

УДК 51-74:665.71

Д.О. СКОБЕЛЕВ, В.П. КОВАЛЕНКО, Е.И. ВЫБОЙЧЕНКО

(ФГУП «Всероссийский научно-исследовательский институт стандартизации материалов и технологий», г. Москва)

Алгоритм формирования ограничительных норм показателей качества нефтепродуктов с использованием метода группового учёта аргументов (далее – МГУА)

Ключевые слова: ограничительные нормы, качество нефтепродуктов, метод группового учёта аргументов.

В статье излагается описание разработанного алгоритма формирования ограничительных норм показателей качества нефтепродуктов с использованием метода группового учёта аргументов. Проанализированы характерные особенности выбора оптимальной модели, описывающей исходные данные. Показано как на основе нескольких моделей в случае их линейного вида можно получить регламентированные значения, ограничивающие интервал, определяющий качество продукции.

При постановке продукции на производство одним из основных этапов процедуры является подтверждение соответствия показателей качества получаемого продукта установленным нормам. Показателем качества продукции называется количественная характеристика одного или нескольких свойств продукции, рассматриваемая применительно к определённым условиям её создания, эксплуатации или потребления [1]. Проведение процедуры контроля и подтверждения соответствия качества невозможно без информации о регламентированном значении показателя качества (ограничительной норме), установленном в нормативной документации. Регламентированные значения, например нефтепродукта, ограничивают некий интервал, обеспечивающий колебание значения показателя качества, связанное с использованием различного сырья в разных регионах, различного технологического оборудования и т.д. Поскольку такие интервалы показателей зачастую выбираются на основании опыта технологов из совокупности статистического материала, представляется актуальной разработка алгоритма научно обоснованного формирования ограничительных норм показателей качества.

Настоящая работа посвящена анализу одного из возможных путей решения этой задачи.

В качестве исходных данных для расчёта необходимо использовать измеренные значения показателя

качества за определённый интервал времени. Эти данные можно представить в виде точечной диаграммы. Тогда все модели, описывающие изменение показателя качества за выбранный временной интервал, условно можно разделить на простые и сложные. Примеры таких моделей показаны на рис. 1 и 2 соответственно для простой и сложной модели. Простая модель в данном случае представляет прямую линию, полученную посредством аппроксимации

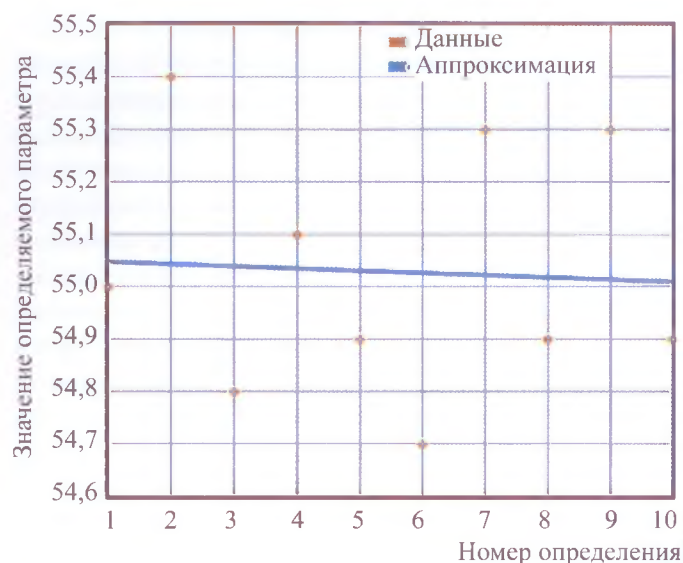


Рис. 1. Пример простой модели

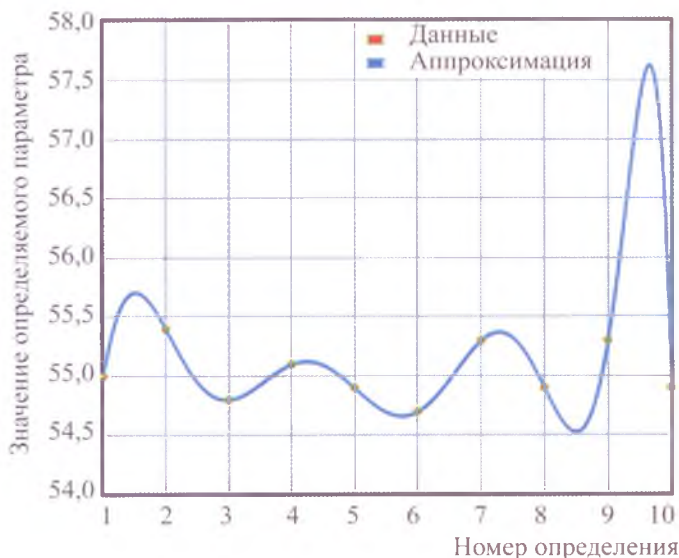


Рис. 2. Пример сложной модели

исходных данных. Под шумом понимают любое внешнее воздействие, способное изменить результаты определения показателя качества. Тогда очевидно, что если изменить хотя бы один результат измерения (см. рис. 1), то простая модель не претерпевает значительных изменений. Это означает, что она не реагирует на шум, но в свою очередь плохо отражает исходные данные. Сложная же модель в рассматриваемом примере – полином 10-го порядка и в отличие от простой модели любое изменение результатов определения показателя качества меняет её кардинально, т.е. сложная модель хорошо отражает исследуемые данные, но предельно чувствительна к шуму.

Использование МГУА позволяет выбрать для описания имеющегося набора данных модель оптимальной сложности из заданного класса. Основным результатом применения теории МГУА состоит в том, что при неточных зашумлённых данных и коротких выборках минимум критерия указывает так называемую нефизическую модель, точность которой выше и структура которой проще полной физической модели [2]. По сути МГУА представляет собой индуктивный подход к процедуре моделирования. Он рассматривает систему, переходя от частного к общему, и синтезирует её путём слияния компонент, разрабатываемых отдельно [3]. Важно понимать, что индуктивный подход имеет дело с фиксированным классом моделей. Класс модели должен отражать её возрастающую сложность. Такой подход не может определить самую оптимальную модель среди всех возможных, он ищет оптимальную модель только в заданном классе. Исходя из изложенного выше, можно выделить три основных этапа моделирования с использованием МГУА:

1. Определение серии моделей возрастающей сложности S . Расчёт коэффициентов для каждой из них.

2. Вычисление значения внешнего критерия селекции для каждой из полученных моделей.

3. Выбор оптимальной модели по минимальному значению критерия.

Далее подробно рассмотрен и проанализирован каждый из этапов применительно к выбранной тематике.

На первом этапе в качестве класса объектов рекомендуем выбрать полиномы одной переменной. Предложено рассмотреть полиномы степени не выше 4, поскольку в этом случае максимальное число членов полинома составляет 5, а соответственно общее число частных моделей 31, что не сильно усложнит расчёт. Помимо этого для полиномов со степенью выше 4 не существует решения в радикалах, т.е. корни полинома нельзя представить в виде формулы, через его коэффициенты с помощью конечного числа арифметических операций. Это означает, что для уравнения 5 степени и выше не всегда можно получить точное решение, что может существенно повлиять на моделирование. Тогда в общем виде серию моделей можно представить как функцию

$$y_m = k_0 + k_1x + k_2x^2 + k_3x^3 + k_4x^4.$$

В зависимости от уровня сложности получим:

для $S = 1$ модели $q_{11} = k_{11}$, $q_{12} = k_{12}x$, $q_{13} = k_{13}x^3$ и т.д.;

для $S = 2$ модели $q_{21} = k_{21} + k'_{21}x$, $q_{22} = k_{22} + k''_{22}x^2$, $q_{23} = k_{23}x + k'_{23}x^3$;

для $S = 3$ модели $q_{31} = k_{31} + k'_{31}x + k''_{31}x^2$,

$q_{32} = k_{32} + k'_{32}x + k''_{32}x^3$, $q_{33} = k_{33} + k'_{33}x + k''_{33}x^4$ и т.д.;

для $S = 4$ модели $q_{41} = k_{41} + k'_{41}x + k''_{41}x^2 + k'''_{41}x^3$,

$q_{42} = k_{42} + k'_{42}x + k''_{42}x^2 + k'''_{42}x^4$ и т.д.;

для $S = 5$ модель

$$q_{51} = k_{51} + k'_{51}x + k''_{51}x^2 + k'''_{51}x^3 + k^{IV}_{51}x^4$$

А общее число всевозможных частных моделей $P_n = 2^n - 1 = 2^5 - 1 = 31$.

Использование МГУА предполагает выбор двух критериев: внутреннего, по которому строят модель для выбранного уровня сложности; внешнего, по которому и выбирают оптимальную модель. Общая классификация критериев, используемых в МГУА, представлена на рис. 3.

Основным способом разделения информации на «внутреннюю» и «внешнюю» в алгоритмах МГУА является разбиение всей имеющейся выборки исходных данных на две и более частей (последователь-

ностей). Для упрощения назовём эти части А и В. Такое разделение позволяет построить разнообразные критерии для количественного сравнения моделей различной структуры [2].

В данном случае в качестве внутреннего критерия необходимо использовать метод наименьших квадратов (МНК), поскольку параметры получаемой функциональной зависимости входят в неё линейно. Использование в расчётах МНК позволяет найти такие значения коэффициентов модели, при которых сумма квадратов отклонений теоретического распределения от эмпирического была бы наименьшей. Иными словами, из всей совокупности линий, которые можно провести, исходя из представленных данных, необходимая линия или кривая выбирается так, чтобы сумма квадратов расстояний по вертикали между точками и этой линией была минимальной [4].

Все внешние критерии можно разделить на две большие группы: критерии точности, выражающие ошибку проверяемой модели на различных частях выборки и критерии согласованности, являющиеся мерой близости оценок, получаемых на различных частях выборки [2]. Из предложенных в литературе был выбран один из критериев точности – критерий регулярности, который является исторически первым критерием селекции [5].

$$K_{reg} = \sum_B (y - y_{mA})^2, \quad (1)$$

где \sum_B – суммирование по точкам части «В» выборки; y – данные наблюдений; y_{mA} – значение модели, полученной на «А» части выборки.

Оптимальная модель, полученная на основании вышеописанного критерия, может быть чувствительна к наличию в исходной совокупности значений, резко отличающихся от остальных (к наличию выбросов). Для сглаживания влияния помех на обеих частях выборки рекомендуется использовать при анализе эти части равноправно. Этого можно достичь, рассчитав в начале критерий регулярности по формуле (1), а затем, поменяв «А» и «В» местами, получить ещё один критерий регулярности. Тогда



Рис. 3. Общая классификация критериев, используемых в МГУА

обобщенный и уже симметричный критерий будет иметь вид

$$K_{reg} = K_{reg_I} + K_{reg_{II}}, \quad (2)$$

где $K_{reg_I} = \sum_B (y - y_{mA})^2$; $K_{reg_{II}} = \sum_A (y - y_{mB})^2$; \sum_A , \sum_B – суммирование по точкам части «А» и «В»

выборки соответственно; y – данные наблюдений; y_{mA} , y_{mB} – значения модели, полученной на «А» и «В» частях выборки соответственно.

Использовать симметричные формы критериев всегда предпочтительнее, учитывая конечность выборки данных: во-первых, снижается чувствительность критериев к способу разделения данных на части, во-вторых, усредняется влияние помех (происходит своеобразная фильтрация) [2].

Стоит отметить, что одним из основных требований, предъявляемых к разрабатываемой модели, является независимость её описания от данных. Для его реализации необходимо иметь критерий, который позволил бы сравнить две модели. Этот критерий называется критерием минимума смещения коэффи-

циентов. Он относится к критериям согласованности и отражает требование того, чтобы лучшая модель имела коэффициенты, оценки которых, полученные на «А» и «В», минимально отличались. Этот критерий позволяет получить несмещённую оценку, т.е. оценку, у которой нет систематической ошибки. Исходя из определения критерия минимума смещения коэффициентов может быть рассчитан по формуле

$$K_{unbias} = \sum (Y_m(A) - Y_m(B))^2, \quad (3)$$

где $Y_m(A)$ и $Y_m(B)$ – модели, полученные на «А» и «В» частях выборки соответственно; \sum – суммирование по всем точкам выборки.

Для получения единственной оптимальной модели необходимо определить совместное использование вышеописанных критериев. Существует два возможных варианта: расчёт комбинированного критерия и последовательный отбор. В первом случае назначают веса λ_1 и λ_2 критериев, причём $\lambda_1 + \lambda_2 = 1$. Тогда комбинированный критерий

$$K = \lambda_1 K_{рек} + \lambda_2 K_{unbias}. \quad (4)$$

В случае же последовательного отбора необходимо отобрать ряд лучших моделей по первому критерию, а затем из полученного набора с использованием второго критерия определить оптимальную модель.

Для решения поставленной задачи необходимо использовать последовательный отбор, поскольку нет информации о степени значимости того или иного критерия при расчёте ограничительных норм.

Теоретически доказано, что при зашумлённых данных и короткой выборке минимум математического ожидания внешнего критерия единственен [2]. Единственность минимума сохраняется и при достаточно больших группах моделей, что и используется для выбора единственной оптимальной модели.

Предложенная методика позволяет получить оптимальную модель в общем виде. Далее коэффициенты этой модели находят с использованием метода наименьших квадратов.

Рассмотрим один из случаев, когда модель представляет собой линейную функцию вида $y = k + k_1x$. В данном случае в виду линейности функции можно произвести оценку точности МНК.

Пусть опытные значения являются точными, а опытные значения x_i имеют случайные ошибки с одинаковой дисперсией σ^2 для всех i . Тогда

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n [y_i - (k_1 x_i + k)]^2, \quad (5)$$

$$S_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2, \quad (6)$$

$$\sigma_{k_1}^2 = \frac{\sigma^2}{n S_x^2}, \quad (7)$$

$$\sigma_k^2 = \frac{\sigma^2}{n} \left(1 + \frac{\bar{x}^2}{S_x^2}\right), \quad (8)$$

$$\sigma_y^2 = x^2 \sigma_{k_1}^2 + \sigma_k^2 + 2x K_{k_1, k}, \quad (9)$$

где n – число измерений; S_x – стандартное отклонение величины x ; $K_{k_1, k}$ – корреляционный момент,

$$K_{k_1, k} = -\frac{\sigma^2 \bar{x}}{n S_x^2}.$$

Оценку абсолютной погрешности расчёта коэффициентов модели можно получить с использованием коэффициента Стьюдента. Эта оценка будет зависеть от доверительной вероятности и числа проведенных экспериментов.

$$\Delta k = t_\alpha \sigma_k, \quad (10)$$

$$\Delta k = t_\alpha \sigma_{k_1}, \quad (11)$$

где t_α – коэффициент Стьюдента [6].

Для возможности применения полученной методики при расчёте ограничительных норм показателей качества нефтепродуктов необходимо по данным нескольких лабораторий построить модели в одной системе координат, как показано на **рис. 4**.

Далее из всех представленных лабораторий определяют наибольшее число измерений и для этого значения проводят вертикальную линию на графике (**рис. 5**). Минимальным значением, от которого необ-

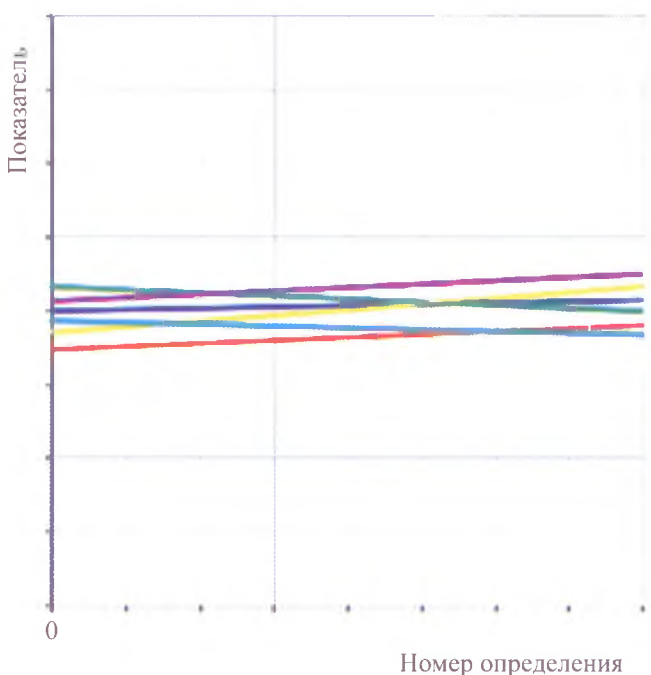


Рис. 4. Пример построения моделей нескольких лабораторий в одной системе координат при расчёте ограничительных норм

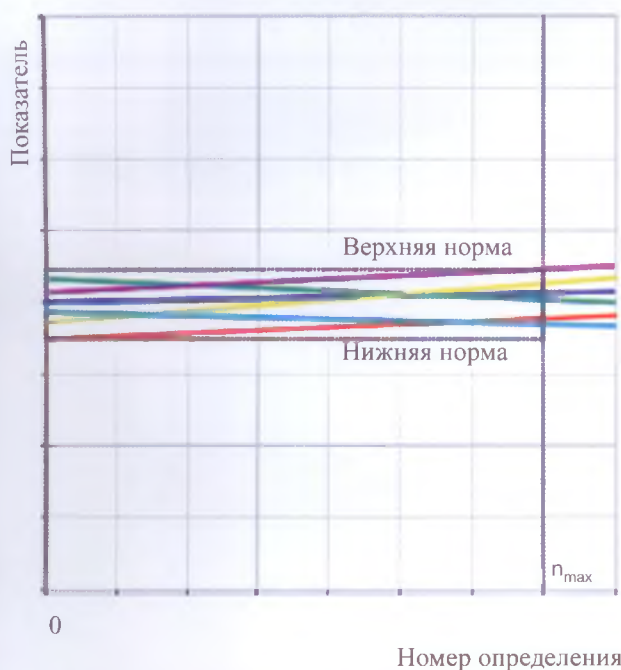


Рис. 5. Пример графического определения ограничительных норм показателей качества

ходимо строить модели, является 0. Исходя из этого, можно графически получить верхнюю и нижнюю ограничительные нормы.

В работе рассмотрен случай получения ограничительных норм показателей качества нефтепродуктов, когда оптимальная модель представляет собой линейную функцию. Если модель нелинейная, то методика расчёта ограничительных норм будет другой.

Рассмотренный алгоритм можно использовать для формирования ограничительных норм для каждого показателя качества нефтепродуктов с учётом данных различных лабораторий. Методика применима как для получения регламентированных значений показателей качества при постановке продукции на производство, так и для внутрилабораторного контроля, прогнозирования изменения показателя качества на короткий промежуток времени, при приёмочном контроле.

Список литературы

- ГОСТ 15467–79 Управление качеством продукции. Основные понятия. Термины и определения
- Ивахненко А.Г., Степашко В.С. Помехоустойчивость моделирования. Киев: Наукова Думка, 1985.
- Звягин Л.С. Принципы системного подхода в моделировании систем // Молодой учёный, 2014. – № 6.
- Елисеева И.И. Эконометрика. М.: Финансы и кредит, 2002.
- Ивахненко А.Г. Метод группового учёта аргументов – конкурент метода стохастической аппроксимации // Автоматика, 1968.
- Шор Я.Б. Статистические методы анализа и контроля качества и надёжности. М.: Советское радио, 1962.

Skobelev D.O., Kovalenko V.P., Vyboychenko E.I.
(Federal State Unitary Enterprise «Russian Research Institute on Standardization of Materials and Technologies», Moscow)

THE ALGORITHM OF FORMATION RESTRICTIVE STANDARD OF PETROLEUM PRODUCTS QUALITY USING GROUP METHOD OF DATA HANDLING

Keywords: restrictive standard, quality of petroleum products, group method of handling.

Abstract.

This article is devoted to the problem of determining the restrictive standards of petroleum product quality. Quality indices of any product are continuously changing during the process of storage, transportation and use. Therefore, quality control is one of the most important component of product life cycle. The procedure of quality control is impossible without information about regulated value of quality index (restrictive standard), which represents the value established in normative documents. This means that today the problem of determining the restrictive standards of product quality is relevant. The central part of the article is devoted to procedure of analysis possible application of group method of data handling in order to use in calculating the regulated values of product quality. Reviewed and analyzed the three main stages of the simulation in compliance with group method of data handling. Among them: determining a number of models increasing complexity, S ; the calculation of the external selection criterion; the choice of the optimal model for minimum criterion. Using group method of data handling assume the choice two criteria: internal, which used to construct the model and external, which used to select the optimal model. This article discusses general classification criteria used in group method of data handling, as well as features of their application in the calculation of restrictive standard. In conclusion, gives an example that illustrates possible outcome of the modeling process. The article considers the case when the optimal model is a linear function. If the model is nonlinear, then the calculation of the restrictive standards will be different from this.

References

- GOST 15367 – 79 Product quality control. Basic concepts. Terms and definitions.
- Ivakhnenko A.G., Stepashko V.S. *Pomekhoustoychivost' modelirovaniya* [Modeling noise stability]. Kiev, Naukova dumka Publ., 1985.
- Zvyagin L.S. The principles of the system approach in modeling systems [Printsipy sistemnogo podkhoda v modelirovanii sistem]. *Molodoy uchenyy – Young scientist*, 2014, no.6.
- Eliseeva I.I. *Ekonometrika* [Econometrics]. Moscow, Finance & credit Publ., 2002.
- Ivakhnenko A. G. A group method of data handling – competitor stochastic approximation method [Metod gruppovogo ucheta argumentov – konkurent metoda stokhasticheskoy aproksimatsii]. *Automatica*, 1968, no.3.
- Shor Ya.B. *Statisticheskie metody analiza i kontrolya kachestva i nadezhnosti* [Statistical methods of analysis and quality control and reliability]. Moscow, Sovetskoe radio Publ., 1962.