

15

Сингулярный спектральный анализ и метод нормированного размаха — метод дифференцирования зашумленных экспериментальных кривых

© М.Л. Бахмутский

Федеральный научный центр Научно-исследовательский институт системных исследований НИЦ „Курчатовский институт“, Москва, Россия
E-mail: mbakhmut@mail.ru

Поступило в Редакцию 28 июля 2025 г.

В окончательной редакции 19 августа 2025 г.

Принято к публикации 25 сентября 2025 г.

Предлагается дифференцировать экспериментальную кривую конечными разностями в физическом пространстве. При этом резко возрастает амплитуда шума, однако это не меняет случайную природу шума. Используя метод сингулярного спектрального анализа, производную представляют в виде суммы аддитивных компонент. У этих компонент методом нормированного размаха оценивают фрактальную размерность (индекс Херста). При синтезе производной суммируют компоненты до тех пор, пока индекс Херста производной не начнет уменьшаться.

Ключевые слова: дифференцирование, сингулярный спектральный анализ, индекс Херста.

DOI: 10.61011/PJTF.2026.03.62176.20454

Необходимость дифференцировать зашумленную экспериментальную кривую возникает во многих задачах: от анализа результатов гидродинамических исследований пластов до задач томографии плазмы и спектрометрии. В работе предлагается устойчивый метод решения этой некорректной задачи, метод свободный от неоднозначности почти общепринятых методов, предварительного сглаживания кривой или увеличения расстояния между экспериментальными точками при вычислении разностной аппроксимации производной. Метод дифференцирования экспериментальной функции заключается в применении к разностной аппроксимации производной метода сингулярного спектрального анализа (SSA) [1,2] для представления ее в виде суммы аддитивных аддитивных компонент и вычисления методом нормированного размаха (*R/S*) [3] фрактальной размерности этой суммы. Таким образом, из суммы исключают шумовые компоненты, оставляя физически значимую часть.

Полагаем, что имеем N значений функции $\tilde{f}_i = f_i + \delta_i$, измеренных в моменты времени t_i ($i = 1-N$). Аппроксимация производной со вторым порядком по Δt в момент t_i имеет вид

$$\tilde{g}_i = \frac{\tilde{f}_{i+1} - \tilde{f}_{i-1}}{2\Delta t} = \frac{f_{i+1} - f_{i-1}}{2\Delta t} + \frac{\delta_{i+1} - \delta_{i-1}}{2\Delta t}, \quad (1)$$

если

$$\Delta t \equiv t_{i+1} - t_i = \text{const.}$$

Пусть $|\delta_i| \leq A$, тогда оценкой ошибки вычисления производной является $\frac{A}{\Delta t} \rightarrow \infty$, если $\Delta t \rightarrow 0$, т. е. разностное дифференцирование приводит к росту амплитуды шума и росту изрезанности кривой. Изрезанность кривой можно оценивать методом *R/S* [3], вычисляя показатель Херста.

Шаг 1. Пусть мы имеем ряд значений \tilde{g}_i , вычисленных по формуле (1): $\tilde{g}_1, \tilde{g}_2, \dots, \tilde{g}_N$. Используя окно длиной m отсчетов, поставим в соответствие этому ряду так называемую траекторную матрицу $\tilde{A} = \{\tilde{a}_{i,j}\}$, где $\tilde{a}_{i,j} = \tilde{g}_{i+j-1}$, $i = 1-m$, $j = 1-n$, $m \leq \left[\frac{N+1}{2}\right]$, $n = N + 1 - m$. В нашем случае для однозначности возьмем окно $m = \left[\frac{N+1}{2}\right]$,

$$\tilde{A} = \begin{pmatrix} \tilde{g}_1 & \tilde{g}_2 & \cdots & \tilde{g}_n \\ \tilde{g}_2 & \tilde{g}_3 & & \tilde{g}_{n+1} \\ \vdots & & \ddots & \vdots \\ \tilde{g}_m & \tilde{g}_{m+1} & \cdots & \tilde{g}_N \end{pmatrix}. \quad (2)$$

По этой матрице можно однозначно восстановить ряд \tilde{g}_j усреднением по формуле

$$\tilde{g}_j = \begin{cases} \frac{1}{j} \sum_{i=1}^j \tilde{a}_{i,j-i+1}, & 1 \leq j \leq m, \\ \frac{1}{m} \sum_{i=1}^m \tilde{a}_{i,j-i+1}, & m \leq j \leq n, \\ \frac{1}{N-j+1} \sum_{i=1}^{N-j+1} \tilde{a}_{i+j-n,n-i+1}, & n \leq j \leq N. \end{cases} \quad (3)$$

Вычислим сингулярное разложение матрицы $\tilde{A} = U\Sigma V^T$, где U — матрица с ортогональными столбцами размерностью m (левые сингулярные векторы), V — матрица правых ортогональных сингулярных векторов размерностью n , $\Sigma = \text{diag}\{\sigma_i\}$, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$ [4].

Шаг 2. Вычисляем последовательность матриц $A^k = \sum_{i=1}^k \sigma_i u_i \times v_i^T$, $k = 1-k_0$.

Здесь u_i и v_i — i -е сингулярные векторы матриц U и V , знак \times — внешнее (кронекерово) умножение,

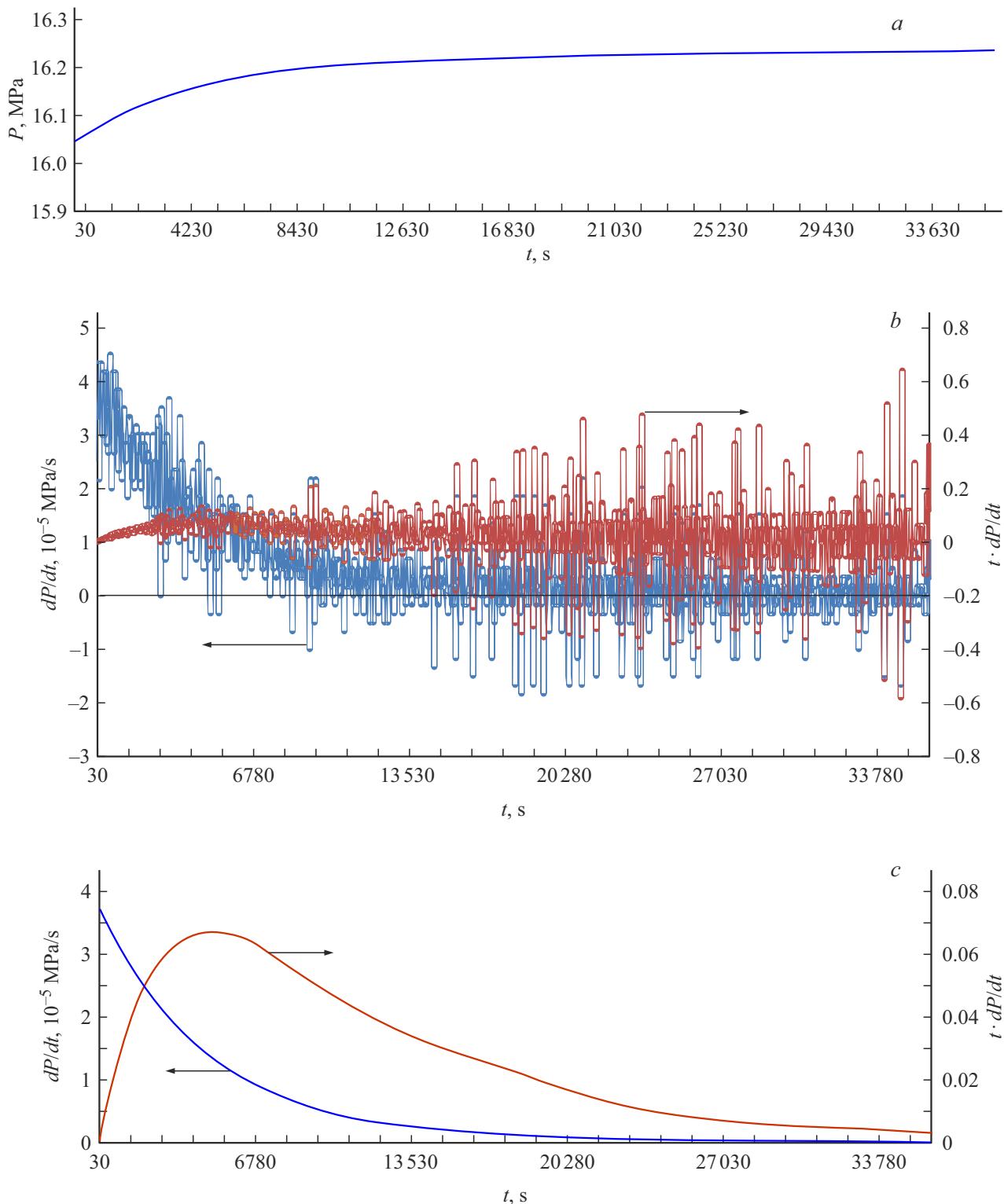


Рис. 1. *a* — кривая восстановления давления скважины одного из месторождений. По оси ординат — давление, по оси абсцисс — время. *b* — график производной, вычисленной по формуле $f'_i = \frac{f_{i+1} - f_{i-1}}{x_{i+1} - x_{i-1}}$. По оси ординат — производная давления, по оси абсцисс — время, по вспомогательной оси ординат — логарифмическая производная. *c* — графики производной (основная ось ординат) и логарифмической производной, восстановленной по первой сингулярной тройке.

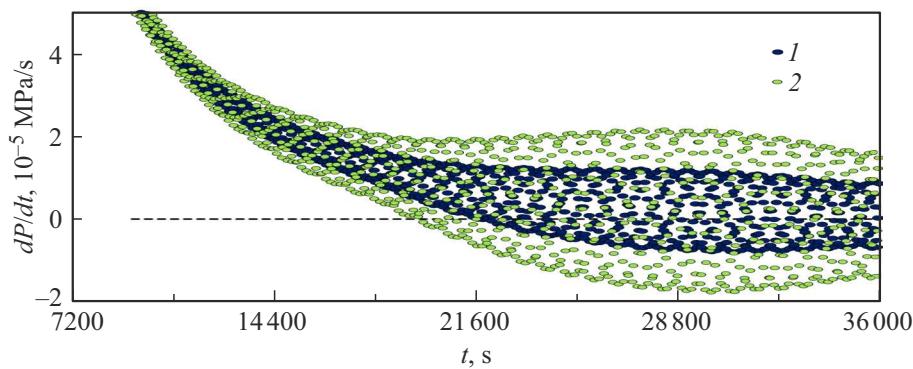


Рис. 2. График производной, восстановленной по двум (первой и второй) и трем первым сингулярным тройкам.

σ_i — сингулярное число. В дальнейшем σ_i, u_i, v_i будем называть сингулярной тройкой. Пользуясь формулой (3), вычислим значения аппроксимации производной g^k , соответствующие этой матрице. Определим фрактальную размерность ряда значений производной, вычисленного с использованием k сингулярных троек. Как известно, $D = 2 - H^k$, где H^k — индекс Херста этого ряда [3].

Шаг 3. Вычислим H^k .

Вычислим

$$\langle g^k \rangle = \frac{1}{N} \sum_{i=1}^N g_i^k, \quad S = \sqrt{\frac{1}{N} \sum_{i=1}^N (g_i^k - \langle g^k \rangle)^2},$$

$$z_m = \sum_{i=1}^m (g_i^k - \langle g^k \rangle), \quad R = \max_{1 \leq m \leq N} (z_m) - \min_{1 \leq m \leq N} (z_m),$$

$$R/S = \left(\frac{\pi}{2} N \right)^{H^k}, \quad (4)$$

отсюда $H^k = \frac{\lg(R/S)}{\lg(\frac{\pi}{2} N)}$.

Число сингулярных троек, по которым восстанавливаются траекторная матрица и производная, пробегает значения от 1 до k_0 . Значение k_0 определяется из неравенства $H^{k_0} > H^{k_0+1}$. При этом не имеет значения зависимость индекса Херста от числа отсчетов. Рассмотрим два примера применения этого алгоритма вычисления производной. В качестве первого примера возьмем измерения кривой восстановления давления, проведенные при гидродинамическом исследовании реальной скважины. При анализе и интерпретации результатов гидродинамических исследований скважин вычисляют логарифмическую производную кривой восстановления давления от времени [5]. На рис. 1, а приведена кривая восстановления давления. Индекс Херста этой кривой $H = 0.8059015$.

На рис. 1, б приведены графики производной, вычисленной по формуле (1), и логарифмической производной. Индекс Херста производной $H = 0.783470335$. Имеем 1203 отсчета измеренной функции. Используя

окно размером 602 отсчета, строим траекторную матрицу для разностной производной, находим ее сингулярное разложение и последовательно по формуле (3) восстанавливаем производную. На рис. 1, с приведены графики производной и логарифмической производной, восстановленной по первой сингулярной тройке. При этом индекс Херста восстановленной производной $H = 0.804599782$. На рис. 2 приведены графики производной, восстановленной соответственно по двум (1) и трем (2) сингулярным тройкам. На графике 1 (рис. 2) видно появление шумовой компоненты, а на графике 2 шумовая компонента усиливается. Показатели Херста кривых 1 и 2 равны 0.804431406 и 0.803925077 соответственно. Таким образом, уменьшение показателя Херста свидетельствует об усилении шума.

Максимальное значение имеет показатель Херста производной, восстановленной по первой сингулярной тройке, и это является результатом (рис. 1, с). Рассмотрим второй пример. Так как процессы фильтрации в пласте описываются линейными или слабо нелинейными параболическими уравнениями, кривые изменения давления являются гладкими. Поэтому в качестве второго примера приведем дифференцирование зашумленной осциллирующей модельной кривой из книги [6]. В качестве модельной рассматривалась задача дифференцирования функции

$$f(t) = \sin(t) + \frac{1}{100} \sin(100t) + h(t), \quad t \in [0, \pi], \quad (5)$$

2001 — отсчет функции, $h(t)$ — шум, т. е. равномерно распределенные случайные числа в диапазоне от -0.01 до 0.01.

На рис. 3, а приведены графики функции и ее производной (дополнительная ось ординат), вычисленной по формуле (1). Дифференцируем функцию по формуле (1) и, используя окно размером $\left[\frac{N+1}{2} \right]$, т. е. в данном случае в 1001 отсчет, находим траекторную матрицу и ее сингулярное разложение. Синтезируя значения производной по формуле (3) и вычисляя показатель Херста, находим производную функции по четырем сингулярным

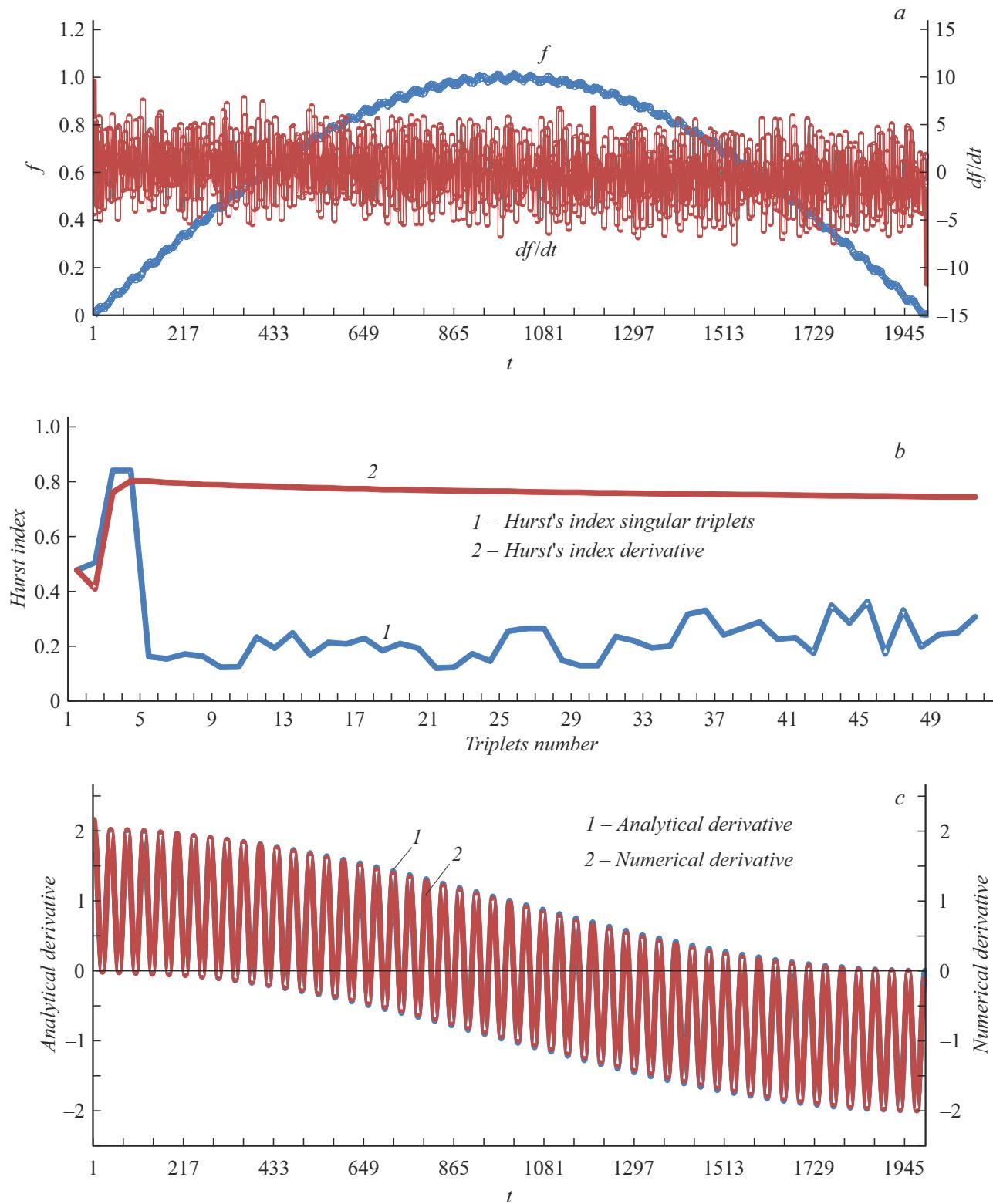


Рис. 3. *a* — модельная функция (основная ось ординат) и ее производная. *b* — показатели Херста синтезированной производной (верхняя линия) и компонент производной, вычисленных по сингулярным тройкам (нижняя линия). По оси абсцисс — номер сингулярной тройки. *c* — сравнение аналитической и разностной производной. Все величины на рисунке безразмерные.

тройкам. На рис. 3, *b* представлены графики показателей Херста в зависимости от числа сингулярных троек в аппроксимации производных (верхняя линия) и показатели Херста для каждой компоненты ряда, построенной по сингулярной тройке. На рис. 3, *c* приведены графики аналитически вычисленной производной (без шума) и разностной производной зашумленной функции. Поскольку на графике аналитическая и разностная производные практически сливаются, приведем количественные результаты оценки близости аналитической и разностной производной.

Сравним аналитическую производную модельной функции без шума и результат вычисления производной предлагаемым методом этой же, но уже зашумленной функции. Обозначим: g_i — значение аналитической производной в точке i (i -я компонента вектора), \check{g}_i — значение разностной производной в точке i (i -я компонента вектора разностной производной). Пусть w_i — вектор разности $w_i = g_i - \check{g}_i$. Тогда $|g| = 44.74371464$, $|\check{g}| = 44.26865779$, $|w| = 1.138447732$. Косинус угла между векторами производных аналитической и зашумленной кривых равен 0.999729803, т. е. угол равен 0.023246921 rad или 1.331989725°.

$$\frac{|g| - |\check{g}|}{|g|} \cdot 100\% = 1.061728687\%,$$

$$\frac{|w|}{|g|} \cdot 100\% = 2.544374649\%.$$

Таким образом, видно, что непосредственное разностное дифференцирование с последующим сингулярным спектральным анализом под контролем вычисляемого индекса Херста позволяет корректно вычислять производную зашумленной экспериментальной функции.

Финансирование работы

Работа выполнена за счет субсидии, выделенной НИЦ „Курчатовский институт“—НИИСИ на выполнение государственного задания по теме № FNEF-2024-0001 („Создание и реализация доверенных систем искусственного интеллекта, основанных на новых математических и алгоритмических методах, моделях быстрых вычислений, реализуемых на отечественных вычислительных системах“).

Конфликт интересов

Автор заявляет, что у него нет конфликта интересов.

Список литературы

- [1] R. Vautard, P. Yiou, M. Ghil, Physica D, **58**, 95 (1992). DOI: 10.1016/0167-2789(92)90103-T
- [2] Н.Э. Голяндина, *Метод „Гусеница“-SSA: анализ временных рядов* (Изд-во ВВМ, СПб., 2004).
- [3] Е. Федер, *Фракталы* (Мир, М., 1991). [J. Feder, *Fractals* (Plenum Press, 1988).]
- [4] Дж. Деммель, *Вычислительная линейная алгебра. Теория и приложения* (Мир, М., 2001). [J.W. Demmel, *Applied numerical linear algebra* (SIAM, 1997).]
- [5] R.N. Horne, *Modern well test analysis. A computer-aided approach* (Petroway, 2002).
- [6] А.Н. Тихонов, В.Я. Арсенин, *Методы решения некорректных задач* (URSS, М., 2022).