



УДК 528.02+543.0

Читайте и узнаете:

- продолжительность периода оценки характеристик измерения;
- что является важнейшим способом применения контрольных карт.

**Ключевые слова:**

измерения, стабильность, оценка, контрольный материал, контрольная карта, погрешность

# Ведение контроля стабильности измерений

*В.И. Дворкин, главный научный сотрудник Института нефтехимического синтеза РАН, профессор, д-р хим. наук*

Важнейшим способом внутрилабораторного контроля качества измерений является оценка стабильности. Ранее нами были описаны модель процесса измерений, положенная в основу контроля стабильности, его основные принципы и организационные аспекты [1]. В настоящей статье рассмотрено, как проводится такой контроль. При этом все сказанное справедливо для любых измерений, хотя в основополагающем стандарте ГОСТ Р ИСО 5725-6-2002<sup>1</sup> все примеры взяты из химического анализа.

**К**онтроль стабильности *при наличии контрольного материала* — самый распространенный вариант. Если таких материалов несколько, контроль ведется аналогично по каждому из них. В этом варианте вначале в пределах некоторого количества серий проводится анализ (исследование) контрольного материала (материалов) для оценки реальных статистических характеристик аналитической методики. Мы будем называть эту стадию «период оценки характеристик». По окончании этого периода в случае соответствия погрешностей анализа предъявляемым к нему требованиям переходят к собственно контролю. При этом, насколько это возможно, отдельно

контролируются разные компоненты погрешности.

Подчеркнем, что *контроль стабильности, как правило, ведется по контрольным картам, контрольные пределы для которых оцениваются экспериментально в самой лаборатории.*

Иногда в лабораториях в качестве контрольных пределов пытаются использовать нормативы, взятые из методик, что близко к подходу ГОСТ Р 50779.43–99 (ИСО 7966–93)<sup>2</sup> с так называемыми приемочными картами, используемыми на производстве при приемке продукции *в дополнение к контрольным картам при выполнении ряда предположений. Контрольные карты более эффективны для целей выявления выхода процесса из подконтрольного состояния.*

### Период оценки характеристик

Это период, за который проводится достаточно много серий измерений контрольного материала. По его результатам:

- оценивают реальные характеристики процесса измерений;
- убеждаются, что они соответствуют предъявляемым требованиям (верификация).

При этом число измерений в серии должно быть таким же, как в дальнейшем при собственно контроле. Выбор продолжительности периода характеристик представляет собой компромисс между желанием получить возможно более точные оценки и необходимостью быстрее перейти к собственно контролю. Точные рекомендации дать невозможно, однако в большинстве случаев считается достаточным иметь результаты 15–20 серий измерений. Полученные в результате метрологические характеристики используются в дальнейшем для построения контрольных карт и собственно контроля.

Обозначим результаты, полученные в период оценки характеристик, как  $X_{ij}$ , где  $j=1,2, \dots, M$  — номер серии, а  $i=1,2, \dots, N$  — номер измерения. Обычно  $M=20-25$ ,  $N=2$  (табл. 1). Точки в нижней строке таблицы говорят о том, что после набора  $M$  серий, то есть окончания

<sup>1</sup> ГОСТ Р ИСО 5725-6-2002 «Точность (правильность и прецизионность) методов и результатов измерений. Часть 6. Использование значений точности на практике» введен в действие Постановлением Госстандарта России № 161-ст от 23.04.2002 г.

<sup>2</sup> ГОСТ Р 50779.43–99 (ИСО 7966-93) «Статистические методы. Приемочные контрольные карты» введен в действие Постановлением Госстандарта России № 686-ст от 23.12.1999 г.

периода оценки характеристик, начинается собственно контроль стабильности, причем данные имеют такую же структуру.

## Фильтрация результатов измерений в ходе оценки характеристик

Часто перед собственно оценкой характеристик выявляют и отбрасывают грубые промахи. При этом используются два основных статистических теста: критерии Граббса и Кохрена.

Критерий Граббса предназначен для выявления грубого промаха в совокупности результатов измерений, имеющих простейшую структуру  $X_1, X_2, \dots, X_p, \dots, X_M$ . В применении к нашим данным (табл. 1)  $X_j = \sum_{i=0}^N X_{ji} / N$ . Пусть величины  $X_j$  расположены по возрастанию. Тогда критерий Граббса рассчитывается как  $G = (X_M - \bar{X}) / s$ , если подозрительно выделяется  $X_M$ , либо как  $G = (\bar{X} - X_1) / s$ , если подозрительно выделяется  $X_1$ . Если  $G > G_{\text{крит}}(\alpha, M)$ , где  $G_{\text{крит}}(\alpha, M)$  — критическое значение статистики Граббса при вероятности ошибки 1-го рода  $\alpha$ , то подозрительно выделяющееся значение считают грубым промахом и удаляют из дальнейших расчетов. Чаще всего  $\alpha$  выбирают равной 0,05 или 0,01.

Критерий Кохрена используется для выявления серий, разброс результатов в которых аномально высок. Предполагается, что число измерений в сериях одинаковое (как это и должно быть). Рассчитывают стандартные отклонения  $s_j$  для каждой  $j$ -й серии и находят максимальное из них  $s_{\text{max}}$ . Рассчитывают величину

$$C = \frac{s_{\text{max}}^2}{s_1^2 + s_2^2 + \dots + s_{\text{max}}^2}. \quad (1)$$

Если найденная величина превышает процентную точку распределения Кохрена при принятой вероятности ошибки первого рода  $\alpha$  (обычно 0,05), то серию с макси-

Табл. 1. Результаты, полученные в период оценки характеристик

Номер серии	Номер измерения					
	1	2	...	i	...	N
1	$X_{11}$	$X_{12}$	...	$X_{1i}$	...	$X_{1N}$
2	$X_{21}$	$X_{22}$	...	$X_{2i}$	...	$X_{2N}$
...	...	...	...	...	...	...
j	$X_{j1}$	$X_{j2}$	...	$X_{ji}$	...	$X_{jN}$
...	...	...	...	...	...	...
M	$X_{M1}$	$X_{M2}$	...	$X_{Mi}$	...	$X_{MN}$
...	...	...	...	...	...	...

мальным стандартным отклонением  $s_{\text{max}}$  отбрасывают. Таблицы критических значений для критериев Граббса и Кохрена имеются, например, в ГОСТ 5725-2-2005.

Оба критерия предполагают нормальное распределение результатов как внутри серий, так и между ними. В общем-то можно проводить фильтрацию последовательно несколько раз (хотя существуют и специальные критерии для отфильтровывания нескольких величин), однако надо очень хорошо подумать, прежде чем отбрасывать более чем 5 % результатов — это может привести к чрезмерному цензурированию выборки и нереальным оценкам метрологических характеристик. Во многих случаях, если какую-либо серию отбрасывают, вместо нее выполняют еще одну.

Если по завершении оценки метрологических характеристик была проведена фильтрация данных по критериям Граббса и Кохрена и затем корректно отслеживалась стабильность по контрольным картам средних и расхождений (размахов) (см. ниже), в дальнейшем по этим критериям промахи обычно не выявляют — «плохие» результаты отбрасываются в ходе контроля.

## Оценка повторяемости (сходимости)

Естественный, не требующий дополнительных измерений способ оценки повторяемости (сходимости) в данном случае следующий. Стан-

дартное отклонение повторяемости  $s_r$  рассчитывают по формуле

$$s_r = \sqrt{\frac{\sum_{j=1}^M \sum_{i=1}^N (X_{ji} - \bar{X}_j)^2}{M(N-1)}}, \quad (2)$$

где  $\bar{X}_j = \sum_{i=1}^N X_{ji} / N$ . При  $N=2$  расчеты можно проводить по формуле

$$s_r = \sqrt{\frac{\sum (X_{j,1} - X_{j,2})^2}{2M}}, \quad (3)$$

где  $X_{j,1}$  и  $X_{j,2}$  — результаты, полученные в  $j$ -й серии.

Для верификации полученную величину  $s_r$  сравнивают с нормативом  $\sigma_r$  — приписанным значением стандартного отклонения повторяемости (сходимости), почти всегда приводимым в методике измерений, по формуле

$$\frac{S_r^2}{\sigma_r^2} < \chi^2(\bar{P}, f) / f, \quad (4)$$

где  $\chi^2(\bar{P}, f)$  есть  $\bar{P} = 1 - \alpha$  квантиль  $\chi^2$  — распределения с числом степеней свободы  $f = M(N-1)$ . В отсутствие специальных требований  $\alpha$  принимают равной 0,05, соответственно доверительная вероятность  $\bar{P} = 1 - \alpha = 0,95$ .

В ГОСТ Р ИСО 5725-6-2002 рекомендован другой способ оценки повторяемости (сходимости) — в одной серии измерений. Для такой оценки также необходим контрольный материал. Это может быть

достаточно однородная рутинная проба, стандартный образец состава и так далее. Этот материал анализируют не менее чем  $N=8-10$  раз в одной аналитической серии и находят оценку стандартного отклонения сходимости  $s_r$  как среднеквадратичное отклонение полученных результатов  $X_1, X_2, \dots, X_r, \dots, X_N$ :

$$s_r = \sqrt{\frac{\sum (X_i - \bar{X})^2}{N-1}}, \quad (5)$$

где  $\bar{X} = \sum X_i / N$ . Далее найденную величину  $s_r$  сравнивают с метрологической характеристикой  $\sigma_r$  по формуле (4). Такой способ плохо сочетается с контролем стабильности, при котором контрольный материал исследуется в нескольких сериях и в серии обычно выполняется по два измерения. Естественно и менее трудоемко выполнять оценку без «лишних» измерений по формулам (2, 3). В специально организованной серии оценка неизбежна, если контроль стабильности не ведется вообще, либо в серии выполняется лишь одно измерение.

Интересно, что проверка по формуле (4) предполагает «презумпцию невиновности» лаборатории: величина  $s_r$ , слегка превышающая норматив  $\sigma_r$ , может оказаться удовлетворительной.

#### Оценка промежуточной прецизионности

Имея данные, приведенные в табл. 1, стандартное отклонение промежуточной прецизионности  $s_{IP}$  можно рассчитать по формуле

$$s_{IP} = \sqrt{\frac{\sum (\bar{X}_j - \bar{\bar{X}})^2}{M-1}} + \left(1 - \frac{1}{N}\right) s_r, \quad (6)$$

где  $\bar{X}_j = \sum_{i=1}^N X_{ji} / N$ ,  $\bar{\bar{X}} = \sum_{j=1}^M \bar{X}_j / M$ .

Норматив для промежуточной прецизионности в методиках практически всегда отсутствует. Можно, однако, предложить количественный тест, а именно: если  $s_{IP} > 1,5s_r$ , то с высокой вероятностью разброс результатов от серии к серии чрезмерно велик, и желательно принять меры по его уменьшению. И уж, конечно, должно выполняться условие  $s_{IP} < \sigma_R$ , где  $\sigma_R$  — норматив для стандартного отклонения воспроизводимости, в противном случае необходимо принять меры к снижению промежуточной прецизионности.

#### Оценка правильности

Как и повторяемость (сходимость), правильность может быть оценена как в нескольких сериях (в ходе контроля стабильности), так и в одной специально организованной серии. Если ведется контроль стабильности, правильность естественно оценивать в нескольких сериях. Но это возможно только в случае, когда используется контрольный материал с известным принятым опорным значением: стандартный образец либо аттестованная смесь; образец, предварительно исследованный в авторитетной лаборатории и т.д.

Оценка правильности в ходе контроля стабильности с использованием СО либо образца сравнения. Проведя измерения в  $M$  сериях, проверяют гипотезу об отсутствии систематической погрешности следующим образом. Рассчитывают

$$t = \left| \bar{\bar{X}} - X_a \right| / \sqrt{\frac{s_{\bar{X}_j}^2}{M} + \frac{(\Delta X_a)^2}{3}}, \quad (7)$$

где  $t$  — критерий Стьюдента,  $X_a$  и  $\Delta X_a$  — соответственно, принятое

для СО опорное значение и его погрешность (неопределенность)<sup>3</sup>,

$$s_{\bar{X}_j} = \sqrt{\frac{\sum (\bar{X}_j - \bar{\bar{X}})^2}{M-1}}. \quad (8)$$

Если  $t > t(\bar{P}, f)$ , при выбранном уровне доверительной вероятности  $\bar{P}$  (обычно 0,975 для двустороннего критерия) и числе степеней свободы  $f=M-1$ , то гипотезу отвергают, что означает наличие систематической составляющей погрешности. В этом случае можно оценить полную систематическую погрешность  $\Delta_c = \bar{\bar{X}} - X_a$  и нижнюю границу доверительного интервала для нее:

$$\Delta_{c,n} = \left| \bar{\bar{X}} - X_a \right| - \frac{s_{\bar{X}_j} t(\bar{P}, f)}{\sqrt{M}} - \Delta X_a. \quad (9)$$

Если в методике указана максимально допустимая величина систематической погрешности  $\Delta_{c,max}$ , то необходимо выполнение неравенства

$$\Delta_{c,n} < \Delta_{c,max}. \quad (10)$$

В противном случае систематическая погрешность превышает допустимую, и необходимо принять меры по ее уменьшению. То же необходимо сделать, если величина  $\Delta_{c,max}$  в методике не приведена.

Случай, когда контроль стабильности проводится с помощью образца, предварительно проанализированного в другой (авторитетной) лаборатории, рассмотрен нами в [2]. Там же описаны способы проверки правильности в отдельном, специально организованном эксперименте.

Проверка соответствия метрологических характеристик, достигнутых лабораторией, нормативным (верификация) проводится по окончании периода оценки характеристик. Однако понятно,

<sup>3</sup> Конечно, всегда лучше использовать такой СО, погрешностью аттестации которого можно пренебречь по принципу пренебрежения малыми погрешностями. Однако это не всегда возможно.

что в реальности опасно дожидаться этого момента — можно окануться «у разбитого корыта», если какая-то характеристика окажется неудовлетворительной. Поэтому естественно проверять качество анализа несколько раз по мере набора данных, например после 4–5, 8–10 и т.д. серий, и продолжать работать только при удовлетворительных результатах. В противном случае набор данных останавливают и принимают меры по повышению качества измерений.

После завершения периода оценки характеристик с удовлетворительными результатами переходят собственно к контролю стабильности с использованием для построения контрольных карт оценки метрологических характеристик, полученных за этот период.

## Контроль стабильности с помощью контрольных карт при наличии контрольного материала

Собственно контроль стабильности предназначен для выявления случаев выхода процесса измерений из статистически контролируемого состояния. Для этого раздельно и разными способами (ведя несколько контрольных карт) фиксируют точность, повторяемость (сходимость) и промежуточную прецизионность результатов анализа. Раздельный контроль позволяет легче выявлять источники погрешностей.

Основным инструментом контроля стабильности являются контрольные карты, впервые предложенные Шухартом [3] и примененные в химическом анализе авторами работы [4]. Контрольные карты — это графики, на которых

по горизонтальной оси откладывают порядковый номер серии измерений (либо дату исследования), а по вертикальной — результаты измерения какого-либо показателя качества исследований. На них в виде горизонтальных прямых наносят «пределы», выход за которые свидетельствует о проблемах со стабильностью процесса измерений. Можно сказать, что на ось ординат контрольных карт наносят результаты измерений величин, которые были бы постоянными в отсутствие погрешностей измерений.

При ведении контрольных карт требуется, чтобы от серии к серии менялись одни и те же условия. Чаще всего изменяются факторы «время» и «калибровка», однако в крупных лабораториях могут меняться также факторы «лаборант» и «оборудование».

Важнейший способ использования контрольных карт — их разглядывание: глазом часто удается выявить изменения процесса измерений, которые с трудом обнаруживаются при статистическом анализе. Однако при рутинной работе из-за нехватки времени больше опираются на формальные признаки («правила»). Можно выделить два вида таких правил.

*Тревожные (или предупредительные) признаки (или критерии)* соответствуют доверительной вероятности выхода процесса из-под контроля  $\bar{P} \approx 0,95$ . Появление тревожного признака (признаков) не требует остановки измерений для устранения проблем: они призваны привлечь внимание исполнителей и стать сигналом для принятия профилактических мер.

*Контрольные признаки (контрольные критерии, границы регулирования)*

свидетельствуют о выходе процесса измерений из подконтрольного состояния с  $\bar{P} > 0,99$ . При их появлении рекомендуется приостановить анализ, выявить и устранить причины сбоя и повторно исследовать пробы из последней, неудачной серии измерений.

Сведения о тревожных и контрольных признаках можно найти также в ГОСТ Р ИСО 7870-2–2015<sup>4</sup>.

Наиболее эффективны и чаще всего используются контрольные карты:

- средних значений Шухарта;
- размахов (расхождений);
- кумулятивных сумм (*CUSUM*);
- текущих расхождений;
- оценок систематической погрешности;
- экспоненциально взвешенных движущихся средних (*EWMR*).

**Контрольные карты средних значений Шухарта** используются для выявления любых видов статистически значимых проявлений нестабильности (разовых выбросов, смещений, дрейфов и т.д.). На этих картах по оси абсцисс откладывают номер серии (дату анализа), а по оси ординат — средние значения, полученные для контрольного материала в  $j$ -й серии  $\bar{X}_j$ . В виде горизонтальных линий наносят общее среднее значение  $\bar{\bar{X}}$ , а также «контрольные пределы» ( $\bar{\bar{X}} \pm 1s$ ), ( $\bar{\bar{X}} \pm 2s$ ), ( $\bar{\bar{X}} \pm 3s$ ), где  $s$  рассчитывается по формуле (10). Величины  $\bar{\bar{X}}$  и  $s$  оцениваются по результатам, полученным в период оценки характеристик (см. выше) и уточняются после набора очередных 20–30 серий. Чаще всего в каждой серии проводится два измерения для каждого контрольного материала, то есть при расчете  $\bar{X}_j$  для каждой  $j$ -й серии усредняются два результата измерений.

<sup>4</sup> ГОСТ Р ИСО 7870-2–2015 «Статистические методы. Контрольные карты. Часть 2. Контрольные карты Шухарта» введен в действие Приказом Росстандарта № 1469-ст от 06.10.2015 г.

В оригинальной работе Шухарта был предложен один тревожный  $1(2s)$  и один контрольный признак  $1(3s)$ , то есть выход результата последнего контрольного измерения за границы  $(\bar{X} \pm 2s)$  и  $(\bar{X} \pm 3s)$  соответственно. Однако быстро выяснилось, что этого недостаточно. Ниже изложены правила Вестгарда [5], рекомендуемые при работе с контрольными картами Шухарта.

В качестве «тревожных», указывающих на возможность выхода процесса измерений из-под контроля, рассматриваются следующие признаки:

-  $1(2s)$  — последняя точка лежит выше контрольного предела  $(\bar{X} + 2s)$  или ниже  $(\bar{X} - 2s)$ ;

-  $2(1s)$  — две или три последовательных точки, включая последнюю, лежат выше контрольного предела  $(\bar{X} + 1s)$  или ниже  $(\bar{X} - 1s)$ ;

-  $7\bar{X}$  — семь или более последовательных точек, включая последнюю, лежат выше (или ниже) среднего значения  $\bar{X}$ ;

-  $4D$  — каждая из четырех и более последовательных точек, включая последнюю, лежит выше (положительный дрейф) или ниже (отрицательный дрейф) предыдущей.

«Контрольные признаки» рассматриваются в случае появления тревожного признака  $1(2s)$ :

-  $1(3s)$  — последняя точка лежит выше контрольного предела  $(\bar{X} + 3s)$  или ниже  $(\bar{X} - 3s)$ ;

-  $2(2s)$  — две последовательных точки, включая последнюю, лежат выше контрольного предела  $(\bar{X} + 2s)$  или ниже  $(\bar{X} - 2s)$ ;

-  $D(4s)$  — разность между последней и предыдущей точками по абсолютной величине превышает  $4s$ ;

-  $4(1s)$  — четыре последовательных точки, включая последнюю, лежат выше контрольного предела  $(\bar{X} + 1s)$  или ниже  $(\bar{X} - 1s)$ ;

-  $10\bar{X}$  — десять последовательных точек, включая последнюю, лежат выше (или ниже) среднего значения  $\bar{X}$ .

Если хотя бы один из этих признаков выявлен, то процесс измерений вышел из-под контроля. Необходимо повторить последние исследования после выявления и устранения источника выявленной погрешности.

Алгоритм (см. рисунок), часто называемый «правилами Вестгарда», имеет ряд преимуществ:

- эффективное выявление даже не слишком больших систематических погрешностей;

- обеспечение малой вероятности ошибочного заключения о выходе измерения из-под контроля;

- облегчение поиска источника неприятностей, если выявлены признаки  $1(3s)$  или  $D(4s)$  — произошел скачок, какое-то событие типа «лопнула пружина дозатора». Выявление же событий  $4(1s)$  или  $10\bar{X}$ , вероятно, говорит о дрейфе или о небольшой по величине постоянной систематической погрешности (например, испаряются градуировочные растворы).

Однако это лишь подсказка; устранение причин выхода процесса измерений из-под контроля требует квалификации и опыта.

Контрольные карты средних значений Шухарта наиболее универсальны, они помогают выявить все виды погрешностей. Однако используются и другие контрольные карты, более эффективные для выявления конкретных видов погрешности.

**Контрольные карты размахов (расхождений)** предназначены для контроля стабильности стандартного отклонения повторяемости (сходимости)  $s_r$ .

Такой контроль возможен только в случае, когда в каждой серии (в условиях повторяемости согласно ГОСТ 5725–2002, ч. 1–6) выполняется два или более измерения для контрольного материала. На контрольные карты размахов (расхождений) на горизонтальную ось наносятся номера серии  $j$ , а на вертикальную — абсолютные величины размахов  $X_{j,max} - X_{j,min}$ , то есть разница между максимальным и минимальным результатами измерений контрольного материала, выполненными в  $j$ -й серии. Чаще всего в серии выполняют два измерения, откладывая

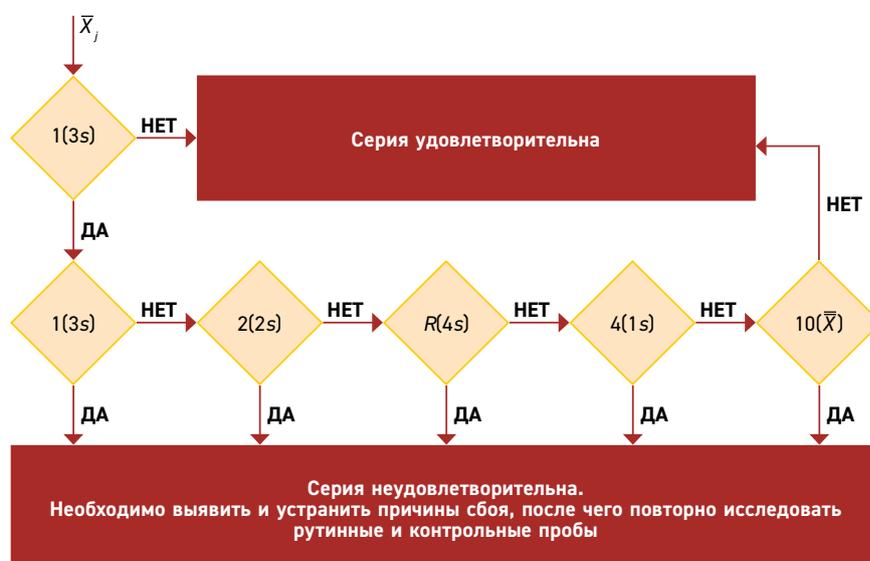


Рисунок. Правила Вестгарда

по оси ординат величины  $|X_1 - X_2|$ . Число измерений в серии может быть и больше (что повышает надежность контроля), однако и стоимость контроля при этом растет<sup>5</sup>. На карту расхождений (размахов) в виде горизонтальных линий наносят среднюю линию и верхние контрольные пределы «+2s» и «+3s» («предел предупреждения» и «предел действия»). Эти величины рассчитывают, умножая оценку стандартного отклонения  $s_r$  (см. выше) на коэффициенты, приведенные в табл. 2<sup>6</sup>. Отсюда понятно, что названия «+2s» и «+3s» не имеют прямого отношения к стандартному отклонению и говорят только о том, что эти линии играют ту же роль, что и контрольные пределы 2s и 3s на контрольных картах средних значений Шухарта.

Как «тревожный признак» для карт расхождений (размахов) обычно рассматривают правило  $R(2s)$  — расхождение (размах) в последней (оцениваемой) серии больше верхнего предела предупреждения «+2s». В качестве «контрольных признаков» для карт размахов (расхождений) используются правила  $R(3s)$  — расхождение (размах) в последней серии больше верхнего предела действия «+3s», и  $2R(2s)$  — расхождение (размах) для двух последних серий больше верхнего предела предупреждения «+2s».

Контрольные карты размахов (расхождений) наиболее эффективно выявляют увеличение разброса результатов в пределах одной серии.

**Контрольные карты кумулятивных сумм (CUSUM)** в основном предназначены для быстрого выявления небольших по величине постоянных систематических погреш-

**Табл. 2. Коэффициенты для расчета среднего значения и контрольных пределов для контрольных карт размахов (расхождений)**

Число измерений в серии $N$	Коэффициент для расчета средней линии ( $d_2$ )	Коэффициент для расчета (верхнего) контрольного предела «2s» ( $D_2(2)$ )	Коэффициент для расчета (верхнего) контрольного предела «3s» ( $D_2$ )
2	1,128	3,686	2,834
3	1,693	4,358	3,469
4	2,059	4,698	3,819
5	2,326	4,918	4,054

ностей и дрейфов. По вертикали на них откладывают кумулятивные суммы  $C_1, C_2, C_3, \dots, C_p, \dots$ , которые рассчитывают по формулам:

$$\begin{aligned} C_1 &= (X_1 - \bar{X}); \\ C_2 &= C_1 + (X_2 - \bar{X}); \dots; \\ C_j &= C_{j-1} + (X_j - \bar{X}); \end{aligned} \quad (11)$$

Здесь  $X_k$  — результаты контроля, полученные с момента начала расчета кумулятивных сумм,  $\bar{X}$  — общее среднее, полученное в предшествующем периоде. Обычно контрольные карты  $CUSUM$ , как и другие карты, начинают вести по окончании периода оценки характеристик. В таком случае  $\bar{X}$  — общее среднее, полученное за этот период,  $X_1$  — первый результат по завершении этого периода. Таким образом, каждая точка карты  $CUSUM$  представляет собой сумму отклонения текущего результата измерения (среднего значения в серии) от среднего и всех предыдущих отклонений. Отклонение от общего среднего «накапливается», что и дало название этому виду контрольных карт. На контрольной карте  $CUSUM$  ось  $X$  соответствует нулевому значению.

Если на карту  $CUSUM$  наносятся собственно величины  $C_i$  (как это делалось, когда их изобрели), то при их визуальном рассмотрении основное внимание уделяют наклону линии, соединяющей точки карты. Если возникает даже небольшая,

но постоянная систематическая погрешность, эта линия будет «ползти» либо вверх (при положительной погрешности), либо вниз — при отрицательной. Для визуального изучения оптимален масштаб, при котором при расстоянии между точками по горизонтали в одну единицу такое же расстояние по вертикали равно  $2s$ . Количество отслеживание угла наклона часто вели с помощью специальных угловых шаблонов, что не слишком удобно.

В настоящее время используются более удобные представления карт  $CUSUM$ . Приведем наиболее эффективный, на наш взгляд, способ, близкий к приведенному в ГОСТ 5725-6-2002.

До тех пор, пока очередной результат  $Y_i$  не выходит за пределы ( $\bar{X} \pm 0,5s$ ), кумулятивные суммы не рассчитываются и не наносятся на контрольную карту. В момент выхода за эти пределы начинают рассчитывать кумулятивные суммы. При этом расчет проводят исходя из того, верхнюю или нижнюю границу пересекла кривая, соединяющая точки на карте Шухарта. Если это ( $\bar{X} + 0,5s$ ), то при расчете кумулятивных сумм из каждого следующего значения вычитают ( $\bar{X} + 0,5s$ ); если пересечена граница ( $\bar{X} - 0,5s$ ), то вычитают ( $\bar{X} - 0,5s$ ).

<sup>5</sup> Существуют и контрольные карты стандартных отклонений, используемые для тех же целей, что и контрольные карты размахов. На них вместо размахов для каждой серии наносятся стандартные отклонения результатов измерений, рассчитанные для данной серии. Однако при числе параллельных измерений менее восьми эффективнее использовать расхождение (размах)  $w$ .

<sup>6</sup> Контрольные карты размахов имеют и нижние контрольные пределы, однако они отличны от нуля лишь при большом числе измерений.

Если после этого очередная кумулятивная сумма окажется больше +5,1s в первом случае или ниже -5,1s во втором, то систематическая погрешность считается достоверно выявленной («контрольный признак» 5,1s). Если же вместо этого кривая, соединяющая кумулятивные суммы, пересечет нулевую линию, их расчет прекращают впредь до нового выхода очередного результата за пределы ( $\bar{X} \pm 0,5s$ ).

Карты *CUSUM* можно считать дополнением к картам Шухарта, они лучше выявляют дрейфы и небольшие по величине смещения (сдвиги среднего значения).

**Контрольные карты текущих расхождений** предназначены для контроля стабильности промежуточной прецизионности. Они полностью аналогичны контрольным картам расхождений (размахов) с тем отличием, что вместо расхождений (размахов) величин  $w$  по вертикали откладывают величины текущих расхождений  $w_j = |\bar{X}_{j+1} - \bar{X}_j|$ , т.е. разности между средними результатами измерений в последовательных сериях. Если каждая аналитическая серия содержит лишь один результат, то он и является средним. Способы расчета контрольных пределов, тревожные и контрольные признаки аналогичны приведенным выше для контрольных карт расхождений (размахов) с той разницей, что вместо расхождения (размаха) рассматривается текущее расхождение.

Контрольные карты текущих расхождений наиболее эффективно выявляют увеличение промежуточной прецизионности (разброса результатов от серии к серии).

**Контрольные карты оценок систематической погрешности** представляют собой видоизмененные карты средних значений Шухарта, на которых по вертикальной оси

вместо средних результатов измерений для подгрупп откладывают оценки систематической погрешности  $\hat{\delta} = \bar{X}_j - \mu$ , где  $\mu$  — условно истинное или принятое опорное значение измеряемой величины. В качестве горизонтальной оси используют 0, а контрольные пределы  $\pm 2s$  и  $\pm 3s$  наносят в виде горизонтальных линий. «Тревожный признак» на контрольной карте оценок систематической погрешности — ситуация, когда очередная точка выходит за пределы  $\pm 2s$ , но лежит в пределах  $\pm 3s$ . «Контрольный признак»: очередная точка лежит вне пределов  $\pm 3s$ . По этим контрольным картам отслеживают степень отклонения результатов исследований от истинного (или принятого опорного) значения измеряемой величины.

При построении карт оценок систематической погрешности используют оценку  $s$ , полученную в данной лаборатории в предшествующий период, то есть это та же величина  $s$ , что и на контрольной карте средних значений Шухарта. Откуда берется величина  $\mu$ , не сказано. Можно выделить три ситуации.

- Величина  $\mu$  получена как среднее значение предыдущих результатов измерений в данной лаборатории (т.е.  $\mu = \bar{X}$ ). В этом случае карта оценок систематической погрешности — та же карта Шухарта.

- Величина  $\mu$  — истинное или принятое опорное значение измеряемой величины для данного контрольного материала — получена иным путем (например, при аттестации стандартного образца, если контроль ведется по СО), однако она несущественно отличается от  $\bar{X}$ . Это означает, что систематическая погрешность данной лаборатории  $\Delta$  невелика. В этом случае карта оценок систематической погрешности может рассматриваться как дополнительная к карте Шухарта.

- Величины  $\mu$  и  $\bar{X}$  существенно различаются, то есть систематическая погрешность данной лаборатории  $\Delta$  велика. В этом случае карту оценок систематической погрешности можно использовать лишь для ориентировки.

Контрольные карты оценок систематической погрешности чаще всего рассматриваются как вспомогательные.

**Карты «Экспоненциально взвешенных движущихся средних» (EWMR)** [6] (оригинальное название — геометрические движущиеся средние, GMA) [7]) отличаются от карт *CUSUM* весовыми коэффициентами, характеризующими вклад в оцениваемые результаты измерений, полученные в предшествующих сериях. Этот вклад тем меньше, чем раньше была проанализирована серия.

Эти карты строят следующим образом: по окончании периода оценки характеристик рассчитывают общее среднее  $\bar{X}$ , затем — значения геометрических движущихся средних для последующих серий:

$$\begin{aligned} X_1^w &= wX_1 + (1-w)X_0^w; \\ X_2^w &= wX_2 + (1-w)X_1^w; \dots, \\ X_k^w &= wX_k + (1-w)X_{k-1}^w; \dots \end{aligned} \quad (12)$$

где  $X_0^w = \bar{X}$  ( $\bar{X}$  — общее среднее, полученное в предшествующем периоде),  $X_k$  — результаты контроля, полученные с момента начала расчета геометрических движущихся средних,  $w$  — весовой фактор. Обычно значение  $w$  выбирают равным 0,20–0,25. Полученные значения  $X_1^w; X_2^w; \dots; X_k^w; \dots$  наносят на контрольную карту *EWMR*, в качестве оси  $X$  используют общее среднее  $\bar{X}$ .

Оценка дисперсии величин  $X_k^w$  равна

$$S_{X_k^w} = S[1 - (1-w)^{2k}] \cdot \frac{w}{2-w}, \quad (13)$$

где  $s$  — стандартное отклонение непреобразованных (получаемых экспериментально) средних значений  $X_k$ .

При больших значениях  $k$  получаем  $S_{X_i}^2 \cong \frac{s^2 w}{2-w}$ . Основным контрольным признаком, используемым при работе с геометрическим движущимся средним, является выход очередного результата за пределы  $(\bar{X} \pm 3ws)$ .

Как видно, при использовании  $EWMR$  в значение последнего результата (по которому ведется контроль) вносят вклад и предыдущие результаты, причем вклад тем меньше, чем дальше эти результаты отстоят от последнего. За счет этого снижается чувствительность контроля к резким (возможно, случайным) изменениям в процессе измерений и повышается эффективность выявления систематических погрешностей. Карты  $EWMR$  не описаны в ГОСТ 5725-6-2002.

Карта  $EWMR$  является промежуточной между картами средних значений Шухарта и картами  $CUSUM$ ; к какой из них она ближе, определяется выбором весового коэффициента  $w$ .

## Контроль стабильности с помощью контрольных карт в отсутствие контрольного материала

Как правило, в большинстве случаев можно найти (приобрести, приготовить самостоятельно) подходящий контрольный материал, т.е. однородное и стабильное вещество, более или менее похожее на анализируемые рутинные пробы. Однако иногда это не удается, чаще всего из-за нестабильности исследуемых веществ. В этом случае возможно применение двух способов контроля стабильности — с помощью контрольных карт расхождений (размахов) и контрольных карт текущих расхождений, ведущихся по результатам исследований рутинных проб.

**Контрольные карты размахов (расхождений)**, построенные по результатам исследований рутинных проб, аналогичны построенным с использованием контрольных материалов, с той разницей, что на них откладываются разности между максимальным и минимальным результатами, полученными для какой-либо рутинной пробы.

Обозначим выбранную в  $j$ -й серии для контроля рутинную пробу индексом  $r_j$ , а полученные для нее результаты измерений — как  $X_{j1}^{pyt}, X_{j2}^{pyt}, \dots, X_{jl}^{pyt}$ . Тогда на контрольную карту будут откладываться разности между максимальной и минимальной из этих величин  $X_{j,max}^{pyt} - X_{j,min}^{pyt}$ . Если, как это чаще всего бывает, для рутинной пробы, выбранной для контроля, в  $j$ -й серии проводится два измерения  $X_{j1}^{pyt} - X_{j2}^{pyt}$ , то на карту откладываются величины  $|X_{j1}^{pyt} - X_{j2}^{pyt}|$ .

Основная трудность здесь заключается в том, что пробы, выбранные для контроля, должны быть более или менее похожими как по свойствам, так и по значению измеряемой величины. Конечно, это не всегда легко обеспечить, особенно для твердых проб, гомогенность которых не гарантирована. В отдельных случаях в конкретной серии измерений вообще не удастся выбрать подходящую пробу (например, в экологической лаборатории в какой-то серии вообще может не оказаться проб с достоверно обнаруживаемой концентрацией аналита).

Требование близости значений определяемого показателя для разных рутинных проб, используемых в процессе контроля, в основном связано с тем, что повторяемость (сходимость) зависит от измеряемой величины. Правда,

влияние этого фактора можно уменьшить (но не устранить), нанося на контрольную карту величины размахов (расхождений), отнесенные к приписанным методикой значениям стандартного отклонения повторяемости (сходимости)<sup>7</sup>.

**Контрольные карты текущих расхождений** также могут вестись по результатам исследований рутинных проб следующим образом:

- в первой серии выбирается и исследуется одна из рутинных проб и по результатам для нее рассчитывается среднее значение. Часть этой пробы сохраняют до второй серии, где вновь исследуют, после чего для нее рассчитывают разность между средними значениями в первой и второй сериях и наносят абсолютную величину этой разности на контрольную карту;

- во второй серии выбирают и исследуют таким же образом другую рутинную пробу, рассчитывают для нее среднее значение и часть ее сохраняют до третьей серии. Исследовав эту пробу в третьей серии и рассчитав среднее, вычитают из него предыдущее среднее и абсолютную величину разности наносят на контрольную карту;

- в третьей серии выбирают и исследуют еще одну рутинную пробу, рассчитывают для нее среднее значение и часть ее сохраняют до четвертой серии и так далее.

Все это можно записать следующим образом. Обозначим среднее значение результатов исследования рутинной пробы, выбранной для контроля в  $j$ -й серии, как  $\bar{X}_{(j),j}^{pyt}$ , а результаты исследования той же пробы в следующей серии как  $\bar{X}_{(j),j+1}^{pyt}$ . На контрольную карту текущих расхождений откладываются абсолютные значения

<sup>7</sup> РМГ 76-2014 «Внутренний контроль качества результатов количественного химического анализа» Приказом Росстандарта № 778-ст от 09.07.2014 г. введены в действие в качестве рекомендаций по метрологии с 01.01.2016 г.

разностей этих величин  $|\bar{X}_{(j),j+1}^{\text{рут}} - \bar{X}_{(j),j}^{\text{рут}}|$ . В остальном контрольная карта текущих расхождений по результатам исследования рутинных проб ведется так же, как карта текущих расхождений, построенная для контрольного материала.

Контрольные карты текущих расхождений по рутинным пробам ведут крайне редко. Дело в том, что отсутствие контрольного материала, как правило, связано с нестабильностью объекта исследований. А для таких объектов почти никогда нельзя гарантировать неизменность определяемого показателя в рутинных пробах до следующей серии измерений.

### IT при внутрилабораторном контроле

Контроль стабильности измерений, выполняемый вручную, как и другие способы внутрилабораторного контроля, всегда был неудобным мероприятием. А с появлением описанных выше усложненных его вариантов, требующих достаточно сложных расчетов и ведения множества контрольных карт, и вовсе стал мало реальным. Проблема была решена с появлением компьютерных программ. Их использование позволяет с легкостью проводить необходимые расчеты и исключить практику построения контрольных карт на бумаге. В настоящее время большинство лабораторий, ведущих контроль измерений по контрольным картам, используют те или иные компьютерные программы.

Существует огромное число компьютерных программ, с помощью которых можно вести контроль стабильности. Иногда это программы, являющиеся неотъемлемой частью прибора. Достаточно часто исполь-

зуются и специализированные программы только для ведения контрольных карт (зачастую созданные прямо в лаборатории). Однако в последнее время все шире применяются программы, предназначенные для компьютеризации лаборатории в целом (LIMS), которые также позволяют компьютеризировать контроль стабильности. В большинстве случаев они устроены так, что при проведении контроля сотрудник должен после получения результатов исследования контрольного материала ввести их в специальный блок программы и, проведя контроль, принять решение о качестве исследований. Достаточно неудобно, причем зачастую этим занимается не лаборант, а специалист по контролю качества.

Существует другой, на наш взгляд, оптимальный подход к компьютеризации контроля стабильности в LIMS. Лаборант вводит в программу первичные результаты измерений как рутинных проб, так и контрольного материала. Компьютер рассчитывает результаты измерений (в единицах измерения), после чего для рутинных проб рассчитываются окончательно приводимые результаты, а результаты для контрольного материала (материалов) наносятся на контрольные карты, анализируются на соответствие контрольным пределам, по правилам Вестгарда и т.д. При выявлении контрольных или «тревожных» признаков программой выдаются предупреждения, и после этого в соответствии с правилами, разработанными в лаборатории для таких случаев, принимается решение о дальнейших действиях. Конечно, все это не мешает в любой момент посмотреть, что происходит на контрольных картах.

Такой подход реализован в программе *QControl*, наиболее распространенной в России и рекомендованной к использованию Госстандартом России<sup>8</sup>.

### Несколько общих замечаний

1. В ситуации, когда контроль стабильности уже ведется, а контрольный материал заканчивается (или подходит к концу срок его годности), незадолго до окончания работы с контрольным материалом (назовем его «выводимым»), но не позже чем за 20–30 серий до этого момента (в зависимости от продолжительности периода оценки характеристик), в анализ вводят материал с близкими свойствами и концентрацией определяемых веществ («вводимый»). На протяжении 20–30 серий измерений выводимый и вводимый контрольные материалы анализируют совместно, а контроль стабильности ведут по выводимому. Для вводимого материала это период оценки характеристик. Процедура, называемая «**перекрыванием**», обеспечивают непрерывность контроля стабильности.

2. Контроль стабильности легко организуется, если исследования выполняются достаточно часто (не реже, чем 1–2 раза в неделю). В противном случае возможны следующие варианты:

- передать этот вид анализа (исследований) в другую (аккредитованную, авторитетную) лабораторию, в которой такой анализ выполняется часто и налажен контроль качества, тогда не надо будет каждый раз мучиться с обеспечением и доказательством качества измерений. К сожалению, это не всегда возможно либо из-за отсутствия подходящей лаборатории, либо из-за противодействия со стороны администрации;

<sup>8</sup> Р 50.2.003-2000 Рекомендации по метрологии «Внутренний контроль качества измерений. Пакет программ *QControl*» введены в действие Постановлением Госстандарта России № 182-ст от 06.07. 2000 г.

- выполнять данный вид исследований чаще, чем это на самом деле необходимо (например, один раз в неделю), контролируя при этом стабильность по контрольным картам. В этом случае лаборатория всегда будет готова выполнить редкое измерение с требуемым качеством. Такой вариант приемлем только в случае простых и дешевых измерений (например, радиоактивности).

Конечно, можно применять и другие способы контроля качества.

3. Контроль стабильности, как и другие способы внутрилабораторного контроля, не является идеальным.

Во-первых, в описанном выше виде он не охватывает процедуру отбора проб, которая при анализе неоднородных объектов вносит существенный, а часто и определяющий вклад в общую погрешность. Стадия отбора пробы нуждается в особых способах планирования и контроля.

Во-вторых, удовлетворительные результаты внутрилабораторного контроля не являются гарантией такого же качества исследования рутинных проб. В частности, различия в качестве анализа контрольных материалов и рутинных проб могут возникать как из-за отличий в составе их матрицы (неадекватность), так и в силу случайных причин.

В-третьих, существует субъективный фактор: выполняющий анализ персонал часто по разным причинам заинтересован в «улучшении» (подгонке) результатов, что, безусловно, снижает эффективность контроля. Иногда это проявляется в том, что результаты группируются вблизи контрольных пределов ( $\bar{X} \pm 2s$ ) на карте средних значений Шухарта.

## Использованная литература:

1. Дворкин В.И. Контроль стабильности измерений. Основные принципы // Контроль качества продукции. 2021, № 3. — С. 44–52.
2. Дворкин В.И. Метрология и обеспечение качества химического анализа. — М.: Техносфера, 2019. — 317 с.
3. Shewhart W.A. Statistical method for the view point of quality control / Ed. Deming W.E. — Pennsylvania: Lancaster Press, 1939. — 155 p.
4. Levey S., Jennings E.R. The use of control charts in the clinical laboratory // Am. J. Clin. Pathol. 1950. V. 20. — P. 1059–1066.
5. Westgard J.O., Barry P.L., Hant M.R., Groth T.A. A multi-rule Shewhart chart for quality control in clinical chemistry // Clin. Chem. 1981. V. 27. — P. 493–501.
6. Hunter J.S. The exponentially weighted moving average // J. of Qual. Technol. 1986. V. 18. P. 203–210.
7. Roberts S.W. Control Chart Tests Based on Geometric Moving Averages // Technometrics. 1959. V. 1, No. 3. — P. 239–242.

## References:

1. Dvorkin V.I. Measurement stability control. Basic principles // Production quality control. 2021, No. 3. — P. 44–52.
2. Dvorkin V.I. Metrology and quality assurance of chemical analysis. — Moscow: Technosphere Publ., 2019— 317 p.
3. Shewhart W.A. Statistical method for the view point of quality control / Ed. Deming W.E. — Pennsylvania: Lancaster Press, 1939. — 155 p.
4. Levey S., Jennings E.R. The use of control charts in the clinical laboratory // Am. J. Clin. Pathol. 1950. V. 20. — P. 1059–1066.
5. Westgard J.O., Barry P.L., Hant M.R., Groth T.A. A multi-rule Shewhart chart for quality control in clinical chemistry // Clin. Chem. 1981. V. 27. — P. 493–501.
6. Hunter J.S. The exponentially weighted moving average // J. of Qual. Technol. 1986. V. 18. P. 203–210.
7. Roberts S.W. Control Chart Tests Based on Geometric Moving Averages // Technometrics. 1959. V. 1, No. 3. — P. 239–242.



## Резюме

**Несмотря на то, что контроль стабильности нельзя признать идеальным, при правильной организации он является важнейшим способом обеспечения качества не только химического анализа, но и других видов лабораторных исследований.**

**TITLE:** \_\_\_\_\_

### Monitoring the stability of measurements

**AUTHOR:** \_\_\_\_\_

**V.I. Dvorkin**, Chief Researcher at the Institute of Petrochemical Synthesis of the Russian Academy of Sciences, Professor, Doctor of Chemical Sciences

**ABSTRACT** \_\_\_\_\_

The most important method of in-laboratory quality control of measurements is the assessment of stability. Earlier, we considered the model of the measurement process, which is the basis for stability control, its basic principles and organizational aspects [1]. This article discusses how such control is carried out. At the same time, all of the above is true for any measurements, although the fundamental standard GOST R ISO 5725-6-2002 all examples are taken from chemical analysis.

**KEYWORDS:** \_\_\_\_\_

measurements, stability, evaluation, control material, control card, error

**SUMMARY** \_\_\_\_\_

Despite the fact that stability control cannot be considered ideal, if properly organized, it is the most important way to ensure the quality of chemical analysis and other types of laboratory tests.