

## Сравнительный анализ методов фильтрации замеров данных скважин сложной конструкции

Давид Юзаевич Шенгелия<sup>1</sup>✉, Игорь Викторович Коваленко<sup>2</sup>,  
Ирина Гелиевна Захарова<sup>1</sup>

<sup>1</sup> Тюменский государственный университет, Тюмень, Россия

<sup>2</sup> Газпромнефть НТЦ, Тюмень, Россия

Контакт для переписки: [d.y.shengeliya@utmn.ru](mailto:d.y.shengeliya@utmn.ru)✉

**Аннотация.** Статья посвящена сравнительному анализу различных методов фильтрации синтетических замеров, имитирующих данные гидродинамических исследований скважин (ГДИС).

Основной целью работы является выявление наиболее эффективных методов фильтрации зашумленных данных ГДИС с позиции сохранения полезной информации и облегчения задачи последующей интерпретации результатов.

В качестве исходных данных использовался набор из 200 синтетических кривых снижения давления (КСД) и кривых восстановления давления (КВД) с различным уровнем искусственно внесенного шума. Для фильтрации данных рассматривались как классические методы (фильтр Каалмана, фильтр Савицкого — Голея, одномерная фильтрация Гаусса), так и численные методы на основе нейронных сетей (автокодировщики) и методов машинного обучения (метод опорных векторов).

В результате сравнительного анализа было продемонстрировано, что результаты применения различных методов фильтрации зависят от типа обрабатываемой кривой (КСД или КВД) и характеристик скважины. Наилучшие показатели по критериям отношения «сигнал — шум» (signal-noise ratio, SNR) и корня среднеквадратической ошибки (RMSE) продемонстрировали современные методы на основе автокодировщиков.

Сделан вывод, что выбор оптимального метода фильтрации требует детального анализа специфики задачи и характеристик исходных данных. Предложено комбинирование различных методов фильтрации для повышения качества обработки и интерпретации данных скважин сложной конструкции.

Полученные результаты имеют практическую значимость, позволяя упростить решение задачи сегментации КСД и КВД, что необходимо для корректной идентификации различных периодов работы скважины в процессе ее исследования.

**Ключевые слова:** гидродинамические исследования скважин, скважины сложной конструкции, фильтр Калмана, фильтр Савицкого — Голея, одномерная фильтрация Гаусса, метод опорных векторов, автокодировщики

**Цитирование:** Шенгелия Д. Ю., Коваленко И. В., Захарова И. Г. 2024. Сравнительный анализ методов фильтрации замеров данных скважин сложной конструкции // Вестник Тюменского государственного университета. Физико-математическое моделирование. Нефть, газ, энергетика. Том 10. № 2 (38). С. 104–120. <https://doi.org/10.21684/2411-7978-2024-10-2-104-120>

Поступила 08.05.2024; одобрена 17.05.2024; принята 24.05.2024

## Comparative analysis of filtering methods for measurement data from complex well configurations

David Yu. Shengeliya<sup>1</sup>✉, Igor V. Kovalenko<sup>2</sup>, Irina G. Zakharova<sup>1</sup>

<sup>1</sup> University of Tyumen, Tyumen, Russia

<sup>2</sup> Gazpromneft Science & Technology Center, Tyumen, Russia

Corresponding author: [d.y.shengeliya@utmn.ru](mailto:d.y.shengeliya@utmn.ru)✉

**Abstract.** This article presents a comparative analysis of various filtering methods for synthetic measurements that simulate data from well test analysis (WTA).

The main objective of this work is to identify the most effective filtering methods for noisy WTA data, with the aim of preserving useful information and facilitating the subsequent interpretation of the results.

The initial dataset consisted of 200 synthetic pressure drawdown (PDD) and pressure buildup (PBU) curves with varying levels of artificially introduced noise. Both classical filtering methods (Kalman filter, Savitzky–Golay filter, one-dimensional Gaussian filtering) and numerical methods based on neural networks (autoencoders) and machine learning (support vector machines) were considered for data filtering.

The comparative analysis demonstrated that the performance of different filtering methods depends on the type of curve (PDD or PBU) and the well characteristics. The best results in terms of signal-to-noise ratio (SNR) and root mean square error (RMSE) were achieved using modern autoencoder-based methods.

The conclusion is that the choice of an optimal filtering method requires a detailed analysis of the specific problem and the characteristics of the input data. A combination of different filtering methods is proposed to improve the quality of processing and interpretation of WTA data for complex well designs.

The obtained results have practical significance, as they can simplify the segmentation of PDD and PBU curves, which is necessary for the correct identification of various operating periods of the well during the investigation process.

**Keywords:** well testing, complex well configurations, Kalman filter, Savitzky–Golay filter, one-dimensional Gaussian filtering, support vector machines, autoencoders

**Citation:** Shengeliya, D. Yu., Kovalenko, I. V., & Zakharova, I. G. (2024). Comparative analysis of filtering methods for measurement data from complex well configurations. *Tyumen State University Herald. Physical and Mathematical Modeling. Oil, Gas, Energy*, 10(2), 104–120. <https://doi.org/10.21684/2411-7978-2024-10-2-104-120>

Received May 8, 2024; Reviewed May 17, 2024; Accepted May 24, 2024

## Введение

Данные замеров с манометра во время проведения гидродинамических исследований скважин (ГДИС) представляют собой последовательность кривых снижения давления (КСД) и кривых восстановления давления (КВД). Эти данные несут в себе важную информацию о свойствах пласта и скважины, необходимую для последующей интерпретации результатов ГДИС [Есипов и др., 2014; Асалузина и др., 2020; Коваленко, 2023].

Однако на практике такие данные часто оказываются зашумленными по ряду причин, в числе которых влияние ствола скважины (ВСС), погрешности измерений на ранних этапах исследования, а также другие факторы (например, скин-фактор) [Unneland и др., 1998]. Шумы существенно затрудняют корректную интерпретацию результатов ГДИС и определение искоемых показателей пласта и скважины.

В связи с этим была определена цель разработки эффективных (с точки зрения устранения шумов при сохранении полезного сигнала) методов фильтрации зашумленных данных замеров с манометра во время проведения ГДИС.

Данное исследование может быть полезным для задачи автоматической сегментации КСД и КВД по следующим причинам:

- эффективная фильтрация позволит устранить шумы и выделить полезный сигнал КСД и КВД;
- сохранение достоверной информации о свойствах пласта в отфильтрованных КСД и КВД обеспечит возможность корректной последующей интерпретации [Бухмастова и др., 2020];
- устранение шумов и аномалий облегчит задачу автоматического распознавания и сегментации КСД и КВД с помощью методов машинного обучения и использования нейронных сетей [Vaferi и др., 2011; Бриллиант и др., 2022].

В данной работе рассматриваются методы фильтрации синтетических данных КСД и КВД, позволяющие устранять шумы с минимальными потерями полезной информации, сохраняя при этом достоверность результатов ГДИС и корректность определения искоемых параметров пласта и скважины.

## Материалы и методы

В качестве исходных данных использовался набор из 200 эталонных кривых — 100 КСД и 100 КВД (1 062 точки для каждой кривой) с различным количеством трещин (от 2 до 21). Данный диапазон основан на фактических данных, полученных при исследовании ряда добывающих скважин на нефтегазовом месторождении X. Нижняя граница установлена для охвата минимального сценария, в то время как верхний предел в 21 трещину позволяет рассмотреть случаи с высокой плотностью трещинообразования. Зашумленные синтетические данные  $P_{noisy}(t)$  формировались путем добавления к эталонным данным  $P_{ideal}(t)$  случайного гауссовского шума  $n(t)$  со средним 0 и заданной дисперсией  $\sigma^2$ , определяемой уровнем шума  $nf$ :

$$P_{noisy}(t) = P_{ideal}(t) + n(t),$$

где  $n(t) \sim N(0, (nf \cdot P_{ideal}(t))^2)$ . Величина  $nf$  варьировалась от 0,02 до 0,2 с шагом 0,02. Такой подход имитирует ситуацию, когда шум измерений имеет большую дисперсию для больших значений измеряемого давления, что полностью соответствует реальным условиям эксплуатации скважин на месторождении X, когда погрешность измерений с манометра пропорциональна самим измеряемым величинам. Кроме того, каждый набор синтетических данных инициализировался со 100 различными случайными числами (random seed) для формирования множества различных шумовых профилей. На рис. 1 и 2 представлены примеры максимально зашумленных данных КСД и КВД соответственно.

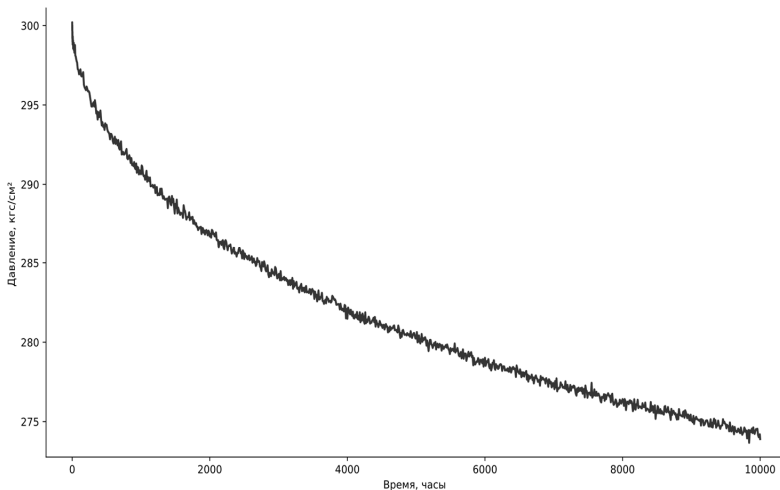


Рис. 1. Пример графика КСД (степень зашумления — 0,2)

Fig. 1. Example of the PDC plot (noise level — 0.2)

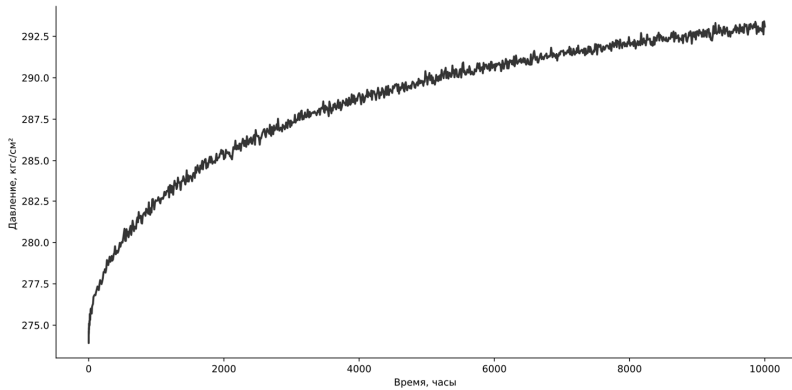


Рис. 2. Пример графика КВД (степень зашумления — 0,2)

Fig. 2. Example of the PRC plot (noise level — 0.2)

В работе были рассмотрены классические и численные методы для фильтрации шумов в данных ГДИС.

### Классические методы

**Фильтр Калмана.** Рекурсивный алгоритм оценивания состояния динамической системы, позволяющий учитывать статистические характеристики шумов и погрешностей [Nikoofard и др., 2015; Заикин, Кареев, 2020].

Он опирается на следующие уравнения:

— состояния:

$$x_{k+1} = A_k x_k + B_k u_k + w_k, \quad (1)$$

где  $x_{k+1}$  — вектор состояния системы в следующий момент времени;  $x_k$  — вектор состояния системы в текущий момент времени;  $A_k$  — матрица перехода состояния;  $B_k$  — матрица управления;  $u_k$  — вектор управляющих воздействий;  $w_k$  — вектор шума процесса;

— наблюдения:

$$z_k = H_k x_k + v_k, \quad (2)$$

где  $z_k$  — вектор измерений;  $H_k$  — матрица наблюдения;  $v_k$  — вектор шума измерения;

— предсказания состояния:

$$\hat{x}_{k+1|k} = A_k \hat{x}_{k|k} + B_k u_k, \quad (3)$$

где  $\hat{x}_{k+1|k}$  — предсказанное состояние в следующий момент времени;  $\hat{x}_{k|k}$  — оцененное состояние в текущий момент времени;

— предсказания ошибки ковариации:

$$P_{k+1|k} = A_k P_{k|k} A_k^T + Q_k, \quad (4)$$

где  $P_{k+1|k}$  — предсказанная ковариация ошибки;  $P_{k|k}$  — оцененная ковариация ошибки в текущий момент времени;  $Q_k$  — ковариация шума процесса;

— обновления состояния:

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(z_k - H_k \hat{x}_{k|k-1}), \quad (5)$$

где  $\hat{x}_{k|k}$  — обновленное состояние в текущий момент времени;  $K_k$  — коэффициент усиления Калмана;

— обновления ковариации ошибки:

$$P_{k|k} = (I - K_k H_k) P_{k|k-1}, \quad (6)$$

где  $P_{k|k}$  — обновленная ковариация ошибки в текущий момент времени;  $I$  — единичная матрица;

— коэффициента усиления Калмана:

$$K_k = P_{k|k-1} H_k^T (H_k P_{k|k-1} H_k^T + R_k)^{-1}, \quad (7)$$

где  $R_k$  — ковариация шума измерения.

В одномерной реализации  $R$  и  $Q$  представляют собой скалярные значения [Арио и др., 2019]. Процесс работы фильтра Калмана, описанного в (1)–(7), состоит из трех этапов:

1. Инициализация параметров:  $P_0 = 0$ .
2. Предсказание состояния:  $Y_{pred} = y_i$ ;  $P_{pred} = P_i + Q$ .
3. Обновление состояния:  $K_i = P_{pred} / (P_{pred} + R)$ ;  $P_i = (1 - K_i) \cdot P_{pred}$ ;  $Y_i = Y_{pred} + K_i \cdot (y_i - Y_{pred})$ .

К преимуществам можно отнести высокую точность и адаптивность, возможность прогнозирования состояния. Недостатки — вычислительная сложность для многомерных систем.

*Фильтр Савицкого — Голея.* Гладкая аппроксимация сглаживающим многочленом для шумоподавления [Сулейманов и др., 2014; Lei и др., 2020]. В общем виде описывается следующим образом:

$$Y_j = \sum_{i=\frac{1-m}{2}}^{\frac{m-1}{2}} C_i y_{j+i}, \quad \frac{m+1}{2} \leq j \leq n - \frac{m-1}{2},$$

где  $Y_j$  — сглаженная точка данных с индексом  $j$ ;  $m$  — ширина окна сглаживания;  $C_i$  — коэффициенты свертки, зависящие от  $m$  и степени полинома.

Из преимуществ — простота реализации и вычисления. Недостаток — возможная потеря высокочастотных составляющих сигнала.

*Одномерная фильтрация Гаусса.* Линейная свертка с весовой функцией, по форме напоминающей гауссовское распределение [Osman, Stewart, 1997].

В основе данного фильтра лежит следующая формула:

$$g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}, \quad (8)$$

где  $x$  — точка, к которой применяется фильтр;  $\sigma$  — стандартное отклонение, определяющее степень сглаживания фильтра.

Для применения фильтра (8) к массиву данных  $f(x)$  используется свертка:

$$g_{\sigma}(x) = \int_{-\infty}^{\infty} f(t)g(x - t). \quad (9)$$

В дискретном случае, когда данные представлены массивом, (9) аппроксимируется как

$$g_{\sigma}(x) = \sum_{t=-\infty}^{\infty} f(t)g(x - t).$$

Основным преимуществом является эффективное подавление выбросов и шумов. К недостаткам относится необходимость подбора параметров ядра.

### Численные методы

*Метод опорных векторов (support vector machine, SVM).* Данный метод является одним из популярных подходов для решения регрессионных задач в рамках машинного обучения (support vector regression, SVR). В его основе лежит построение гиперплоскости в многомерном пространстве признаков, которая максимально близко аппроксимирует исходные данные [Khoukhi и др., 2011; Ansari, Gholami, 2015].

Математически это можно представить следующим образом. Дано обучающее множество  $\{(x_i, y_i), i = 1 \dots n\}$ , где  $x_i$  — вектор входных признаков, а  $y_i$  — соответствующее значение целевой переменной.

Целевая функция метода SVR имеет вид

$$\min \left( \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N (\xi_i + \xi_i^*) \right) \quad (10)$$

при ограничениях:

$$y_i - wx_i - b \leq \varepsilon + \xi_i, \quad (11)$$

$$wx_i + b - y_i \leq \varepsilon + \xi_i^*, \quad (12)$$

$$\xi_i, \xi_i^* \geq 0, \quad (13)$$

где  $w$  — вектор весов гиперплоскости;  $b$  — смещение гиперплоскости;  $\varepsilon$  — величина допустимой ошибки;  $\xi_i, \xi_i^*$  — величины отклонений от целевой функции (слабые и сильные ошибки соответственно);  $C$  — параметр регуляризации, определяющий компромисс между сложностью модели и точностью аппроксимации. Важной особенностью SVR является использование ядерных функций (kernel functions), которые позволяют работать с нелинейными зависимостями между признаками и целевой переменной. Наиболее популярные ядерные функции — линейная, полиномиальная, радиальная базисная (RBF) и сигмоидная [Awad, Khanna, 2015].

Из основных преимуществ — эффективность в высокоразмерных пространствах без перегрузки модели, устойчивость к выбросам, а также возможность использования нелинейных ядер [Aung и др., 2020]. Из недостатков — высокая чувствительность к выбору гиперпараметров для целевой функции (10) с учетом ограничений (11)–(13) и высокая вычислительная сложность при большом объеме данных.

*Автокодировщики.* Общий принцип работы автокодировщиков (специализированных искусственных нейронных сетей [Kuester и др., 2021; Jiang и др., 2022]) заключается в следующем:

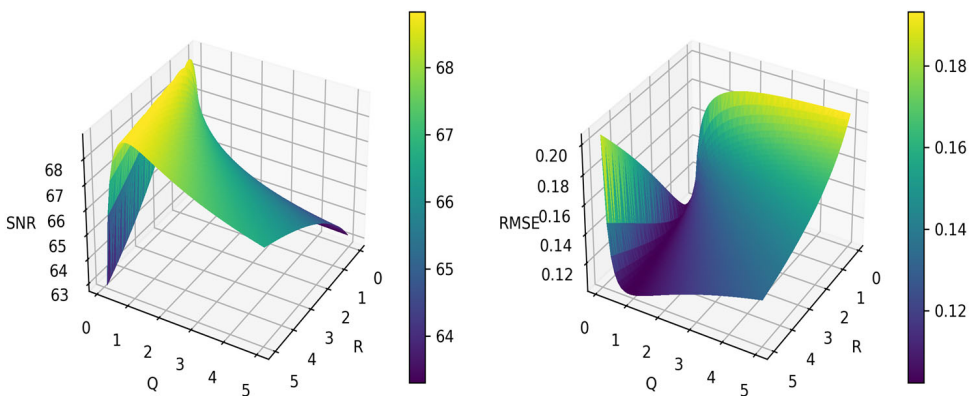
1. На нейронную сеть подается пара входных сигналов — в чистом виде и с шумом.
2. Автокодировщик пытается реконструировать исходный чистый сигнал на выходе.
3. В процессе автокодировщик обучается более устойчивому внутреннему представлению данных, игнорируя шум.

Для объективного сравнения методы тестировались на синтетических данных с известными свойствами шума и полезного сигнала ГДИС. Качество фильтрации оценивалось метриками RMSE и SNR. Первая метрика показывает среднее отклонение отфильтрованных значений давления от истинных значений давления в системе, вторая метрика характеризует степень подавления случайных шумов и помех при замерах синтетических данных давления.

## Результаты и обсуждение

### Фильтр Калмана

На рис. 3 приведен график, показывающий зависимость значений RMSE и SNR от параметров  $R$  и  $Q$ . В результате вычислительных экспериментов для используемой выборки было определено, что наилучшие результаты для характеристик качества фильтрации достигаются при значениях  $Q \in [0,1; 1]$  и  $R \in [0,5; 4,9]$ .



**Рис. 3.** Значения SNR и RMSE в зависимости от параметров  $R$  и  $Q$

**Fig. 3.** SNR and RMSE values as a function of  $R$  and  $Q$  parameters



### Фильтр Савицкого — Голея

На рис. 4 представлен график, который демонстрирует влияние двух параметров на результаты применения сглаживающего фильтра к данным: размер окна, к которому применяется фильтр, и степень полинома, используемого для аппроксимации данных внутри окна. Из графика (рис. 4) видно, что для исследуемого набора данных наилучшие результаты фильтрации (наименьшее значение RMSE и наибольшее значение SNR) были достигнуты при ширине окна  $W_{size} \in [67; 119]$ . Это составляет от 5 до 9% от общего размера выборки данных.

Необходимо отметить, что степень полинома, которым аппроксимируются данные внутри окна, оказывает влияние на результат. Зависимости, представленные на рис. 4, показывают, что оптимальными являются полиномы третьей, четвертой и пятой степеней. В некоторых случаях сопоставимые результаты показывают и полиномы второй степени.

### Одномерный фильтр Гаусса

На рис. 5 приведен график, показывающий зависимость значений RMSE и SNR от значения  $\sigma$ . В результате вычислительных экспериментов для используемой выборки было определено, что наилучшие результаты фильтр демонстрирует при значениях  $\sigma \in [7,5; 9]$ .

При меньших значениях  $\sigma$  фильтр недостаточно хорошо удаляет шумовую составляющую, что приводит к высокому RMSE. При больших  $\sigma$  происходит чрезмерное сглаживание и теряется полезный сигнал, что отражается в низком SNR.

### Метод опорных векторов

На рис. 6 приведен график, показывающий зависимость значений RMSE и SNR от значений  $C$  (параметр регуляризации L2) и  $\epsilon$  (порог допустимой ошибки). В результате вычислительных экспериментов для используемой выборки было определено, что наилучшим ядром для фильтрации является `rbf` (ядро, основанное на радиально-базисной функции) и что наилучшие результаты метод опорных векторов демонстрирует при значениях  $C \in [60\ 000; 80\ 000]$  и  $\epsilon \in [0,01; 0,1]$ .

### Автокодировщики

Denoising Autoencoder (DAE) — это специализированная архитектура искусственной нейронной сети, разновидность автокодировщика, основанная на полносвязных слоях.

Convolutional Autoencoder (CAE) — разновидность автокодировщика, использующая свертки вместо полносвязных слоев. Из ключевых аспектов CAE можно отметить свойственную структурированным данным локальную корреляцию для более эффективного кодирования, сохранение пространственной и топологической структуры входных данных, а также меньшее число параметров по сравнению с полносвязными автокодировщиками.

В таблицах 1 и 2 приведены архитектуры CAE и DAE, применяемые для фильтрации данных КСД и КВД.

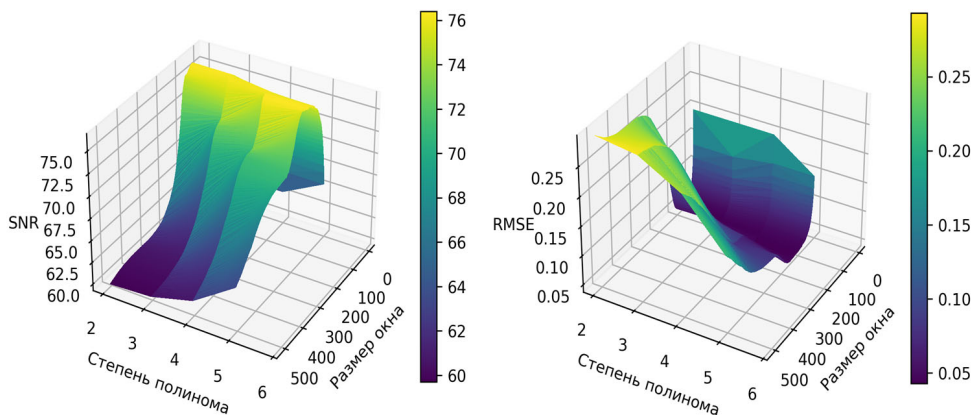


Рис. 4. Значения SNR и RMSE в зависимости от размера окна и степени полинома

Fig. 4. SNR and RMSE values as a function of window size and polynomial degree

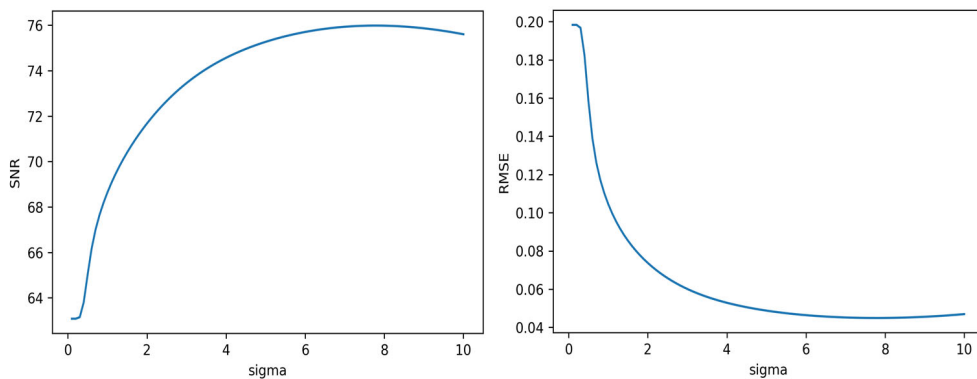


Рис. 5. Значения SNR и RMSE в зависимости от значения  $\sigma$

Fig. 5. SNR and RMSE values as a function of  $\sigma$  value

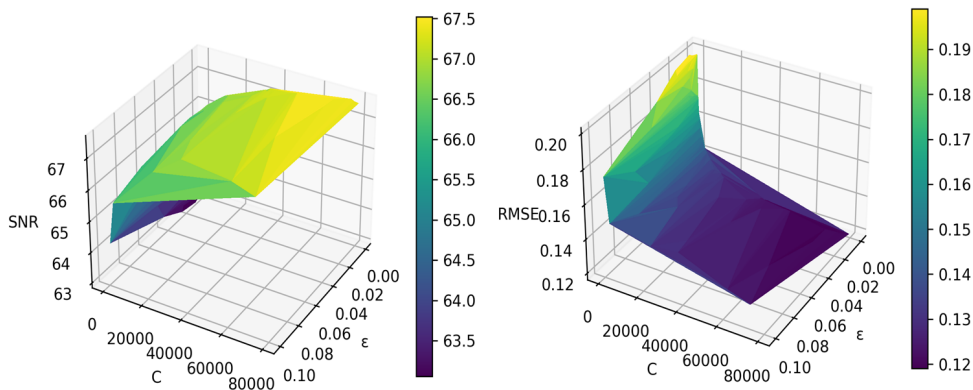


Рис. 6. Значения SNR и RMSE в зависимости от значений C и  $\epsilon$

Fig. 6. SNR and RMSE values as a function of the values of C and  $\epsilon$

**Таблица 1.** Архитектура CAE для фильтрации КСД и КВД**Table 1.** CAE architecture for filtering PDCs and PRCs

Тип слоя	Входные данные	Выходные данные	Параметры
<b>Кодировщик</b>			
Conv1D	(1, 2, 1 062)	(16, 1, 532)	Размер ядра — $3 \times 3$ , шаг — 2, pad — 1
ReLU → Dropout ( $p = 0,2$ )			
Conv1D	(16, 1, 532)	(32, 1, 266)	Размер ядра — $3 \times 3$ , шаг — 2, pad — 1
ReLU → Dropout ( $p = 0,2$ )			
<b>Декодировщик</b>			
ConvTranspose1D	(32, 1, 266)	(16, 1, 532)	Размер ядра — $3 \times 3$ , шаг — 2, pad — 1, output_pad — 1
ReLU			
ConvTranspose1D	(16, 1, 532)	(1, 2, 1062)	Размер ядра — $3 \times 3$ , шаг — 2, pad — 1, output_pad — 1
Sigmoid			

**Таблица 2.** Архитектура DAE для фильтрации КСД и КВД**Table 2.** DAE architecture for filtering PDCs and PRCs

Тип слоя	Входные данные	Выходные данные
<b>Кодировщик</b>		
Dense	(2, 1 062)	(2, 532)
ReLU		
Dense	(2, 532)	(2, 266)
ReLU		
Dense	(2, 266)	(2, 133)
ReLU		
<b>Декодировщик</b>		
Dense	(2, 133)	(2, 266)
ReLU		
Dense	(2, 266)	(2, 532)
ReLU		
Dense	(2, 532)	(2, 1 062)

### Сравнительный анализ методов фильтрации

В таблице 3 приведены сравнительные показатели SNR и RMSE для типового случая КСД (7 трещин, полудлина трещины — 100 м, скин-фактор = 0,13, ВСС = 0,2). Лучшие результаты SNR и RMSE достигнуты при размере окна 145 и полиноме пятой степени для фильтра Савицкого — Голея, при  $\sigma = 7,8$  для фильтра Гаусса и при архитектурах DAE и CAE.

**Таблица 3.** Сравнительная характеристика различных методов фильтрации замеров КСД

**Table 3.** Comparative characterization of different filtering methods for PDC measurements

Метод фильтрации	Параметры	SNR, дБ	RMSE, кгс/см <sup>2</sup>
Фильтр Калмана	$Q = 0,8; R = 4$	68,3693	0,1079
	$Q = 0,7; R = 3,7$	68,3749	0,1078
	$Q = 0,8; R = 4,3$	68,3753	0,1078
Фильтр Савицкого — Голея	окно = 73 отсчета; степень = 3	76,4635	0,04249
	окно = 117 отсчета; степень = 4	76,4687	0,04246
	окно = 145 отсчета; степень = 5	76,4687	0,04246
Фильтр Гаусса	$\sigma = 8,5$	75,942	0,0451
	$\sigma = 8$	75,98	0,0449
	$\sigma = 7,8$	75,984	0,0445
Метод опорных векторов, ядро — rbf	$C = 76\ 000; \varepsilon = 0,001$	67,6105	0,11774
	$C = 74\ 000; \varepsilon = 0,001$	67,5889	0,11804
	$C = 68\ 000; \varepsilon = 0,01$	67,5878	0,11805
DAE	—	77,2319	0,04031
CAE	—	77,5634	0,03985

В таблице 4 приведены значения показателей SNR и RMSE для типового случая КВД (9 трещин, полудлина трещины — 150 м, скин-фактор = 0,1, ВСС = 0,2). На данном примере видно, что по сравнению со случаем КСД фильтр Гаусса и фильтр Савицкого — Голея выдают примерно одинаковые значения RMSE и SNR как между собой, так и в рамках различных параметров для каждого из них. Фильтр Калмана и метод опорных векторов показывают себя более стабильно, сохраняя сопоставимый уровень RMSE и SNR, однако автокодировщики DAE и CAE показывают наилучшие результаты.

**Таблица 4.** Сравнительная характеристика различных методов фильтрации замеров КВД**Table 4.** Comparative characterization of different filtering methods for PRC measurements

Метод фильтрации	Параметры	SNR, дБ	RMSE, кгс/см <sup>2</sup>
Фильтр Калмана	$Q = 0,8; R = 4$	65,7198	0,149
	$Q = 0,7; R = 3,7$	65,612	0,1509
	$Q = 0,8; R = 4,3$	65,5797	0,1514
Фильтр Савицкого — Голея	окно = 73 отсчета; степень = 3	63,2204	0,1987
	окно = 117 отсчета; степень = 4	63,1856	0,1995
	окно = 145 отсчета; степень = 5	63,1743	0,1997
Фильтр Гаусса	$\sigma = 8,5$	63,3549	0,1955
	$\sigma = 8$	63,3693	0,1953
	$\sigma = 7,8$	63,3812	0,195
Метод опорных векторов, ядро — rbf	$C = 76\ 000; \varepsilon = 0,001$	64,8324	0,1622
	$C = 74\ 000; \varepsilon = 0,001$	64,8216	0,1625
	$C = 68\ 000; \varepsilon = 0,01$	64,8138	0,1627
DAE	—	68,0144	0,1342
CAE	—	68,1621	0,1325

## Заключение

Применение различных методов фильтрации синтетических данных для скважин сложной конструкции показало, что эффективность того или иного метода зависит от типа обрабатываемых данных, а также от характеристик скважины.

Классические методы фильтрации продемонстрировали приемлемую результативность для определенных случаев и конфигураций скважин, однако современные методы, основанные на применении нейронных сетей, показали лучшие результаты при сопоставимой производительности по сравнению с классическими методами.

Таким образом, выбор оптимального метода фильтрации требует тщательного анализа специфики задачи и характеристик исходных данных, а комбинирование традиционных и современных методов фильтрации (пример для КВД: сверточный автокодировщик — для первичной фильтрации, фильтр Калмана — опционально для более точной фильтрации с учетом различных параметров) позволяет повысить качество обработки и интерпретации данных скважин сложной конструкции. Данное исследование является одним из основных шагов к улучшению качества сегментации кривых давления на периоде работы скважины сложной конструкции.

## СПИСОК ИСТОЧНИКОВ

- Асалхузина Г. Ф., Давлетбаев А. Я., Хабибуллин И. Л., Ахметова Р. Р. 2020. К вопросу выбора длительности режимов при гидродинамических исследованиях скважин на установившихся режимах закачки в низкопроницаемых коллекторах // Вестник Тюменского государственного университета. Физико-математическое моделирование. Нефть, газ, энергетика. Том 6. № 1 (21). С. 135–149. <https://doi.org/10.21684/2411-7978-2020-6-1-135-149>
- Бриллиант Л. С., Дулкарнаев М. Р., Данько М. Ю., Елишева А. О., Набиев Д. Х., Хуторная А. И., Мальков И. Н. 2022. Управление добычей нефти на основе нейросетевой оптимизации режимов работы скважин на участке опытно-промышленных работ пласта ЮВ1 Ватгеганского месторождения ТПП «Повхнефтегаз» // Георесурсы. Том 24. № 1. С. 3–15. <https://doi.org/10.18599/grs.2022.1.1>
- Бухмастова С. В., Фахреева Р. Р., Питюк Ю. А., Давлетбаев А. Я., Азарова Т. П., Фаргер Д. В., Якупов Р. Ф. 2020. Апробация методов MLR и CRMIP при исследовании взаимодействия скважин // Нефтяное хозяйство. № 8. С. 58–62. <https://doi.org/10.24887/0028-2448-2020-8-58-62>
- Есипов Д. В., Куранаков Д. С., Лапин В. Н., Чёрный С. Г. 2014. Математические модели гидроразрыва пласта // Вычислительные технологии. Том 19. № 2. С. 33–61.
- Заикин А. А., Кареев И. А. 2020. Моделирование потоков жидкости в нефтеносных пластах с помощью фильтра Калмана // Итоги науки и техники. Серия «Современная математика и ее приложения. Тематические обзоры». Том 175. С. 27–35. <https://doi.org/10.36535/0233-6723-2020-175-27-35>
- Коваленко И. В. 2023. Гидродинамическое моделирование горизонтальных скважин с многостадийным гидроразрывом пласта с учетом потерь давлений на конвергенцию линий тока // Нефтепромысловое дело. № 2 (650). С. 26–28. [https://doi.org/10.33285/0207-2351-2023-2\(650\)-26-28](https://doi.org/10.33285/0207-2351-2023-2(650)-26-28)
- Судейманов Б. А., Дышин О. А., Исаев Р. Ж. 2014. Интерпретация кривых восстановления давления на основе анализа детерминированных моментов забойного давления // Нефтепромысловое дело. № 1. С. 12–23.
- Ansari H. R., Gholami A. 2015. An improved support vector regression model for estimation of saturation pressure of crude oils // Fluid Phase Equilibria. Vol. 402. Pp. 124–132. <https://doi.org/10.1016/j.fluid.2015.05.037>
- Apio A., Dambros J. W., Farenzena M., Trierweiler J. O. 2019. Comparison of Kalman filter-based approaches for permanent downhole gauge pressure estimation in offshore oil production // Journal of Petroleum Science and Engineering. Vol. 182. Article 106254. <https://doi.org/10.1016/j.petrol.2019.106254>
- Aung Z., Mikhaylov I. S., Thu Aung Y. 2020. Application of support vector system for solving problems of classification and forecasting of oil wells // 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus). Pp. 568–572. <https://doi.org/10.1109/EIConRus49466.2020.9039343>
- Awad M., Khanna R. 2015. Support vector regression // Efficient Learning Machines. Berkeley: Apress. Pp. 67–80. [https://doi.org/10.1007/978-1-4302-5990-9\\_4](https://doi.org/10.1007/978-1-4302-5990-9_4)
- Jiang J., Ren H., Zhang M. 2022. A convolutional autoencoder method for simultaneous seismic data reconstruction and denoising // IEEE Geoscience and Remote Sensing Letters. Vol. 19. Article 7503405. <https://doi.org/10.1109/LGRS.2021.3073560>

- Khoukhi A., Oloso M., Elshafei M., Abdurraheem A., Al-Majed A. 2011. Support vector regression and functional networks for viscosity and gas/oil ratio curves estimation // *International Journal of Computational Intelligence and Applications*. Vol. 10. No. 03. Pp. 269–293. <https://doi.org/10.1142/S1469026811003100>
- Kuester J., Gross W., Middelman W. 2021. 1D-convolutional autoencoder based hyperspectral data compression // *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XLIII-B1-2021. Pp. 15–21. <https://doi.org/10.5194/isprs-archives-XLIII-B1-2021-15-2021>
- Lei J., Fan Y., Li Y., Xu T. 2020. Data analysis of hydraulic fracturing pressure in unconventional oil and gas fields // *IOP Conference Series: Earth and Environmental Science*. Vol. 558. No. 2. Article 022057. <https://doi.org/10.1088/1755-1315/558/2/022057>
- Nikooifard A., Aarsnes U. J. F., Johansen T. A., Kaasa G.-O. 2015. Estimation of states and parameters of a drift-flux model with unscented Kalman filter // *IFAC-PapersOnLine*. Vol. 48. No. 6. Pp. 165–170. <https://doi.org/10.1016/j.ifacol.2015.08.026>
- Osman M. S., Stewart G. 1997. Pressure data filtering and horizontal well test analysis case study // *Middle East Oil Show and Conference (15–18 March 1997, Bahrain)*. Paper SPE-37802-MS. <https://doi.org/10.2118/37802-MS>
- Unneland T., Manin Y., Kuchuk F. 1998. Permanent gauge pressure and rate measurements for reservoir description and well monitoring: Field cases // *SPE Reservoir Evaluation & Engineering*. Vol. 1. No. 3. Pp. 168–176. <https://doi.org/10.2118/38658-PA>
- Vaferi B., Eslamloueyan R., Ayatollahi Sh. 2011. Automatic recognition of oil reservoir models from well testing data by using multi-layer perceptron networks // *Journal of Petroleum Science and Engineering*. Vol. 77. No. 3–4. Pp. 254–262. <https://doi.org/10.1016/j.petrol.2011.03.002>

## References

- Asalkhuzina, G. F., Davletbaev, A. Ya., Khabibullin, I. L., & Akhmetova, R. R. (2020). On the selection of suitable operate durations for injection tests in low permeability reservoirs. *Tyumen State University Herald. Physical and Mathematical Modeling. Oil, Gas, Energy*, 6(1), 135–149. <https://doi.org/10.21684/2411-7978-2020-6-1-135-149> [In Russian]
- Brilliant, L. S., Dulkarnaev, M. R., Danko, M. Yu., Elisheva, A. O., Nabiev, D. Kh., Khutornaya, A. I., & Malkov, I. N. (2022). Oil production management based on neural network optimization of well operation at the pilot project site of the Vatyeganskoe field (Territorial Production Enterprise Povkhneftegaz). *Georesources*, 24(1), 3–15. <https://doi.org/10.18599/grs.2022.1.1> [In Russian]
- Bukhmastova, S. V., Fakhreeva, R. R., Pityuk, Yu. A., Davletbaev, A. Ya., Azarova, T. P., Farger, D. V., & Yakupov, R. F. (2020). Approbation of MLR and CRMIP methods in research of well interference. *Oil Industry Journal*, (8), 58–62. <https://doi.org/10.24887/0028-2448-2020-8-58-62> [In Russian]
- Esipov, D. V., Kuranakov, D. S., Lapin, V. N., & Cherny, S. G. (2014). Mathematical models of hydraulic fracturing. *Computational Technologies*, 19(2), 33–61. [In Russian]
- Zaikin, A. A., & Kareev, I. A. (2020). Modeling fluid flows in oil fields using the Kalman filter. *Itogi nauki i tekhniki. Seriya "Sovremennaya matematika i ee prilozheniya. Tematicheskie obzory"*, 175, 27–35. <https://doi.org/10.36535/0233-6723-2020-175-27-35> [In Russian]



- Kovalenko, I. V. (2023). Hydrodynamic simulation of horizontal wells by multi-stage hydraulic fracturing of a formation taking into account pressure losses due to streamline convergence. *Oilfield Engineering*, (2), 26–28. [https://doi.org/10.33285/0207-2351-2023-2\(650\)-26-28](https://doi.org/10.33285/0207-2351-2023-2(650)-26-28) [In Russian]
- Suleimanov, B. A., Dyshin, O. A., & Isaev, R. Zh. (2014). Interpretation of pressure built-up curves on the basis of analysis of bottomhole pressure deterministic moments. *Oilfield Engineering*, (1), 12–23. [In Russian]
- Ansari, H. R., & Gholami, A. (2015). An improved support vector regression model for estimation of saturation pressure of crude oils. *Fluid Phase Equilibria*, 402, 124–132. <https://doi.org/10.1016/j.fluid.2015.05.037>
- Apio, A., Dambros, J. W., Farenzena, M., & Trierweiler, J. O. (2019). Comparison of Kalman filter-based approaches for permanent downhole gauge pressure estimation in offshore oil production. *Journal of Petroleum Science and Engineering*, 182, Article 106254. <https://doi.org/10.1016/j.petrol.2019.106254>
- Aung, Z., Mikhaylov, I. S., & Thu Aung, Y. (2020). Application of support vector system for solving problems of classification and forecasting of oil wells. In *2020 IEEE Conference of Russian young researchers in electrical and electronic engineering (EIConRus)* (pp. 568–572). <https://doi.org/10.1109/EIConRus49466.2020.9039343>
- Awad, M., & Khanna, R. (2015). Support vector regression. In *Efficient learning machines* (pp. 67–80). Apress. [https://doi.org/10.1007/978-1-4302-5990-9\\_4](https://doi.org/10.1007/978-1-4302-5990-9_4)
- Jiang, J., Ren, H., & Zhang, M. (2022). A convolutional autoencoder method for simultaneous seismic data reconstruction and denoising. *IEEE Geoscience and Remote Sensing Letters*, 19, Article 7503405. <https://doi.org/10.1109/LGRS.2021.3073560>
- Khokhi, A., Oloso, M., Elshafei, M., Abdulraheem, A., & Al-Majed, A. (2011). Support vector regression and functional networks for viscosity and gas/oil ratio curves estimation. *International Journal of Computational Intelligence and Applications*, 10(03), 269–293. <https://doi.org/10.1142/S1469026811003100>
- Kuester, J., Gross, W., & Middelman, W. (2021). 1D-convolutional autoencoder based hyperspectral data compression. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B1-2021, 15–21. <https://doi.org/10.5194/isprs-archives-XLIII-B1-2021-15-2021>
- Lei, J., Fan, Y., Li, Y., & Xu, T. (2020). Data analysis of hydraulic fracturing pressure in unconventional oil and gas fields. *IOP Conference Series: Earth and Environmental Science*, 558(2), Article 022057. <https://doi.org/10.1088/1755-1315/558/2/022057>
- Nikoofard, A., Aarsnes, U. J. F., Johansen, T. A., & Kaasa, G.-O. (2015). Estimation of states and parameters of a drift-flux model with unscented Kalman filter. *IFAC-PapersOnLine*, 48(6), 165–170. <https://doi.org/10.1016/j.ifacol.2015.08.026>
- Osman, M. S., & Stewart, G. (1997, March 15–18). *Pressure data filtering and horizontal well test analysis case study* [Conference paper SPE-37802-MS]. Middle East Oil Show and Conference, Bahrain. <https://doi.org/10.2118/37802-MS>
- Unneland, T., Manin, Y., & Kuchuk, F. (1998). Permanent gauge pressure and rate measurements for reservoir description and well monitoring: Field cases. *SPE Reservoir Evaluation & Engineering*, 1(3), 168–176. <https://doi.org/10.2118/38658-PA>
- Vaferi, B., Eslamloueyan, R., & Ayatollahi, Sh. (2011). Automatic recognition of oil reservoir models from well testing data by using multi-layer perceptron networks. *Journal of Petroleum Science and Engineering*, 77(3–4), 254–262. <https://doi.org/10.1016/j.petrol.2011.03.002>



## Информация об авторах

*Давид Юзаевич Шенгелия*, аспирант кафедры программного обеспечения, Школа компьютерных наук, Тюменский государственный университет, Тюмень, Россия  
d.y.shengeliya@utmn.ru, <https://orcid.org/0009-0004-5222-9672>

*Игорь Викторович Коваленко*, кандидат технических наук, руководитель программы развития продуктов, Газпромнефть НТЦ, Тюмень, Россия  
kovalenko.iv@gazpromneft-ntc.ru

*Ирина Гелиевна Захарова*, кандидат физико-математических наук, профессор кафедры программного обеспечения, Школа компьютерных наук, Тюменский государственный университет, Тюмень, Россия  
i.g.zakharova@utmn.ru, <https://orcid.org/0000-0002-4211-7675>

## Information about the authors

*David Yu. Shengeliya*, Postgraduate Student, Department of Software, School of Computer Science, University of Tyumen, Tyumen, Russia  
d.y.shengeliya@utmn.ru, <https://orcid.org/0009-0004-5222-9672>

*Igor V. Kovalenko*, Cand. Sci. (Tech.), Product Development Manager, Gazpromneft Science & Technology Center, Tyumen, Russia  
kovalenko.iv@gazpromneft-ntc.ru

*Irina G. Zakharova*, Cand. Sci. (Phys.-Math.), Professor, Department of Software, School of Computer Science, University of Tyumen, Tyumen, Russia  
i.g.zakharova@utmn.ru, <https://orcid.org/0000-0002-4211-7675>