

# ПОТОКОВЫЙ АНАЛИЗ ДАННЫХ КОНТРОЛЯ КАЧЕСТВА С ЦЕЛЬЮ ВЫЯВЛЕНИЯ СБОЕВ ТЕХНОЛОГИЧЕСКОГО ПРОЦЕССА

**Р.О. РЕЗАЕВ**, канд. физ. мат. наук, генеральный директор ООО «Проектирование материалов», г. Москва, Россия; доцент ТПУ, г. Томск, Россия; науч. сотр. IFW Dresden; г. Дрезден, Германия;  
**А.А. ДМИТРИЕВ**, главный технолог ООО «Проектирование материалов», г. Москва, Россия.

*Ключевые слова: контрольная карта, статистический метод контроля, контроль качества, бетонный завод*

*Keywords: control chart, statistical control method, quality control, concrete plant*



В статье обсуждается проблема применения методов статистического контроля к анализу данных мониторинга прочности. В частности, представляет интерес определение момента времени, когда в технологическом процессе произошли изменения. К таким изменениям относятся – смена характеристик сырьевых материалов, выход из контролируемого режима работы параметров оборудования и т.д. Классические методы тестирования выборки на принадлежность данных одной функции распределения (тест Колмогорова-Смирнова и связанная с ним контрольная карта) не обладают достаточной чувствительностью к детектированию отклика в данных от небольших (~ 5-10%) изменений технологического процесса. Предлагается для этих целей использование алгоритма CUSUM (cumulative sum) – в работе приведено обоснование алгоритма, его валидация с помощью имитационного моделирования и применение к реальным производственным данным. Продемонстрировано, что применение алгоритма обеспечивает возможность детектирования моментов времени изменения параметров технологического процесса. Показано, как построение регрессионных моделей на основе гауссовских процессов помогает повысить достоверность потенциальных точек сбоя технологических процессов, обнаруженных алгоритмом CUSUM.

The paper discusses the problem of applying statistical control methods to the analysis of strength monitoring data. In particular, it is of interest to determine the point in time when changes occurred in the technological process. Such changes include – a change in the characteristics of raw materials, a deviation from a controlled mode of functioning of equipment parameters, etc. The classical methods of testing a sample for whether the data belongs to one distribution function (the Kolmogorov-Smirnov test and the associated control chart) do not have sufficient sensitivity to detect the response in the data from small (~ 5-10%) changes in the technological process. For these purposes, the use of the CUSUM (cumulative sum) algorithm is suggested – the paper provides a reasoning for the algorithm, its validation using simulation modeling and application to real production data. It has been demonstrated that the application of the algorithm provides the possibility of detecting time points of changes in the parameters of the technological process. Gaussian processes are applied to fit data and prove change point detection.

## Введение

Методы управления качеством и его контроля для бетонных смесей и железобетонных изделий (ЖБИ) занимают одно из приоритетных мест в иерархии вопросов организации серийного производства [см., например, 1-2].

Существуют различные стандарты, описывающие регламент требований для обеспечения различного уровня качества продукции. Так, большое внимание уделяется различным проблемам, связанным со статистическими методами контроля качества. В контексте производства бетонных изделий, смесей предлагается к ознакомлению, например, ГОСТ Р ИСО 12491-2011.

Статистические методы являются достаточно эффективным инструментом, обеспечивающим надежной количественной обратной связью управляющий персонал с производственной реальностью.

Основной целью внедрения таких методов в практику является контроль вариативности характеристик сырьевых материалов и параметров технологических процессов/оборудования. Любое производство — это динамический процесс, сопряженный со случайными факторами, и грамотно поставленный контроль качества должен «уметь отделять» случайные факторы от систематических. Постановка такой задачи без каких-либо дополнительных условий является стандартной для математической статистики. Однако нестандартной такая задача становится при следующей формулировке — обнаружение в данных контроля изменения систематической компоненты **за минимальное время** с момента появления этого изменения.

Суть проблемы иллюстрируется следующим практическим примером. Завод выпускает бетонные изделия со средней прочностью, например, 30 МПа. Различные случайные факторы вызывают флуктуации прочности вокруг этой величины, и динамика фактического ее значения может выглядеть, например, так — [27, 29, 33, 29, 31, 26, 32 ...] МПа. Разброс значений прочности описывается дисперсией, и устойчивая, контролируемая ситуация характеризуется постоянным средним значением и дисперсией.

Частота появления и амплитудные значения отклонений прочности от среднего значения хорошо прогнозируются, и состав

бетона можно подбирать с максимально экономичным запасом прочности.

Предположим, что в дозаторе цемента произошел небольшой сбой, который в целом не повлиял на функциональность — базово устройство решает задачу дозирования цемента в миксер, — но каждый раз дозировка производится с заниженным значением расхода цемента. Такую ситуацию на практике может спровоцировать, например, засор, блокирующий полное открытие затвора, или нарушение работы датчиков веса. Результирующая последовательность значений прочности становится следующей — [27, 28, 33, 29, 30, 26, 31 ...] МПа и, на первый взгляд, мало изменилась относительно предыдущей. Однако важно то, что систематическая компонента — среднее значение — изменилась в меньшую сторону, а дисперсия осталась прежней. Это приводит к тому, и частота и амплитудные значения отклонений также изменились, что в результате повышает вероятность такого нежелательного события как «сброс прочности». Нужно своевременно и надежно детектировать момент времени, когда произошло изменение среднего значения. Проблема состоит в том, что для оценки среднего значения необходимо иметь выборку, состоящую, как минимум, из нескольких десятков элементов, взятых из одного и того же распределения. На практике это возможно не всегда, часто имеет место ситуация, когда решение нужно принимать буквально по нескольким значениям измеряемой величины, поэтому проблема усугубляется тем, что возникающие отклонения необходимо детектировать, в том числе, и по малой выборке данных. Отметим, что существует такое решение, как ужесточение входного контроля данных, что, очевидно, позволит уменьшить дисперсию и на этом фоне гораздо отчетливее будет виден отклик в данных контроля от изменения систематических компонент даже при небольшой выборке. Чем строже контроль, тем ощутимее эффект, однако обратная сторона этого решения состоит в резком удорожании себестоимости выпускаемых изделий. Многие предприятия находятся в конкурентной борьбе по фактору цены, что в результате может ослабить их позиции. Поэтому полезным и разумным с практической точки зрения, исходя из фактических рыночных реалий, предлага-

ется решение в виде поиска оптимального баланса между, с одной стороны, допустимым разбросом характеристик сырьевых материалов и параметров технологического процесса, и, с другой стороны, достаточным запасом прочности для компенсации максимальных отклонений.

Один из вариантов реализации такого решения основан на разработке специальных контрольных карт, определяющих допустимые границы колебаний контролируемых величин [2]. В менеджменте качества такие карты часто именуется картами Шухарта.

В данной работе мы представим алгоритм построения аналога контрольной карты с целью детектирования факта изменения систематической компоненты в наблюдаемых данных за минимальное время с момента начала этих изменений.

### Контрольные карты

Контрольная карта представляет собой статистический инструмент, используемый для мониторинга и управления технологическим процессом во времени. Это графическое представление данных процесса, которое помогает идентифицировать любые отклонения или изменения параметров вероятностных законов, характеризующих его динамику в моменте.

Основная цель использования карт – подать своевременный сигнал обратной связи оператору о потенциальном сбое или поломке. Поскольку техника построения карт базируется только на статистических закономерностях, привязка к какому-либо конкретному технологическому процессу отсутствует, что позволяет их использовать в совершенно различных отраслях и задачах. Общая схема строения контрольной карты представлена на рис. 1 – определяется средняя линия значений (зеленая) и контрольные границы для отклонений (красные). Далее построенная карта применяется для потокового контроля данных процесса – измеряемая интересующая величина последовательно во времени откладывается на контрольной карте. Если наблюдаемая точка несколько раз подряд выходит за контрольные границы, это является сигналом появления сбоя в процессе.

С момента их появления в 20-х годах 20-го века были разработаны различные

типы контрольных карт, классифицирующиеся по типу случайного процесса, описывающего наблюдаемую величину. Существуют, например, контрольные карты для отслеживания непрерывных данных, таких как температура или вес, дискретных данных, таких как количество дефектов в продукте и т.д.

В свою очередь, случайные процессы подразделяются на длительные/короткие стационарные и нестационарные. Отслеживая процесс и предпринимая корректирующие действия, когда это необходимо, контрольные карты помогают гарантировать, что процесс или система работают в стабильном, управляемом режиме и соответствуют желаемым стандартам качества.

Тем не менее, хотя контрольные карты и являются мощным инструментом для мониторинга, при стандартном построении у них есть несколько недостатков, которые ограничивают их использование в практике производства бетонных смесей/изделий из бетона. К таковым, в частности, относятся:

во-первых, требование наличия достаточно большой выборки в однородных условиях для определения границ контроля, что сложно обеспечить на практике – в процессе набора статистики условия могут измениться, что исказит реальные контрольные границы.

Во-вторых, контрольные карты дают слабый отклик на хотя и относительно небольшие, но требующие учета изменения параметров технологического процесса.

В-третьих, при построении контрольных границ используется предположение о нормальном законе распределения, что далеко не всегда справедливо.

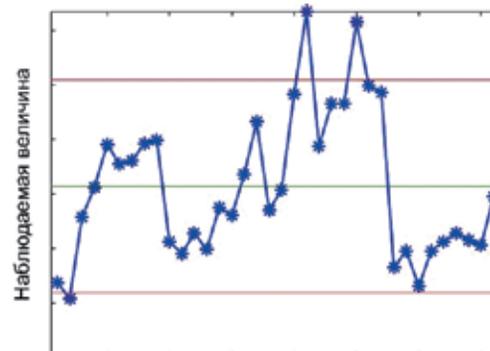


Рис. 1. Схема работы с контрольной картой – красные линии определяют граничные значения наблюдаемой величины, зеленая линия – среднее значение. Сигнал оператору подается при пересечении наблюдаемых данных красных линий.

Указанные проблемы частично можно решить, если вместо контрольных карт использовать алгоритм детектирования «точки перехода» (eng. – «change point detection») параметров вероятностной модели, порождающей данные процесса. В текущей работе мы рассмотрим простейший вариант такого алгоритма – CUSUM-тест (eng. CUmulative SUM), предложенный в 1954 г. Е.С. Пэйджем (университет Кэмбридж) [3].

### Алгоритм детектирования точки перехода CUSUM. Вероятностная модификация

Объектом действия алгоритма CUSUM является случайный временной ряд (случайный процесс) – последовательность случайных величин, индексированных моментами времени, в которые эта величина была порождена в соответствии с некоторой вероятностной моделью. Временной ряд называется стационарным в том случае, если параметры модели не зависят от времени. В противном случае ряд – нестационарный. CUSUM детектирует переход от одного стационарного ряда к другому в данных одной природы. На рис. 2 показан пример случайного временного ряда, порожденного нормальным законом распределения, параметры которого – математическое ожидание и дисперсия – в моменты времени, указанные зеленой линией, претерпевали скачкообразные изменения.

Временной ряд строится следующим образом: в каждый момент времени  $t_i$  ( $i$  – индекс времени, например, номер дня) производится единичное измерение случайной величины, плотность распределения вероятности которой описывается нор-

мальным законом. Полученное значение откладывается в соответствующей системе координат, как на рис. 2, и алгоритм детектирования точки перехода должен уловить момент времени, когда параметры нормального закона изменили свои значения. Очевидно, что избежать задержки между моментом времени, когда параметры фактически изменились, и моментом времени, когда удалось детектировать это изменение, не получится. Поэтому с практической точки зрения необходимо минимизировать эту задержку.

Общая идея алгоритма состоит в следующем. Если среднее значение временного ряда является постоянным значением равным нулю – кумулятивная сумма отсчетов сходится к случайной величине, распределенной по нормальному закону с нулевым средним значением. Соответственно, если вдруг в какой-то момент времени среднее значение окажется отличным от нулевого и превышает заранее установленное пороговое значение, то потенциально это может означать, что в наблюдаемом ряду произошло изменение параметров вероятностной модели временного ряда (возникла точка перехода). Сходимость суммы случайных величин к нормальному закону распределения является следствием центральной предельной теоремы [4]. Выбор порогового значения алгоритма определяется в соответствии со стандартными подходами тестирования гипотез в статистике [5]. Опишем последовательные шаги алгоритма.

Во-первых, поскольку в общем случае среднее значение стационарного временного ряда не равно нулю, необходимо стандартизировать наблюдаемую величину:

$$z_i = \frac{x_i - \hat{\mu}_x}{\hat{\sigma}_x}$$

где символ шляпки над величинами означает оценки среднего –  $\mu$ , и дисперсии –  $\sigma$ , вычисленных по выборке  $X = \{x_n, \dots, x_m\}$ ,  $x_i$  – наблюдаемые фактические реализации временного ряда. Тогда, в соответствии с центральной предельной теоремой, справедливо следующее соотношение:

$$S_T = \sum_{i=1}^T z_i \sim N(0, T) \Rightarrow \tilde{S}_T = \frac{1}{\sqrt{T}} S_T \sim N(0, 1).$$

Разделив кумулятивную сумму  $S_T$  на квадратный корень длины временного промежутка, по которому ведется суммирование, получается стандартное нормальное

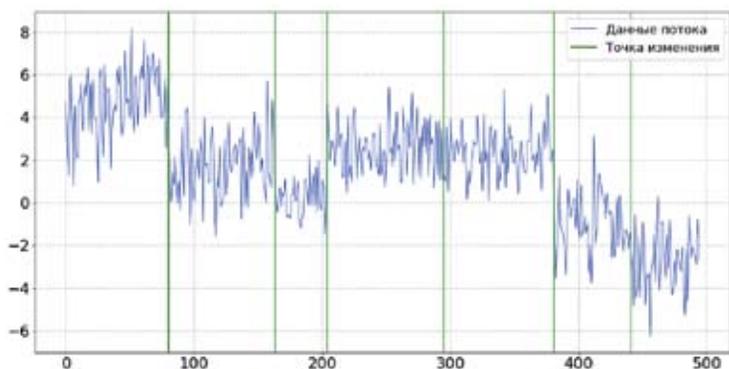


Рис. 2. Пример результата смены параметров вероятностной модели, порождающей случайный временной ряд. Зеленые линии отмечают моменты времени изменения дисперсии и среднего значения. Задача алгоритма CUSUM – определить положение зеленых линий с наименьшей задержкой.

распределение с нулевым значением среднего и единичной дисперсией. Следовательно, к полученному результату можно применять различные методики тестирования гипотез – вероятность того, что стандартизированная величина  $\tilde{S}_T$  примет значение меньше, чем заданное  $c$  (в соответствии с некоторым критерием значимости) определяется

$$F(c) \approx P(\tilde{S}_T < c)$$

функцией нормального распределения вероятности. Отметим, что в текущей формулировке алгоритм сможет детектировать только те отклонения, которые дают вклад отсчетов в отрицательную сторону относительно среднего значения. Чтобы была возможность детектировать и положительные сдвиги, необходимо рассчитать вероятность того, что

$$P(|\tilde{S}_T| \geq |c|)$$

абсолютные значения  $\tilde{S}_T$  превышают заданное критерием значимости значение. Не вдаваясь в детали математических выкладок, приведем результирующее выражение:

$$P(|\tilde{S}_T| \geq |c|) = 2(1 - P(\tilde{S}_T < |c|)) = 2(1 - F(|c|)).$$

Именно последнее выражение и является базовым в нашей работе для построения двустороннего алгоритма CUSUM, способного детектировать изменения параметров технологического процесса с заданным критерием значимости. Для каждой новой точки во временном ряде мы напрямую получаем оценку того, насколько верна гипотеза сдвига систематической компоненты, связанной с этой точкой. Как только критерий значимости пересекается, мы детектируем смену режима генерации наблюдаемой величины и подаем сигнал оператору. Далее алгоритм запускается с началом отсчета в точке смены режима.

Реализация алгоритма по шагам выглядит следующим образом:

1. Определение критерия значимости  $p_{lim} \in (0, 1)$  и объем первоначальной выборки  $N$ , необходимой для оценки среднего значения –  $T_1$  (если брать в единицах времени).

2. Сбор наблюдений  $x_t$  до тех пор, пока  $T < T_1$ .

3. В момент  $T_1$  рассчитать:

$$\hat{\mu}_x = \frac{1}{T_1} \sum_{t=1}^{T_1} x_t, \hat{\sigma}_x = \sqrt{\frac{1}{T_1 - 1} \sum_{t=1}^{T_1} (x_t - \hat{\mu}_x)^2}$$

4. Далее в каждый последующий момент времени для новой точки, рассчитывать величину  $\tilde{S}_T$  и вероятность  $p_T = 2(1 - F(|\tilde{S}_T|))$ .

5. Если в какой-то момент времени обнаруживается, что  $p_T > p_{lim}$ , зафиксировать точку перехода, обозначить начало нового стационарного процесса – идем в п. 2.

## Валидация алгоритма

Проведем имитационное моделирование для оценки параметров  $N$ ,  $p_{lim}$  алгоритма CUSUM. Будем считать, что на некотором модельном производстве прочность выпускаемого изделия описывается различными вероятностными распределениями – в период 1-23 дня – среднее и дисперсия равны 32 и 3,2 МПа (синее распределение на рис. 3), в период 24-50 дней – 34 и 3,4 МПа (зеленое распределение на рис. 3), и в период 51-79 дней – 30 и 3,0 МПа (красное распределение на рис. 3), соответственно:

$$R_{1-23} \sim \frac{1}{3,2 \times \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{R-32}{3,2}\right)^2\right),$$

$$R_{24-50} \sim \frac{1}{3,4 \times \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{R-34}{3,4}\right)^2\right),$$

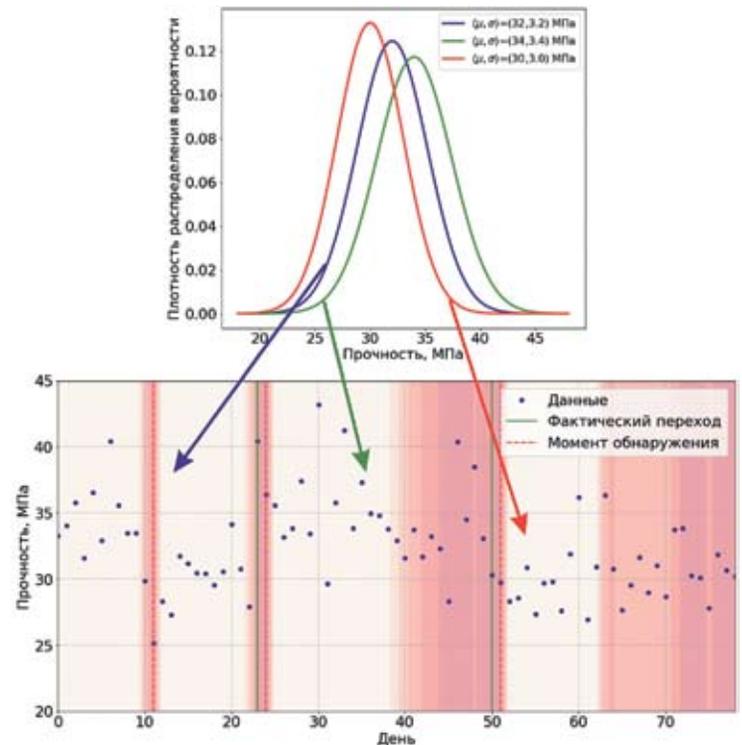


Рис. 3. Применение алгоритма детектирования изменений параметров распределения вероятности. Алгоритм CUSUM имеет следующие параметры:  $N=10$ ,  $p_{lim}=0,1$ . Интенсивность розового оттенка характеризует отклик  $p_T$  от CUSUM, до тех пор, не приблизится к пороговому значению  $p_{lim}$ .

$$R_{51-79} \sim \frac{1}{3,0 \times \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{R-30}{3,0}\right)^2\right)$$

Визуально выборка по 3-м различным распределениям на рис. 3 слабо отличается друг от друга. Однако минимальная требуемая прочность, обеспечивающая вероятность 0,95 по второй выборке – 27,2 МПа, а по третьей – 24 МПа. Фактически в период 50-79 дней можно с высокой долей вероятности утверждать, что бетон будет иметь чаще случаи «сброса прочности».

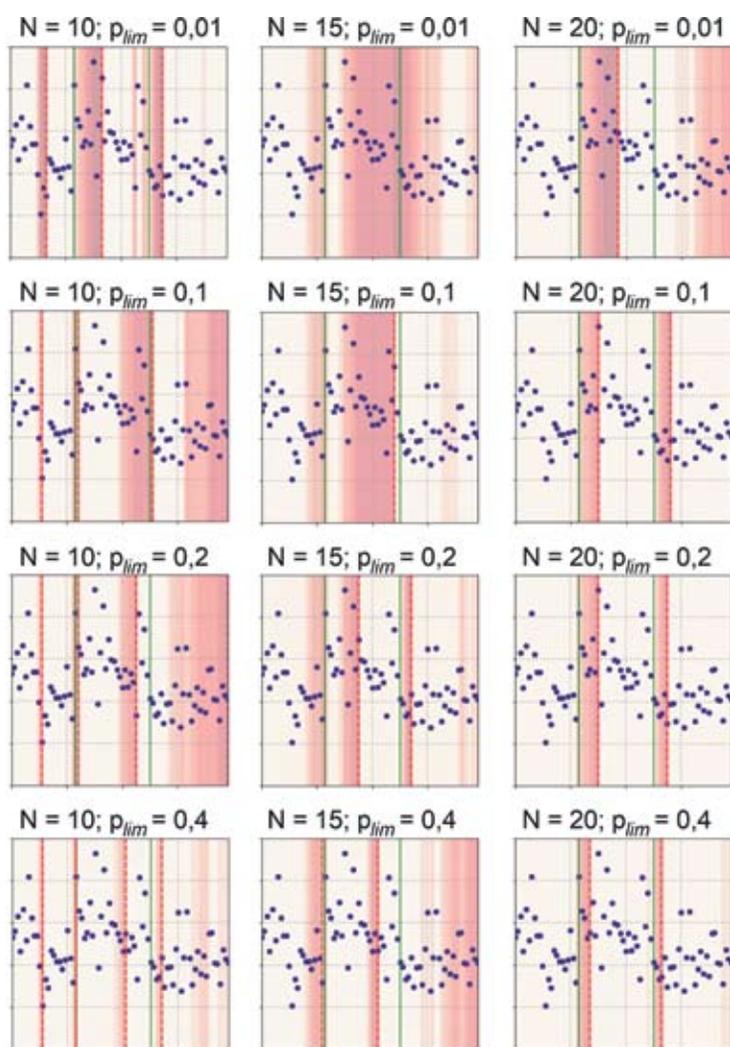


Рис. 4. Результаты работы алгоритма CUSUM с различными параметрами  $N, p_{lim}$ . Параметры  $N, p_{lim} = 10; 0,1$  обеспечивают наилучший результат.

ТАБЛИЦА 1. РЕЗУЛЬТАТЫ РАБОТЫ АЛГОРИТМА CUSUM С РАЗЛИЧНЫМИ ПАРАМЕТРАМИ  $N, p_{lim}$ .

N	День детектирования изменений												
	10				15				20				
$p_{lim}$	0,01	0,1	0,2	0,4	0,01	0,1	0,2	0,4	0,01	0,1	0,2	0,4	
Факт	23	33	24	24	-	-	48	35	-	37	30	30	27
изм-я	50	55	51	78	54	-	-	54	-	-	56	55	53
Ложный сигнал	13	11	11,45	11,23,41	-	-	-	22,42	-	-	-	-	-

Результаты работы алгоритма CUSUM показаны на рис. 3, где отмечена успешная детекция точек изменения параметров вероятностного закона спустя 1 день от этого изменения. При этом появилось одно ложное срабатывание.

Отметим, что наличие ложных срабатываний в небольшом количестве не является проблемой, в отличие от ситуации, когда был бы пропущен переход. Поэтому будем допускать, что 1-2 ложных срабатывания являются допустимой величиной.

На рис. 4 приведены результаты работы CUSUM на данных рис. 3 с различными параметрами  $N, p_{lim}$  с целью определить их наилучшую комбинацию. Результаты рис. 4 для удобства представлены в табл. 1, из которой видно, что наилучшая комбинация  $N, p_{lim} = 10; 0,1$ . С найденными параметрами применим данный алгоритм к реальной производственной ситуации.

### Практическое применение к производственным данным

В табл. 2 представлены производственные данные ежедневного контроля прочности в период 1-82 дней. В 27-й день в составах поменялся расход цемента – уменьшили на 20 кг, далее на 61-й день изменился режим ТВО – длительность прогрева была увеличена на 2 часа. В табл. 3 по каждому из периодов приведены статистические оценки средней прочности и ее дисперсии. Далее в табл. 4 показаны результаты применения к данным таблицы 2 алгоритма CUSUM в последовательном режиме.

Графический анализ данных и результаты работы алгоритма CUSUM представлены на рис. 5. Зелеными линиями отмечены дни фактических изменений, произошедших на производстве, красными пунктирными линиями отмечены моменты времени, когда CUSUM обнаружил отклик от указанных изменений.

Несмотря на определенную задержку – смена расхода цемента была зафиксиро-

ТАБЛИЦА 2. ДАННЫЕ ЕЖЕДНЕВНОГО КОНТРОЛЯ ПРОЧНОСТИ

Дни	1	2	3	4	5	6	7	8	9	10	11	12	13	14
R <sub>28</sub> , МПа	49,2	53,3	54,3	49,9	54,6	51,6	58,8	54,6	46,2	44,2	49,7	48,7	36,0	59,3
15	16	17	18	19	20	21	22	23	24	25	26	27		
53,6	58,4	45,3	47,7	53,9	51,3	41,7	53,5	57,1	56,0	57,9	56,6	59,6; поменялся расход цемента		
28	29	30	31	32	33	34	35	36	37	38	39	40	41	42
49,5	46,3	42,6	44,2	44,6	45,8	48,9	40,7	45,1	51,9	45,2	37,9	45,0	53,6	55,6
43	44	45	46	47	48	49	50	51	52	53	54	55	56	57
56,6	55,2	44,6	45,9	45,7	45,0	47,5	52,4	53,0	53,8	45,4	52,1	48,3	50,6	43,2
58	59	60	61			62	63	64	65	66	67	68	69	70
40,9	44,4	50,9	48,9; изменился режим ТВО			48,3	48,0	40,8	46,7	38,4	39,0	50,8	40,1	45,8
71	72	73	74	75	76	77	78	79	80	81	82			
47,2	42,7	47,4	42,7	53,7	47,6	53,0	55,8	46,4	51,5	48,6	46,0			

ТАБЛИЦА 3. СТАТИСТИЧЕСКИЕ ОЦЕНКИ СРЕДНЕГО И ДИСПЕРСИИ В РАЗЛИЧНЫЕ ПЕРИОДЫ

Период дней с 1-го по 27-й		Период дней с 28-го по 61-й		Период дней с 62-го по 82	
среднее, МПа	ср.кв.откл., МПа	среднее, МПа	ср.кв.откл., МПа	среднее, МПа	ср.кв.откл., МПа
51,0	5,8	47,1	4,7	46,6	4,8

рована на 7-й день, а смена режима ТВО – на 9-й, тем не менее, алгоритм успешно справился с поставленной задачей. На 48-й день было сгенерировано ложное событие, хотя, поскольку данные производственные, нельзя исключать, что в этот момент также было какое-то изменение, не учтенное в табл. 2. Результат текущей работы имеет практическое значение и представляет интерес для развития данной методики к различным производственным данным.

### Регрессия на основе гауссовского процесса (кригинг)

Использование гауссовских процессов для построения регрессионной модели набирает все большую популярность в последнее время [6]. Данный метод обладает тем преимуществом, что позволяет, во-первых, отказаться от конкретного вида аналитических зависимостей для аппроксимации данных, а, во-вторых, учитывать вероятностную природу данных при решении проблемы прогнозирования. При регрессии наблюдаемые данные рассматриваются как элементы выборки N-мерного гауссовского распределения, где N – количество элементов выборки:

$$f(x') \sim N(\mu, K(\theta, x, x'))$$

где  $\mu$  – математическое ожидание,  $\theta$  – набор гиперпараметров ядра  $K(\theta, x, x')$ ,  $x'$  – независимые переменные, отвечающие

ТАБЛИЦА 4. ФАКТИЧЕСКИЕ ИЗМЕНЕНИЯ И СПОСОБНОСТЬ CUSUM ИХ ОБНАРУЖИТЬ

Фактическое изменение, день	27	61	Ложное срабатывание
Сигнал детектирования, день	34	72	48

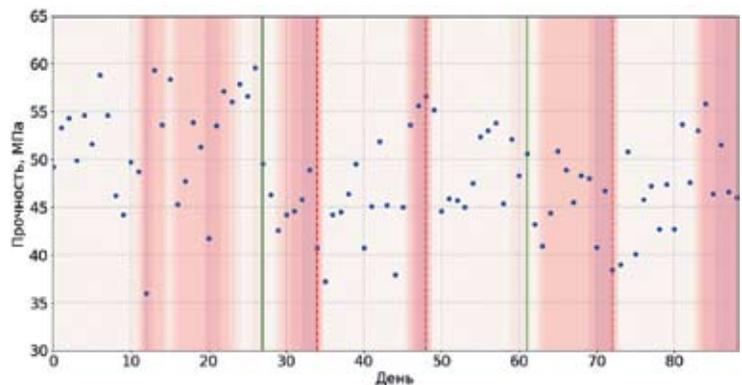


Рис. 5. Результат применения алгоритма CUSUM к производственным данным табл. 2.

наблюдаемым точкам, задающего ковариационную матрицу. Количественные значения гиперпараметров находятся исходя из максимизации функции правдоподобия  $\mathcal{L}(\theta | y) = \prod p(y = Y_i | \theta = \{\mu, \sigma\})$ . Конкретный вид ядра ковариационной матрицы задается исходя из экспоненциального класса:

$$K(\theta, x, x') = \exp\left(-\frac{(x - x')^2}{2\theta^2}\right).$$

Параметр  $\theta$  в выбранном в таком виде ядре отражает степень корреляции между «соседними» точками  $x$ . В общем случае его физическая интерпретация отсутствует,

однако в нашем частном случае наиболее наглядной является интерпретация  $\theta$  как «длины корреляции», т.е. промежуток времени в течение которого в наблюдаемых величинах можно проследить взаимную корреляцию.

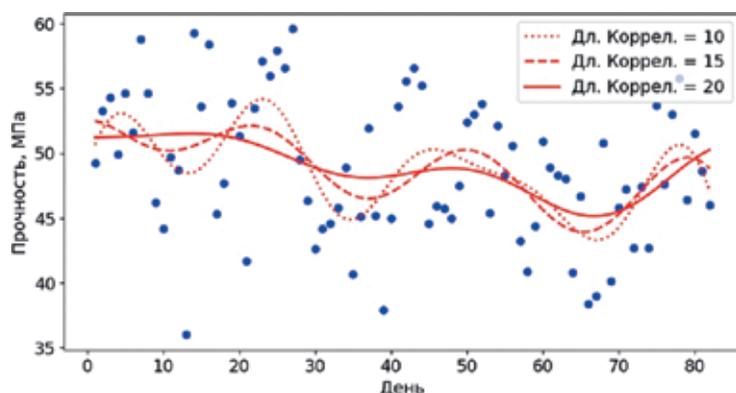


Рис. 6. Регрессия производственных данных таблицы 2 на основе гауссовских процессов.

На рис. 6 представлены три регрессионные модели на основе гауссовского процесса с длиной корреляции 10, 15 и 20 дней, оценки параметра математического ожидания  $\mu = 32; 34,6; \text{ и } 36,4$  МПа, соответственно. Наиболее «сглаженной» является зависимость, соответствующая длине корреляции 20 дней — области перехода отражаются как монотонно убывающие или возрастающие участки, которых для данной зависимости 3: с 20 по 30-й день, с 50-го по, ориентировочно, 65-й, и последний — с 65-го до конца. При этом положение этих участков находится в согласии с регрессионными моделями, построенными для длин корреляции 10 и 15 дней. Таким образом, анализ регрессионных моделей показывает, что в интервале с 20-го по 30-й и с 50-го по 65-й дни, скорее всего, произошли какие-то изменения параметров технологического процесса, что и было в действительности (см. табл. 2). Анализируя потоковые случайные величины экспресс методами, например, алгоритмами CUSUM, можно извлекать потенциально значимую информацию, а проводя дальнейшую сегментацию данных посредством регрессионных моделей на базе гауссовского процесса — повышать достоверность сделанных выводов.

## Заключение

Проблемы нестабильности характеристик сырьевых материалов и сбоев в технологических процессах являются одними из самых актуальных в производственной практике. Как следствие, свойства выпускаемой продукции также будут нестабильными и постоянно отклоняться от средних значений. В таких случаях необходимо по наблюдаемым значениям этих свойств определить момент, когда эти отклонения выйдут из-под контроля, что критически повышает риск серьезного брака. На практике это означает, что либо в параметрах технологического процесса произошел серьезный сбой, либо характеристики сырьевых материалов кардинально изменились. В обоих случаях требуется внешнее вмешательство для устранения причин, вызвавших сбой. Алгоритм CUSUM предназначен для решения подобных задач. В данной работе показан пример того, каким образом этот алгоритм можно успешно применять для работы с производственными данными по выпуску бетонных смесей/изделий. В частности, по данным реальной выборки были определены фактически случившееся изменения в технологическом процессе — уменьшение расхода цемента и изменение режима ТВО.

## Библиографический список

1. Батяновский Э.И. *Технология производства железобетонных изделий*. — Минск, 2018, уч. пособие. — 304 с.
2. Касторных Л.И., Трищенко И.В., Гикало М.А. *Контроль и оценка прочности бетона на заводах сборного и товарного бетона // ИВД, 4(27), 2013.*
3. Page E.S. *Continuous inspection scheme // Biometrika, 41(1/2), 1954, p. 100-115.*
4. Кибзун А.И., Горяинова Е.Р., Наумов А.В., Сиротин А.Н. *Теория вероятностей и математическая статистика. Базовый курс с примерами и задачами // М.: Физматлит, 2002. — 224 с.*
5. Гателюк О.В., Манюкова Н.В. *Проверка статистических гипотез // Лань, 2022. — 112 с.*
6. V.L. Deringer, A.P. Bartok, N. Bernstein, D.M. Wilkins, M. Ceriotti, and G. Csanyi. *Gaussian Process Regression for Materials and Molecules, Chem. Rev., 2021, 121, 16, 10073-10141.*