

Исследование влияния зашумленности исходных данных на точность определения коэффициентов взаимовлияния скважин методом линейной регрессии

Родион Михайлович Ганопольский✉

Тюменский государственный университет, Тюмень, Россия

Контакт для переписки: g.m.ganopolskij@utmn.ru✉

Аннотация. Для прогнозирования работы добывающей скважины используются гидродинамические симуляторы. Их работа занимает много времени, поэтому постоянно появляются альтернативные упрощенные методы, например прокси-модели. Сначала эти модели настраиваются на исторические данные, а затем дают прогнозы. Но очень часто исходные данные недостаточно точны и полны. Необходимо изучить, насколько ошибка в каждом исходном параметре влияет на качество прогноза.

В статье построен алгоритм определения коэффициентов взаимовлияния скважин методом линейной регрессии. Проведено исследование устойчивости полученного решения к зашумлению исходных данных при различных условиях. Получен прогноз дебита на добывающей скважине, который далее сравнивается с точным значением. Проанализировано, в каких исходных данных погрешность больше всего влияет на ошибку прогноза. Предложены возможные варианты обеспечения устойчивости.

Ключевые слова: прогноз добычи, прокси-модели, объемно-резистивная модель CRM, коэффициенты взаимовлияния скважин, метод линейной регрессии, устойчивость к зашумленности данных

Цитирование: Ганопольский Р. М. 2023. Исследование влияния зашумленности исходных данных на точность определения коэффициентов взаимовлияния скважин методом линейной регрессии // Вестник Тюменского государственного университета. Физико-математическое моделирование. Нефть, газ, энергетика. Том 9. № 1 (33). С. 107–115. <https://doi.org/10.21684/2411-7978-2023-9-1-107-115>

Поступила 20.02.2023; одобрена 27.03.2023; принята 31.03.2023

Influence of noisy input data in accuracy of well interference coefficients by linear regression method

Rodion M. Ganopolskij✉

University of Tyumen, Tyumen, Russia

Corresponding author: r.m.ganopolskij@utmn.ru✉

Abstract. Hydrodynamic simulators are used to predict the operation of a production well. Their work demands high performance. Alternative simplified methods are appearing constantly, for example, proxy models. Firstly, these models are tuned to historical data and then they make predictions. The source data is quite inaccurate and incomplete often. It is necessary to study how an input parameter error affects quality of forecast.

This article shows an algorithm for determining the well interference coefficients by linear regression method. Study was made of the stability of the obtained solution to the noise of the initial data under various conditions. A production well forecast has been obtained, which is then compared with the exact value. The error in which initial data has the greatest impact on the forecast error is analyzed. Possible options for ensuring sustainability are proposed.

Keywords: production forecast, proxy models, capacity resistive model CRM, well interference coefficients, linear regression method, resistance to data noise

Citation: Ganopolskij, R. M. (2023). Influence of noisy input data in accuracy of well interference coefficients by linear regression method. *Tyumen State University Herald. Physical and Mathematical Modeling. Oil, Gas, Energy*, 9(1), 107–115. <https://doi.org/10.21684/2411-7978-2023-9-1-107-115>

Received February 20, 2023; Reviewed March 27, 2023; Accepted March 31, 2023

Введение

Для оптимизации гидродинамической симуляции нефтегазового пласта используются различные методы. Одни из таких методов — семейство 0-мерных прокси-моделей CRM (capacity resistive model — объемно-резистивная модель), основанных на решении уравнения материального баланса [Sayarrou, 2008; Sayarrou и др., 2009; Yousef и др., 2006]:

$$q_{jk} = q_{j(k-1)} e^{-\frac{\Delta t}{\tau_j}} + \left(1 - e^{-\frac{\Delta t}{\tau_j}}\right) \sum_{i=1}^N f_{ij} I_{ik}, \quad (1)$$

где q_{jk} — дебит на добывающей скважине j в момент времени k ; I_{ik} — расход (приемистость) на нагнетательной скважине i ; N — количество нагнетательных скважин; Δt — шаг по времени; параметры f_{ij} — коэффициенты взаимовлияния скважин (их сумма не обязательно равна единице); τ_j — временная константа, определяемая как $\tau_j = cV/j$, где c — коэффициент сжимаемости, V — поровый объем, J — коэффициент продуктивности.

Параметры f_{ij} показывают, какой вклад вносит приемистость нагнетающей скважины i в дебит добывающей скважины j . Константу τ_j иногда называют временем запаздывания, учитывающим в модели изменение порового объема из-за давления или скорости распространения возмущения между скважинами. Этот параметр имеет диапазон значений от 0 до ∞ .

Как правило, полная модель CRM учитывает больше физических параметров, например забойное давление и т. п. [Sayarroug, 2008; Sayarroug и др., 2009; Yousef и др., 2006]. Мы будем использовать формулу (1) для первоначальной оценки влияния зашумленности данных. Есть еще более простая модель, которая пренебрегает константами τ_j , но в этом случае устойчивость метода возможно рассчитать аналитически.

Все параметры модели (1) можно определить для заданного месторождения только по известной истории закачки и добычи с помощью различных численных методов оптимизации. Один из таких способов — метод наименьших квадратов [Демиденко, 1981]. Найденные коэффициенты подставляются в формулу (1), затем рассчитывается прогноз добычи при известных приемистостях. Мы будем использовать метод линейной регрессии для поиска коэффициентов, но для этого сначала придется зависимость линеаризовать.

Известно, что исходные (промысловые) данные по месторождению даны с какой-то погрешностью, часто за какие-то дни данные отсутствуют. Рассмотрим, как неточность в этих данных влияет на погрешность параметров модели CRM и на дальнейшую ошибку прогноза.

Линеаризация зависимости

Исходная зависимость (1) не является линейной от искомых параметров, поэтому преобразуем ее. Для удобства перейдем от коэффициента τ к безразмерному параметру γ . В дальнейшем будем считать всё только для одной добывающей скважины, поэтому не будем использовать индекс j .

$$\gamma = \frac{1}{1 - e^{-\frac{\Delta t}{\tau}}}.$$

Новый коэффициент меняется в диапазоне значений от 0 до 1. Подставим его в формулу (1) и преобразуем полученное выражение:

$$q_k - q_{k-1} = -\frac{1}{\gamma} q_{k-1} + \frac{1}{\gamma} \sum_{i=1}^N f_i I_{ik}.$$

Введем обозначение $\Delta q_k = q_k - q_{k-1}$ и преобразуем:

$$q_{k-1} = -\gamma \Delta q_k + \sum_{i=1}^N f_i I_{ik}. \quad (2)$$

В таком виде мы получаем линейную зависимость дебита в предыдущий момент времени от других значений. На самом деле, если нам известно изменение и предыдущее значение, то ничего больше для вычисления следующего значения не нужно. Но такая зависимость (2) позволяет найти коэффициенты взаимовлияния методом линейной регрессии из известной истории добычи, что нам и требуется. Метод линейной регрессии дает решение [Демиденко, 1981]:

$$x = (A^T A)^{-1} A^T b, \quad (3)$$

где x — вектор искоемых коэффициентов $x = (f_1, \dots, f_N, \gamma)$; A — матрица из строк, содержащих значения I_{ik} и Δq_k ; b — вектор значений дебита в предыдущий момент.

Влияние зашумленности на расчет коэффициентов

Синтетические данные — семь нагнетательных скважин. Заданные истории нагнетательных скважин изображены на графиках на рис. 1.

Дебит на добывающей скважине определяется по формуле (1) с помощью коэффициентов, показанных в таблице 1. Результат расчета показан на рис. 2.

Далее с помощью метода линейной регрессии (3) из истории добывающей скважины на рис. 2 были получены значения параметров соотношения (2). Результат расчетов показан в таблице 2.

Отметим, что некоторые коэффициенты получились отрицательными, хотя и вблизи исходных. С помощью формулы (1) и рассчитанных коэффициентов был построен дебит добывающей скважины на весь период. График совпадает с графиком на рис. 2, обе линии сливаются.

Для проверки устойчивости метода сделаем равномерный случайный разброс в исходных данных (максимальные значения 3, 5 и 10%) с помощью формулы $\tilde{x} = x(1 - \Delta x + 2 \text{rand}(\Delta x))$, где Δx — максимальная заданная погрешность, x — исходное значение, \tilde{x} — зашумленная величина, rand — генератор случайных чисел от 0 до аргумента. Затем используем данные с внесенными погрешностями для определения коэффициентов взаимовлияния (3) и расчета прогноза по формуле (1).

На рис. 3 видно, как при увеличении зашумленности меняется прогноз.

Для выявления параметра, требующего больше точности, внесем погрешность разного порядка отдельно в данные по добывающей скважине и по нагнетательным. Из графиков на рис. 4 видно, что большее влияние имеет точность исходных данных по добывающей скважине: зашумленность значений нагнетательных скважин мало влияет на ошибку прогноза.

Стоит отметить, что даже при достаточно точном прогнозе полученные коэффициенты взаимовлияния и временная константа могут сильно отличаться от точных значений.

Для каждого заданных максимальных погрешностей были определены средние ошибки прогноза — дисперсия по всему временному периоду (таблица 3). Видно, что средняя ошибка больше коррелирует с максимальной погрешностью, вносимой в историю добывающей скважины. Но большие ошибки, накладываемые на данные по нагнетательным скважинам, тоже оказывают заметное влияние на дисперсию прогноза.

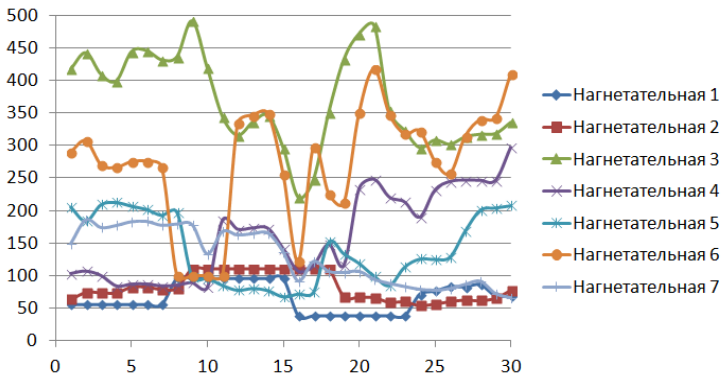


Рис. 1. История семи нагнетательных скважин

Fig. 1. History of seven injection wells

Таблица 1. Исходные коэффициенты CRM

Table 1. Initial CRM coefficients

f_1	f_2	f_3	f_4
0	0	0,036	0,055
f_5	f_6	f_7	γ
0,023	0,132	0	3,96

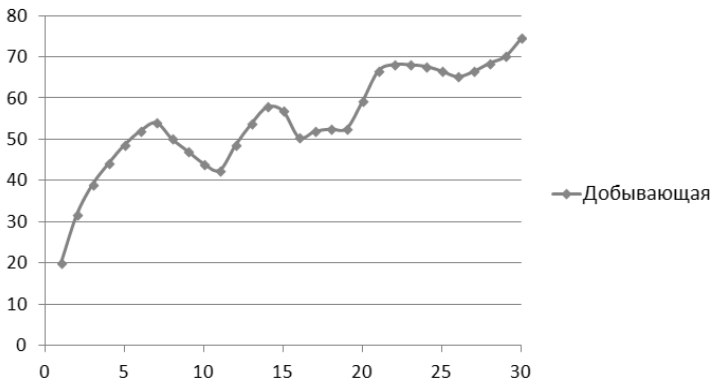


Рис. 2. Рассчитанный дебит добывающей скважины

Fig. 2. Calculated production well flow rate

Таблица 2. Коэффициенты CRM, определенные методом линейной регрессии

Table 2. CRM coefficients determined by linear regression

f_1	f_2	f_3	f_4
-5,48867E-14	4,03566E-14	0,036456	0,055307
f_5	f_6	f_7	γ
0,023069	0,131548	-1,16573E-15	3,9556277

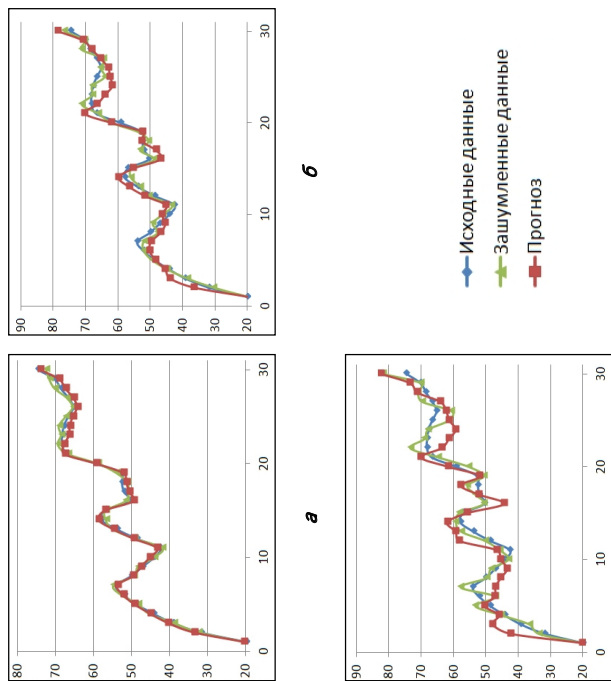


Рис. 3. Исходные и зашумленные данные (прогноз). Максимальный случайный разброс: а — 3%; б — 5%; в — 10%

Fig. 3. Initial and noisy data (the forecast). Maximum random spread: а — 3%; б — 5%; в — 10%

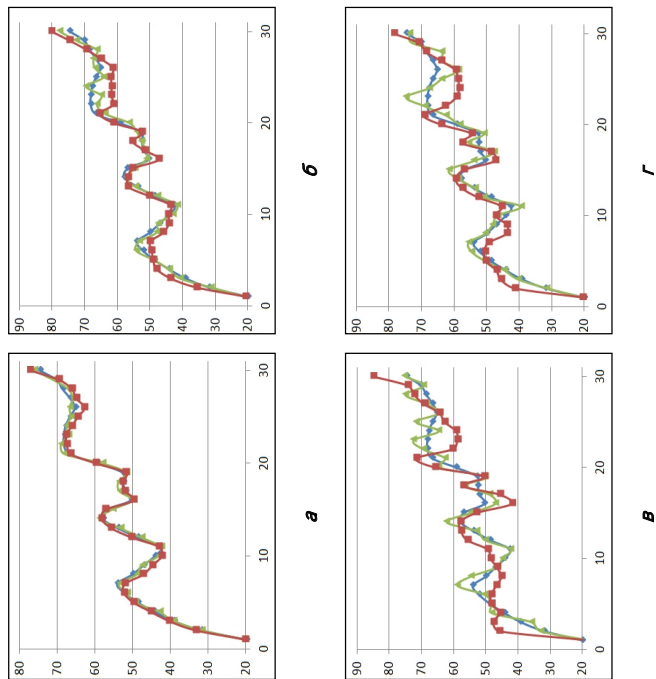


Рис. 4. Исходные и зашумленные данные (прогноз). Случайный разброс для нагнетательных-добывающей скважин: а — 10-3%; б — 10-5%; в — 0-10%; г — 5-10%

Fig. 4. Initial and noisy data (the forecast). Random spread for injection-production wells: а — 10-3%; б — 10-5%; в — 0-10%; г — 5-10%

Таблица 3. Средняя ошибка прогноза для разных погрешностей
Table 3. Average forecast error for different errors

Максимальная погрешность, %		Средняя ошибка прогноза, %
Нагнетательные	Добывающая	
3	3	2,3
10	3	2,7
5	5	5,5
10	5	5,5
0	10	10,0
5	10	11,0
10	10	11,0

Обсуждение

Прямое применение метода линейной регрессии иногда дает отрицательные коэффициенты взаимовлияния метода CRM, что не соответствует действительности. Это можно избежать модифицированным методом, вносящим ограничения в искомое решение [Данько и др., 2019; Демиденко, 1981; Ручкин и др., 2018; Степанов и др., 2020; Степанов и др., 2021].

Рис. 4 позволяет увидеть, что большее влияние на точность определения параметров метода CRM оказывает погрешность в данных истории добычи. Если данные дебита добывающей скважины достаточно точны, то зашумленность приемистости производит малый эффект. Отсюда следует, что и для других методов определения коэффициентов взаимовлияния будет проявляться похожая зависимость от погрешности исходных данных (например, метод наименьших квадратов, искусственные нейронные сети и др.), только, вероятно, другого порядка. Поэтому необходимо провести похожие проверки влияния погрешностей для других методов. В дальнейшем это позволит выбрать самый устойчивый метод и проверить его сначала на полной модели CRM, а затем на модификациях, в частности двухфазных [Olenchikov, Posvyanskii, 2019; Kim и др., 2012].

Заключение

Необходимо отметить, что полученная оценка влияния справедлива для зависимостей (1) и (2), где для определения дебита в любой взятый момент времени используются приемистости за тот же промежуток и дебит за предыдущий период. Благодаря выражению (1) можно построить линейную зависимость, где используются данные по всем скважинам за несколько временных отметок, что позволит нивелировать погрешности. Тем не менее рекомендуется также применить другие способы линеаризации формулы (1), чтобы сравнить устойчивости разных подходов.

Сильно отличающиеся коэффициенты взаимовлияния, определенные с помощью метода линейной регрессии, дают очень близкие прогнозы. Этот результат, в свою очередь, дает возможности для преобразования выражения (1), введения новых численных комплексов и дальнейшей линеаризации.

СПИСОК ИСТОЧНИКОВ

- Данько М. Ю., Бриллиант Л. С., Завьялов А. С. 2019. Применение метода динамического материального баланса и CRM-метода (capacitance-resistive models) к подсчету запасов Ачимовских и Баженовских коллекторов // Недропользование XXI век. № 4 (80). С. 76–85.
- Демиденко Е. З. 1981. Линейная и нелинейная регрессия. М.: Финансы и статистика. 302 с.
- Ручкин А. А., Степанов С. В., Князев А. В., Степанов А. В., Корытов А. В., Авсянко И. Н. 2018. Исследование особенностей оценки взаимовлияния скважин на примере модели CRM // Вестник Тюменского государственного университета. Физико-математическое моделирование. Нефть, газ, энергетика. Том 4. № 4. С. 148–168. <https://doi.org/10.21684/2411-7978-2018-4-4-148-168>
- Степанов С. В., Тырсин А. Н., Ручкин А. А., Поспелова Т. А. 2020. Использование энтропийного моделирования для анализа эффективности системы заводнения // Нефтяное хозяйство. № 6. С. 62–67. <https://doi.org/10.24887/0028-2448-2020-6-62-67>
- Степанов С. В., Бекман А. Д., Ручкин А. А., Поспелова Т. А. 2021. Сопровождение разработки нефтяных месторождений с использованием моделей CRM: монография. Тюмень: Экспресс. 298 с.
- Kim J. S., Lake L.W., Edgar T. F. 2012. Integrated capacitance-resistance model for characterizing waterflooded reservoirs // 2012 IFAC Workshop on Automatic Control in Offshore Oil and Gas Production (31 May — 1 June 2012, Trondheim, Norway). <https://doi.org/10.13140/2.1.2060.0964>
- Olenchikov D., Posvyanskii D. 2019. Application of CRM-like models for express forecasting and optimizing field development // SPE Russian Petroleum Technology Conference (22–24 October 2019, Moscow, Russia). <https://doi.org/10.2118/196893-MS>
- Sayarpour M. 2008. Development and Application of Capacitance-Resistive Models to Water/CO₂ Floods. Ph. D diss. The University of Texas at Austin. 236 p.
- Sayarpour M., Zuluaga E., Kabir C. S., Lake L. W. 2009. The use of capacitance-resistance models for rapid estimation of waterflood performance and optimization. Journal of Petroleum Science and Engineering. Vol. 69. No. 3–4. Pp. 227–238. <https://doi.org/10.1016/j.petrol.2009.09.006>
- Yousef A. A., Gentil P., Jensen J. L., Lake L. W. 2006. A capacitance model to infer interwell connectivity from production- and injection-rate fluctuations // SPE Reservoir Evaluation & Engineering. Vol. 9. No. 6. Pp. 630–646. SPE-10.2118. <https://doi.org/10.2118/95322-PA>

References

- Danko, M. Yu., Brilliant, L. S., & Zavyzlov, A. S. (2019). Application of dynamic material balance method and CRM method (capacitance-resistive models) for reserves assessment in Achimov and Bazhenov reservoirs. *Nedropolzovanie XXI vek*, (4), 76–85. [In Russian]
- Demidenko, E. Z. (1981). *Linear and non-linear regression*. Finansy i statistika. [In Russian]
- Ruchkin, A. A., Stepanov, S. V., Knyazev, A. V., Stepanov, A. V., Korytov, A. V., & Avsyanko, I. N. (2018). Applying CRM model to study well interference. *Tyumen State University Herald. Physical and Mathematical Modeling. Oil, Gas, Energy*, 4(4), 148–168. <https://doi.org/10.21684/2411-7978-2018-4-4-148-168> [In Russian]

- Stepanov, S. V., Tyrsin, A. N., Ruchkin, A. A., & Pospeiova, T. A. (2020). Using entropy modeling to analyze the effectiveness of the waterflooding system. *Oil Industry*, (6), 62–67. <https://doi.org/10.24887/0028-2448-2020-6-62-67> [In Russian]
- Stepanov, S. V., Bekman, A. D., Ruchkin, A. A., & Pospelova, T. A. (2021). *Maintenance of oil field development using CRM models*. Express. [In Russian]
- Kim, J. S., Lake, L. W., & Edgar, T. F. (2012, May 31 – June 1). *Integrated capacitance-resistance model for characterizing waterflooded reservoirs* [Conference paper]. 2012 IFAC Workshop on Automatic Control in Offshore Oil and Gas Production, Trondheim, Norway. <https://doi.org/10.13140/2.1.2060.0964>
- Olenchikov, D., & Posvyanskii, D. (2019, October 22–24). *Application of CRM-like models for express forecasting and optimizing field development* [Conference paper]. SPE Russian Petroleum Technology Conference, Moscow, Russia. <https://doi.org/10.2118/196893-MS>
- Sayarpour, M. (2008). *Development and application of capacitance-resistive models to water/CO₂ floods* [Doctoral dissertation, The University of Texas at Austin].
- Sayarpour, M., Zuluaga, E., Kabir, C. S., & Lake, L. W. (2009). The use of capacitance-resistance models for rapid estimation of waterflood performance and optimization. *Journal of Petroleum Science and Engineering*, 69(3–4), 227–238. <https://doi.org/10.1016/j.petrol.2009.09.006>
- Yousef, A. A., Gentil, P., Jensen, J. L., & Lake, L. W. (2006). A capacitance model to infer interwell connectivity from production- and injection-rate fluctuations. *SPE Reservoir Evaluation & Engineering*, 9(6), 630–646. <https://doi.org/10.2118/95322-PA>

Информация об авторе

Родион Михайлович Ганопольский, кандидат физико-математических наук, заведующий кафедрой моделирования физических процессов и систем, Физико-технический институт, Тюменский государственный университет, Тюмень, Россия
r.m.ganopolskij@utmn.ru

Information about the author

Rodion M. Ganopolskij, Cand. Sci. (Phys.-Math.), Head of the Department of Modeling of Physical Processes and Systems, Institute of Physics and Technology, University of Tyumen, Tyumen, Russia
r.m.ganopolskij@utmn.ru