas hydrate powder formation – Ultimate solution in natural gas flow assurance // Fuel. 2015. Vol. 146. P. 1-5. DOI: 10.1016/j.fuel.2014.12.078

15. *Sloan Jr. E.D., Koh C.A.* Clathrate Hydrates of Natural Gases. CRC press, 2007. 752 p. DOI: 10.1201/9781420008494

16. *Koh C.A., Westacott R.E., Zhang W. et al.* Mechanisms of gas hydrate formation and inhibition // Fluid Phase Equilibria. 2002. Vol. 194-197. P. 143-151. DOI: 10.1016/S0378-3812(01)00660-4

17. *Huangjing Zhao, Minwei Sun, Firoozabadi A.* Anti-agglomeration of natural gas hydrates in liquid condensate and crude oil at constant pressure conditions // Fuel. 2016. Vol. 180. P. 187-193. DOI: 10.1016/j.fuel.2016.03.029

18. *Li Wan, Na Zhang, De-Qing Liang.* Inhibition effects of polysaccharides for gas hydrate formation in methane–water system // Journal of Molecular Liquids. 2019. Vol. 292. № 111435. DOI: 10.1016/j.molliq.2019.111435

19. *Liang Yang, Zhenzhen Liu, Daoping Liu et al.* Enhanced natural gas hydrates formation in the suspension with metal particles and fibers // Journal of Molecular Liquids. 2020. Vol. 301. № 112410. DOI: 10.1016/j.molliq.2019.112410

20. *Hammerschmidt E.* Formation of Gas Hydrates in Natural Gas Transmission Lines // Industrial & Engineering Chemistry. 1934. Vol. 26. Iss. 8. P. 851-855. DOI: 10.1021/ie50296a010

21. *Kondori J., Zendehboudi S., James L.* New insights into methane hydrate dissociation: Utilization of molecular dynamics strategy // Fuel. 2019. Vol. 249. P. 264-276. DOI: 10.1016/j.fuel.2019.02.125

22. *Caili Dai, Yong Hu, Yining Wu et al.* Effects of structural properties of alcohol molecules on decomposition of natural gas hydrates: A molecular dynamics study // Fuel. 2020. Vol. 268. № 117322. DOI: 10.1016/j.fuel.2020.117322

23. *Bingbing Chen, Huiru Sun, Kehan Li et al.* Experimental investigation of natural gas hydrate production characteristics via novel combination modes of depressurization with water flow erosion // Fuel. 2019. Vol. 252. P. 295-303. DOI: 10.1016/j.fuel.2019.04.120

24. *Yutaek Seo, Seong-Pil Kang.* Inhibition of methane hydrate re-formation in offshore pipelines with a kinetic hydrate inhibitor // Journal of Petroleum Science and Engineering. 2012. Vol. 88-89. P. 61-66. DOI: 10.1016/j.petrol.2011.11.001

25. *Ju Dong Lee, Englezos P.* Unusual kinetic inhibitor effects on gas hydrate formation // Chemical Engineering Science. 2006. Vol. 61. Iss. 5. P. 1368-1376. DOI: 10.1016/j.ces.2005.08.009

26. *Bo-Hui Shi, Shang-Fei Song, Xiao-Fang Lv et al.* Investigation on natural gas hydrate dissociation from a slurry to a water-in-oil emulsion in a high-pressure flow loop // Fuel. 2018. Vol. 233. P. 743-758. DOI: 10.1016/j.fuel.2018.06.054

27. *Bahadori A.* Fluid Phase Behavior for Conventional and Unconventional Oil and Gas Reservoirs. Gulf Professional Publishing, 2017. 554 p.

28. *Soroush E., Mesbah M., Shokrollahi A. et al.* Evolving a robust modeling tool for prediction of natural gas hydrate formation conditions // Journal of Unconventional Oil and Gas Resources. 2015. Vol. 12. P. 45-55. DOI: 10.1016/j.juogr.2015.09.002

29. *Mesbah M., Soroush E., Rezakazemi M.* Development of a least squares support vector machine model for prediction of natural gas hydrate formation temperature // Chinese Journal of Chemical Engineering. 2017. Vol. 25. Iss. 9. P. 1238-1248. DOI: 10.1016/j.cjche.2016.09.007

30. *Berge B.K.* Hydrate Predictions on a Microcomputer // Petroleum Industry Application of Microcomputers, 18-20 June 1986, Silvercreek, CO, USA. OnePetro, 1986. № SPE-15306-MS. DOI: 10.2118/15306-MS

31. *Motiee M.* Estimate possibility of hydrates // Hydrocarbon Processing. 1991. Vol. 70. № 7. P. 98-99.

32. *Sun C.-Y., Chen G.-J., Lin W., Guo T.-M.* Hydrate Formation Conditions of Sour Natural Gases // Journal of Chemical & Engineering Data. 2003. Vol. 48. Iss. 3. P. 600-602. DOI: 10.1021/je020155h

33. *Towler B., Mokhatab S.* Quickly estimate hydrate formation conditions in natural gases // Hydrocarbon Processing. 2005. Vol. 84. № 4. P. 61-63.

34. *Salufu S.O., Nwakwo P.* New Empirical Correlation for Predicting Hydrate Formation Conditions // SPE Nigeria Annual International Conference and Exhibition, 5-7 August 2013, Lagos, Nigeria. OnePetro, 2013. № SPE-167571-MS. DOI: 10.2118/167571-MS

35. *Safamirzaei M.* Predict gas hydrate formation temperature with a simple correlation. URL: https://gasprocessingnews.com/articles/2015/08/predict-gas-hydrate-formation-temperature-with-a-simple-correlation/ (дата обращения 09.02.2025).

36. *Elgibaly A.A., Elkamel A.M.* A new correlation for predicting hydrate formation conditions for various gas mixtures and inhibitors // Fluid Phase Equilibria. 1998. Vol. 152. Iss. 1. P. 23-42. DOI: 10.1016/S0378-3812(98)00368-9

37. *Zahedi G., Karami Z., Yaghoobi H.* Prediction of hydrate formation temperature by both statistical models and artificial neural network approaches // Energy Conversion and Management. 2009. Vol. 50. Iss. 8. P. 2052-2059. DOI: 10.1016/j.enconman.2009.04.005

38. *Riazi S.H., Heydari H., Ahmadpour E. et al.* Development of novel correlation for prediction of hydrate formation temperature based on intelligent optimization algorithms // Journal of Natural Gas Science and Engineering. 2014. Vol. 18. P. 377-384. DOI: 10.1016/j.jngse.2014.03.012

Статья опубликована в открытом доступе по лицензии CC BY 4.0

39. *Hesami S.M., Dehghani M., Kamali Z., Bakyani A.E.* Developing a simple-to-use predictive model for prediction of hydrate formation temperature // International Journal of Ambient Energy. 2017. Vol. 38. Iss. 4. P. 380-388. DOI: 10.1080/01430750.2015.1100678

40. *Amar M.N.* Prediction of hydrate formation temperature using gene expression programming // Journal of Natural Gas Science and Engineering. 2021. Vol. 89. № 103879. DOI: 10.1016/j.jngse.2021.103879

41. *Mesbah M., Habibnia S., Ahmadi S. et al.* Developing a robust correlation for prediction of sweet and sour gas hydrate formation temperature // Petroleum. 2022. Vol. 8. Iss. 2. P. 204-209. DOI: 10.1016/j.petlm.2020.07.007

42. *Hosseini M., Leonenko Y.* A reliable model to predict the methane-hydrate equilibrium: An updated database and machine learning approach // Renewable and Sustainable Energy Reviews. 2023. Vol. 173. № 113103. DOI: 10.1016/j.rser.2022.113103

43. *Karaköse M., Yücel Ö.* Predictive Modeling of the Hydrate Formation Temperature in Highly Pressurized Natural Gas Pipelines // Energies. 2024. Vol. 17. Iss. 21. № 5306. DOI: 10.3390/en17215306

44. *Khan N., Khan Ya., Zhang Lei et al.* Investigating a User-Friendly Machine Learning Approach to Predict the Gas Hydrate Formation Temperature Available to Purchase // Journal of the Geological Society of India. 2026. Vol. 102 (4). P. 555-565. DOI: 10.17491/jgsi/2026/174383

45. *Aleem W., Ahmad S., Qamar S. et al.* Exploring Machine Learning Techniques for Accurate Prediction of Methane Hydrate Formation Temperature in Brine: A Comparative Study // Arabian Journal for Science and Engineering. 2024. Vol. 49. Iss. 6. P. 8983-8996. DOI: 10.1007/s13369-024-09030-5

46. *Khafaie M., Masoumeh M., Azimi A., Mohammadi A.* Comparing ANN with PSO-ANN in Prediction of Natural Gas Hydrate Formation Parameters // Journal of the Japan Petroleum Institute. 2025. Vol. 68. Iss. 2. P. 53-60. DOI: 10.1627/jpi.68.53

47. *Mahmood Riyadh Atta, Akram Fadhl Al-Mahmodi, Bhajan Lal et al.* Artificial Intelligence and Machine Learning in Thermodynamic Gas Hydrate Studies: A Review // Energy & Fuels. 2025. Vol. 39. Iss. 38. P. 18287-18310. DOI: 10.1021/acs.energyfuels.5c02863

48. *Mokwenye P.O.* Evaluation Of Gas Hydrate In Gas Pipeline Transportation: A Thesis Submitted to the Graduate Faculty of the University of North Dakota In partial fulfillment of the requirements for the degree of Master of Science. Grand Forks, 2020. 107 p. (№ 3383). URL: https://commons.und.edu/theses/3383/ (дата обращения 09.02.2025).

49. *Azizi J., Azizi J., Davarnejad R.* Prediction of Gas Hydrate Formation using HYSYS Software // International Journal of Engineering, Transactions C: Aspects. 2014. Vol. 27. Iss. 9. P. 1325-1330.

50. *Bahadori A., Vuthaluru H.B.* A novel correlation for estimation of hydrate forming condition of natural gases // Journal of Natural Gas Chemistry. 2009. Vol. 18. Iss. 4. P. 453-457. DOI: 10.1016/S1003-9953(08)60143-7

51. *Wilcox W.I., Carson D.B., Katz D.L.* Natural Gas Hydrates // Industrial & Engineering Chemistry. 1941. Vol. 33. Iss. 5. P. 662-665. DOI: 10.1021/ie50377a027

52. *Guang-Jin Chen, Tian-Min Guo.* Thermodynamic modeling of hydrate formation based on new concepts // Fluid Phase Equilibria. 1996. Vol. 122. Iss. 1-2. P. 43-65. DOI: 10.1016/0378-3812(96)03032-4

53. *Jøssang A., Stange E.* A new predictive activity model for aqueous salt solutions // Fluid Phase Equilibria. 2001. Vol. 181. Iss. 1-2. P. 33-46. DOI: 10.1016/S0378-3812(00)00515-X

54. *Sloan E.D.* Clathrate hydrates of natural gases. Marcel Dekker, New York, 1990. 641 p.

55. *Nasrifar K., Moshfeghian M.* A model for prediction of gas hydrate formation conditions in aqueous solutions containing electrolytes and/or alcohol // The Journal of Chemical Thermodynamics. 2001. Vol. 33. Iss. 9. P. 999-1014. DOI: 10.1006/jcht.2000.0811

56. *Deaton W.M., Frost E.M.* Gas hydrates and their relation to the operation of natural-gas pipe mines. American Gas Association. United States. Bureau of Mines. Monograph 8. P. 101.

57. *Hashemkhani M., Soleimani R., Fazeli H. et al.* Prediction of the binary surface tension of mixtures containing ionic liquids using Support Vector Machine algorithms // Journal of Molecular Liquids. 2015. Vol. 211. P. 534-552. DOI: 10.1016/j.molliq.2015.07.038

58. *Suykens J.A.K., Van Gestel T., De Brabanter J. et al.* Least Squares Support Vector Machines. World Scientific, 2002. 308 p. DOI: 10.1142/5089

59. *Suykens J.A.K., Vandewalle J.* Least Squares Support Vector Machine Classifiers // Neural Processing Letters. 1999. Vol. 9. Iss. 3. P. 293-300. DOI: 10.1023/A:1018628609742

60. *Suykens J.A.K., De Brabanter J., Lukas L., Vandewalle J.* Weighted least squares support vector machines: robustness and sparse approximation // Neurocomputing. 2002. Vol. 48. Iss. 1-4. P. 85-105. DOI: 10.1016/S0925-2312(01)00644-0

61. *Varamesh A., Hemmati-Sarapardeh A., Keshavarz Moraveji M., Mohammadi A.H.* Generalized models for predicting the critical properties of pure chemical compounds // Journal of Molecular Liquids. 2017. Vol. 240. P. 777-793. DOI: 10.1016/j.molliq.2017.05.023

62. *Varamesh A., Hemmati-Sarapardeh A., Dabir B., Mohammadi A.H.* Development of robust generalized models for estimating the normal boiling points of pure chemical compounds // Journal of Molecular Liquids. 2017. Vol. 242. P. 59-69. DOI: 10.1016/j.molliq.2017.06.039

63. *Ningbo Zhao, Xueyou Wen, Jialong Yang et al.* Modeling and prediction of viscosity of water-based nanofluids by radial basis function neural networks // Powder Technology. 2015. Vol. 281. P. 173-183. DOI: 10.1016/j.powtec.2015.04.058

64. *Faris H., Aljarah I., Mirjalili S.* Evolving Radial Basis Function Networks Using Moth–Flame Optimizer // Handbook of Neural Computation. Elsevier, 2017. P. 537-550. DOI: 10.1016/B978-0-12-811318-9.00028-4

65. *Hemmati-Sarapardeh A., Varamesh A., Husein M.M., Karan K.* On the evaluation of the viscosity of nanofluid systems: Modeling and data assessment // Renewable and Sustainable Energy Reviews. 2018. Vol. 81. Part 1. P. 313-329. DOI: 10.1016/j.rser.2017.07.049

66. *Turnbull D., Elkan C.* Fast recognition of musical genres using RBF networks // IEEE Transactions on Knowledge and Data Engineering. 2005. Vol. 17. Iss. 4. P. 580-584. DOI: 10.1109/TKDE.2005.62

67. *Mehrjoo H., Riazi M., Nait Amar M., Hemmati-Sarapardeh A.* Modeling interfacial tension of methane-brine systems at high pressure and high salinity conditions // Journal of the Taiwan Institute of Chemical Engineers. 2020. Vol. 114. P. 125-141. DOI: 10.1016/j.jtice.2020.09.014

68. *Riazi M., Mehrjoo H., Nait Amar M. et al.* Modeling the density of acid gases at extensive ranges of pressure and temperature conditions // Journal of Petroleum Science and Engineering, 2021. Vol. 207. № 109063. DOI: 10.1016/j.petrol.2021.109063

69. *Rostami A., Hemmati-Sarapardeh A., Shamshirband S.* Rigorous prognostication of natural gas viscosity: Smart modeling and comparative study // Fuel. 2018. Vol. 222. P. 766-778. DOI: 10.1016/j.fuel.2018.02.069

70. *Zendehboudi A., Tatar A., Xianting Li.* A comparative study and prediction of the liquid desiccant dehumidifiers using intelligent models // Renewable Energy. 2017. Vol. 114. Part B. P. 1023-1035. DOI: 10.1016/j.renene.2017.07.078

Статья опубликована в открытом доступе по лицензии CC BY 4.0

71. *Riazi M., Mehrjoo H., Nakhaei R. et al.* Modelling rate of penetration in drilling operations using RBF, MLP, LSSVM, and DT models // Scientific Reports. 2022. Vol. 12. № 11650. DOI: 10.1038/s41598-022-14710-z

72. *Dabiri M.-S., Hadavimoghaddam F., Ashoorian S. et al.* Modeling liquid rate through wellhead chokes using machine learning techniques // Scientific Reports. 2024. Vol. 14. № 6945. DOI: 10.1038/s41598-024-54010-2

73. *Xinyuan Zou, Yingting Zhu, Jing Lv et al.* Toward Estimating CO_2 Solubility in Pure Water and Brine Using Cascade Forward Neural Network and Generalized Regression Neural Network: Application to CO_2 Dissolution Trapping in Saline Aquifers // ACS Omega. 2024. Vol. 9. Iss. 4. P. 4705-4720. DOI: 10.1021/acsomega.3c07962

74. *Nanda A.K., Neelakandan S., Gupta S. et al.* Multi-layer perceptron's neural network with optimization algorithm for greenhouse gas forecasting systems // Environmental Challenges. 2023. Vol. 11. № 100708. DOI: 10.1016/j.envc.2023.100708

75. *MacKay D.J.C.* Bayesian Interpolation // Neural Computation. 1992. Vol. 4. Iss. 3. P. 415-447. DOI: 10.1162/neco.1992.4.3.415

76. *Makomere R., Rutto H., Koech L., Banza M.* The use of artificial neural network (ANN) in dry flue gas desulphurization modelling: Levenberg–Marquardt (LM) and Bayesian regularization (BR) algorithm comparison // The Canadian Journal of Chemical Engineering. 2023. Vol. 101. Iss. 6. P. 3273-3286. DOI: 10.1002/cjce.24715

77. *Chavoshi S., Safamirzaei M., Pajoum Shariati F.* Evaluation of Empirical Correlations for Predicting Gas Hydrate Formation Temperature // Gas Processing Journal. 2018. Vol. 6. Iss. 2. P. 15-36. DOI: 10.22108/gpj.2018.112052.1036

78. *Shateri M., Ghorbani S., Hemmati-Sarapardeh A., Mohammadi A.H.* Application of Wilcoxon generalized radial basis function network for prediction of natural gas compressibility factor // Journal of the Taiwan Institute of Chemical Engineers. 2015. Vol. 50. P. 131-141. DOI: 10.1016/j.jtice.2014.12.011

79. *Benamara C., Gharbi K., Nait Amar M., Hamada B.* Prediction of Wax Appearance Temperature Using Artificial Intelligent Techniques // Arabian Journal for Science and Engineering. 2020. Vol. 45. Iss. 2. P. 1319-1330. DOI: 10.1007/s13369-019-04290-y

80. *Castillo E., Hadi A.S., Conejo A., Fernández-Canteli A.* A General Method for Local Sensitivity Analysis With Application to Regression Models and Other Optimization Problems // Technometrics. 2004. Vol. 46. Iss. 4. P. 430-444. DOI: 10.1198/004017004000000509

81. *Rousseeuw P.J., Leroy A.M.* Robust Regression and Outlier Detection. John Wiley & Sons, 2003. 360 p.

82. *Mehrjoo H., Riazi M., Rezaei F. et al.* Modeling interfacial tension of methane-brine systems at high pressure high temperature conditions // Geoenergy Science and Engineering. 2024. Vol. 242. № 213258. DOI: 10.1016/j.geoen.2024.213258

83. *Dabiri M.-S., Haji-Hashemi R., Ashoorian S. et al.* Modeling drilling fluid density at high-pressure high-temperature conditions using advanced machine-learning techniques // Geoenergy Science and Engineering. 2025. Vol. 244. № 213369. DOI: 10.1016/j.geoen.2024.213369

***Авторы*: Мохсен Риази,** *магистр* (*Университет им. Шахида Бахонара в Кермане, Керман, Иран*), **Хосейн Мехраджу,** *PhD* (*Университет им. Шахида Бахонара в Кермане, Керман, Иран*), **Мохаммад-Сабер Дабири,** *магистр* (*Университет им. Шахида Бахонара в Кермане, Керман, Иран*), **Джамильям Исмаилова,** *PhD, доцент* (*Казахский национальный исследовательский технический университет им. К.И.Сатпаева, Алматы, Республика Казахстан*), **Масуд Риази,** *PhD, доцент* (*Назарбаев Университет, Астана, Республика Казахстан*), *masoud.riazi@nu.edu.kz*.

Авторы заявляют об отсутствии конфликта интересов.

Статья опубликована в открытом доступе по лицензии CC BY 4.0