Применение символьной регрессии для модификации формулы Вильсона

Е. В. Колдоба*, А. Д. Говоров‡

С помощью символьной регрессии в формулу Вильсона для констант фазового равновесия (K-values) введено давление схождения. Полученные корреляции хорошо передают фазовое поведение газонефтяного месторождения заданного состава во всём диапазоне давлений.

Ключевые слова: символьная регрессия, формула Вильсона, давление схождения, K-values .

Введение. Нефть и газ содержат сотни компонентов, поэтому при фазовых переходов необходимо учитывать как можно больше количество компонентов – это можно сделать с помощью констант фазового равновесия (K-values). В газонефтяных задачах K-values — не константы, а функции давления, температуры и состава; при высоких давлениях они стремятся к единице при приближении к давлению схождения P_{conv}. При низких давлениях хорошо работает формула Вильсона, при высоких — её модификация Уитсона—Торпа, однако последняя ухудшает точность при низких давлениях. В данной работе с помощью символьной регрессии [1] в формулу Вильсона вводится P_{conv}, чтобы обеспечить корректную работу во всём диапазоне давлений.

Формула Вильсона для K-values имеет вид [2]:

$$K_i^{Wil} = \frac{p_{ci}}{P} \exp\left[5.37\,(1+\omega_i)\left(1-\frac{T_{ci}}{T}\right)\right], \quad i = 1, \ldots, N. \tag{1}$$

Согласно этой формуле при заданных давлении P и температуре T каждая функция K_i^{Wil} зависит только от характеристик i-го компонента: критических температуры T_{ci} и давления p_{ci}, ацентрического фактора ω_i, и не зависит от концентраций и взаимодействий с другими компонентами.

* *Колдоба Елена Валентиновна* — канд. физ.-мат. наук, доцент каф. вычислительной механики мех.-мат. ф-та МГУ, e-mail: elenakoldoba@mail.ru.

Koldoba Elena Valentinovna — Candidate of Physical and Mathematical Sciences, Docent, Lomonosov Moscow State University, Faculty of Mechanics and Mathematics, Chair of Computational Mechanics.

‡ *Говоров Александр Дмитриевич* — аспирант, кафедра вычислительной механики мех.-мат. ф-та МГУ, e-mail: alexandrdg@gmail.com.

Govorov Alexander Dmitrievich — Graduate student, Lomonosov Moscow State University, Faculty of Mechanics and Mathematics, Chair of Computational Mechanics.

Опубликовано на условиях лицензии CC-BY 4.0
Published under CC-BY 4.0

Известно, что при низких давлениях, когда смесь ещё можно считать идеальной, значения K-values монотонно убывают. При высоких давлениях функции K-values немонотонны и стремятся к единице. И этот предел достигается при некотором давлении схождения P_{conv}: $K_i \to 1$ при $P \to P_{\text{conv}}$ для всех i. Величина P_{conv} зависит от концентраций смеси и от взаимодействия компонентов.

Модификация формулы Вильсона для высоких давлений, предложенная Уитсоном—Торпом имеет вид [2]:

$$K_i^{Tor} = \left(\frac{p_{ci}}{P_{\text{conv}}}\right)^{A-1} \cdot \frac{\exp\left[5.37\,A(1+\omega_i)\left(1-T_{ri}^{-1}\right)\right]}{p_{ri}},\; A = 1 - \left(\frac{P-p_0}{P_{\text{conv}}-p_0}\right)^{\alpha}, \tag{2}$$

где $\alpha = 0.6$, $p_{ri} = P/p_{ci}$, $T_{ri} = T/T_{ci}$, $p_0 = 1$ бар; $A \to 1$ при низких давлениях и $A \to 0$ при $P \to P_{\text{conv}}$.

Символьная регрессия. Дано множество наблюдений:
$\{(P^{(j)}, T^{(j)}, \omega_i^{(j)}, p_{ci}^{(j)}, T_{ci}^{(j)}, P_{\text{conv}}^{(j)}, K_i^{(j)})\}_{j=1}^{M}$.

Требуется найти аналитическое выражение F такое, что

$$K_i^{(j)} \approx F\left(P^{(j)}, T^{(j)}, \omega_i^{(j)}, p_{ci}^{(j)}, T_{ci}^{(j)}, P_{\text{conv}}^{(j)}\right) \quad \text{для всех } j \text{ и } i.$$

Искомая функция F должна удовлетворять предельным переходам:

$$F\left(P \to 1 \text{ бар}\right) \approx K_i^{\text{Wil}}, \quad F\left(P \to P_{\text{conv}}\right) \to 1 \quad \text{для всех } i,$$

а также условиям монотонности: для легких компонентов: $F > 1$, $\frac{\partial F}{\partial P} < 0$ во всем диапазоне P; для всех компонентов $\frac{\partial F}{\partial P} < 0$ при низких давлениях; для тяжелых компонентов $F < 1$, $\frac{\partial F}{\partial P} > 0$ при высоких давлениях.

Символьная регрессия [1] минимизирует функционал:
$\mathcal{L} = \text{MSE}\left(K_i^{\text{pred}}, K_i^{\text{true}}\right) + \lambda\,\mathcal{R}$.

Здесь K_i^{true} — точное значение K_i i-го компонента из t-Navigator; K_i^{pred} — предсказание формулы-кандидата, предлагаемой символьной регрессией на текущем шаге эволюции; $\mathcal{R}$ — штраф за сложность; λ — параметр парсимонии; MSE вычисляется по обучающей выборке.

По завершении поиска формируется фронт Парето, элементы которого характеризуются компромиссом между точностью, сложностью. В пространство поиска входят базовые операторы $+, -, \times, /, \exp, \log, (\cdot)^a$, а также пользовательский оператор $(p_{ci}/P_{\text{conv}})^{A-1}$, где A – подвыражение, которое алгоритм выбирает из пространства поиска.

Выражение для физического штрафа:
$\lambda_{\text{conv}}\, w(P)\left(K_i^{\text{pred}}(P) - 1\right)^2$, $w(P) = \exp\left[-\frac{P_{\text{conv}}-P}{\Delta P}\right]$, где $\Delta P > 0$ — масштаб давления, λ_{conv} — коэффициент штрафа. При $P \to P_{\text{conv}}$ вес $w(P) \to 1$, а при удалении от P_{conv} убывает.

Расчет. Для газоконденсатного месторождения, состав которого представлен в таб.1, по композиционной модели симулятора t-Navigator при разных температурах вычислялись точные значения K_i^{true} и определялись давления схождения P_{conv}. При моделировании использовалось уравнение состояния Пенга-Робинсона, переменные нормировались к порядку $O(1)$.

Таблица 1: Параметры смеси

Компоненты	Концентрации	M_i (г/моль)	P_{ci} (бар)	T_{ci} (K)	ω_i
C1-C2, N_2	0.389665	17.1046	45.976	199.83	0.015
C3-C4	0.080687	51.6441	39.126	397.01	0.171
C5-C6	0.059501	78.2886	32.981	483.39	0.254
C7-C13	0.241202	126.491	27.055	615.45	0.431
C14-C20	0.108817	227.602	16.926	793.73	0.752
C21+	0.120128	473.914	9.171	919.58	1.243

Параметры численного эксперимента: $N = 6$, диапазон давлений $P = 1 \div 150$ бар, диапазон температур $T = 273 \div 800$ K, число точек обучения – 5000, число точек вблизи P_{conv} – 500, $\lambda_{\text{conv}} = 100$, $\lambda = 0.005$, число итераций = 50, число популяций = 25, размер популяции = 50, максимальный размер дерева $\nu = 30$, $\Delta P = 15$ бар. На фронте Парето выбиралась формула, которая имела приемлемую точность и удовлетворяла условиям (3)-(6).

В итоге получена формула следующего вида:

$$K_i = \left(\frac{p_{ci}}{P_{\text{conv}}}\right)^{A-1} \cdot \frac{\exp\left[5.37\, A\,(1+\omega_i)\left(1-T_{ri}^{-1}\right)\right]}{p_{ri}}, \quad A = 1 - \left(\frac{P}{P_{\text{conv}}}\right)^{\alpha},$$

где $\alpha = 0.415$.

На рис. 1 представлены зависимости K-values от давления при $T = 703$ K и $P_{\text{conv}} = 107$ бар, рассчитанные разными способами для легкого (а) и тяжелого(б) компонента. Для рассматриваемого месторождения формула Вильсона дает достаточно большие ошибки. Метод Уитсона-Торпа улучшает зависимость K-values. Расчеты, выполненные с помощью символьной регрессии дают результаты, которые ближе всего к точным значениям.

Заключение. Символьная регрессия позволила ввести P_{conv} в формулу Вильсона и получить зависимость, корректно работающую во всём диапазоне давлений и более близкую к точным значениям, чем модели Вильсона и Уитсона—Торпа.

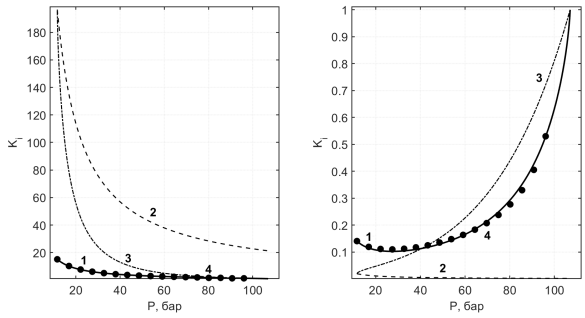


Рис. 1: Значения K-values: 1 — точные, 2 — Вильсон, 3 — Уитсон—Торп, 4 — символьная регрессия.

Список литературы

[1] M. Cranmer, “Interpretable Machine Learning for Science with PySR and SymbolicRegression.jl”, 2023, arXiv: 2305.01582.

[2] T.Ahmed, *Working Guide to Vapor-Liquid Phase Equilibria Calculations*, Gulf Professional Publishing, Amsterdam, 2009, 200 pp.

Статья поступила 6 ноября 2026 г.

Application of symbolic regression for the modification of Wilson’s formula

E. V. Koldoba, A. D. Govorov

Using symbolic regression, the convergence pressure has been incorporated into Wilson’s correlation for phase equilibrium constants (K-values). The resulting correlations accurately represent the phase behavior of an oil-and-gas reservoir with a given composition across the entire pressure range.

Keywords: symbolic regression, Wilson formula, convergence pressure, distribution coefficients .

References

[1] M. Cranmer, “Interpretable Machine Learning for Science with PySR and SymbolicRegression.jl”, 2023, arXiv: 2305.01582.

[2] T.Ahmed, *Working Guide to Vapor-Liquid Phase Equilibria Calculations*, Gulf Professional Publishing, Amsterdam, 2009, 200 pp.

Received on November 6, 2026