

DOI: 10.15593/2499-9873/2021.2.04

УДК 519.6

И.С. Можаровский^{1,2}, С.А. Самотылова^{2,3}, Е.С. Баулин⁴

¹Владивостокский государственный университет экономики
и сервиса, Владивосток, Россия

²Институт автоматизации и процессов управления ДВО РАН,
Владивосток, Россия

³Дальневосточный федеральный университет, Владивосток, Россия

⁴ООО «Центр цифровых технологий», Москва, Россия

ИСПОЛЬЗОВАНИЕ АЛГОРИТМА УСЛОВНЫХ ЧЕРЕДУЮЩИХСЯ МАТЕМАТИЧЕСКИХ ОЖИДАНИЙ ДЛЯ ПОСТРОЕНИЯ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ С ЦЕЛЬЮ ОЦЕНКИ ПОКАЗАТЕЛЕЙ КАЧЕСТВА ВЫХОДНЫХ ПРОДУКТОВ

С целью повышения эффективности производства и улучшения качества товарного продукта рассматривается задача построения математических моделей для оценки показателей качества выходных продуктов в условиях нелинейности технологического процесса. Использование существующих параметрических подходов при построении таких моделей ограничено необходимостью получения аналитически заданной функциональной зависимости (известной или заданной структуры) с последовательным уточнением значений ее коэффициентов. В свою очередь, это приводит к получению математической модели для оценки показателей качества выходных продуктов низкого качества. В случае нелинейности технологического процесса наиболее перспективными являются непараметрические методы, такие как алгоритм чередующихся условных математических ожиданий. Предлагается использовать алгоритм условных чередующихся математических ожиданий для определения структуры связей, вид которых первоначально неизвестен, и выявить нелинейные функциональные зависимости на основе преобразований используемых переменных. Показана эффективность предлагаемого подхода при построении математической модели на синтетическом примере. Предлагаемый подход протестирован на массообменном технологическом объекте процесса производства высокооктановой добавки бензинов – метил-трет-бутилового эфира. Удалось повысить точность математической модели для оценки концентрации метанола в выходном продукте до 67,2 % MAE, в сравнении с моделью, полученной методом нейронной сети.

Ключевые слова: математическая модель, оценка показателей качества, нелинейность технологического процесса, непараметрические методы, нейронная сеть прямого распространения, обобщенно-регрессионная нейронная сеть, алгоритм условных чередующихся математических ожиданий, массообменный технологический объект, ректификационная колонна, метил-трет-бутиловый эфир.

I.S. Mozharovskii^{1,2}, S.A. Samotylova^{2,3}, E.S. Baulin⁴

¹Vladivostok State University of Economics and service,
Vladivostok, Russian Federation

²Institute of Automation and Control Processes Far Eastern Branch
of the Russian Academy of Sciences, Vladivostok, Russian Federation

³Far Eastern Federal University, Vladivostok, Russian Federation

⁴LLL "Center of Digital Technology", Moscow, Russian Federation

USING THE ALGORITHM OF CONDITIONAL ALTERNATING MATHEMATICAL EXPECTATIONS TO MATHEMATICAL MODELS DESIGN TO ESTIMATE THE QUALITY INDICATORS OF OUTPUT PRODUCTS

In order to increase production efficiency and improve the quality of an output product, the problem of mathematical models designs for estimating the quality indicators of output products in conditions of nonlinearity of the technological process is considered. The use of existing parametric approaches in the design of such models is limited by the need to obtain an analytically specified functional dependence (known the model structure) with a sequential refinement of the values of its coefficients. In turn, this leads to obtaining a mathematical model for estimating the quality indicators of output products of low-accuracy. In the case of nonlinearity of the technological process, the most promising are nonparametric methods such as the algorithm of alternating conditional mathematical expectations. It is proposed to use an algorithm of conditional alternating mathematical expectations to determine the structure of relationships between variables, the form of which is initially unknown and to identify nonlinear functional dependencies based on transformations of the variables used. The effectiveness of the proposed approach in designing a mathematical model using a synthetic example is shown. The proposed approach was tested on a mass-transfer technological object of the production process of a high-octane gasoline additive – methyl-tert-butyl ether. It was possible to improve the accuracy of the mathematical model for estimating the concentration of methanol in the output product up to 67.2 % MAE in comparison with the model obtained by the neural network.

Keywords: mathematical model, estimation of quality indicators, nonlinearity of the technological process, nonparametric methods, feed forward neural network, generalized regression neural network, algorithm of conditional alternating mathematical expectations, mass-transfer technological object, distillation column, methyl-tert-butyl ether.

Введение

В связи с повышающимися требованиями к качеству основных видов нефтепродуктов предприятия нефтеперерабатывающей и нефтехимической промышленности вынуждены непрерывно повышать экономическую эффективность производства и качество выпускаемой продукции. Эффективность производства может быть улучшена с помощью систем виртуального мониторинга и контроля показателей качества выходного продукта. Для этого необходима разработка точных математических моделей для оценки показателей качества выходных

продуктов (ММОПК), описывающих нелинейные процессы, проходящие в ректификационных колоннах (РК).

Для построения ММОПК используются результаты лабораторных исследований и данные со встроенных измерительных приборов ректификационных колонн (РК). Если включить адекватную математическую модель для оценки качества выходного продукта РК в систему усовершенствованного управления технологическим процессом (СУУТП), то возможен мониторинг и контроль технологического процесса (ТП) в режиме реального времени.

Методы, используемые для построения ММОПК выходных продуктов ректификационных колонн, можно разделить на два типа: параметрические и непараметрические. К параметрическим относятся регрессионные методы, такие как метод наименьших квадратов [1, 2], робастная регрессия [3, 4], гребневая регрессия [5–9], проекции на скрытые (латентные структуры) [10–12] и т.д. Такой подход построения ММОПК основан на предположении о возможности получения аналитически заданной функциональной зависимости (известной или заданной структуры) с последовательным уточнением значений ее коэффициентов. В реальных условиях большинство массообменных технологических объектов являются слабо формализуемыми из-за недостаточности имеющихся знаний о них и среде, в которой они функционируют, поэтому такой подход обычно не вносит ясности в выбор структуры модели. В случае нелинейности технологического процесса наиболее перспективными являются непараметрические методы [13], в частности нечеткая логика, искусственные нейронные сети (НС) [14] и алгоритм чередующихся условных математических ожиданий.

Математические модели, построенные с помощью нечетких алгоритмов, часто успешно подходят для построения моделей сложных объектов, в которых входные данные могут быть неточными и слабо формализованными [15–17]. Основными недостатками такого подхода являются сложность и субъективность формирования правил нечеткой логики, на основе которых будет функционировать модель, а также выбор вида и параметров ее функций принадлежности. Для этого требуются эмпирические знания об объекте исследования и широкая обучающая выборка данных.

Подход к построению ММОПК на основе нейронных сетей хорошо зарекомендовал себя для объектов, имеющих выраженные связи

выходной переменной с входными. В случае, когда такие связи неявные (слабые), требуются дополнительные исследования, направленные на определение структуры и параметров прогнозирующего алгоритма нейронных сетей, от которых зависит успешность разрешения поставленной задачи [18].

Наиболее универсальным является вероятностный метод оценки идентифицируемости зависимостей из данных источника [19], однако он имеет существенный недостаток, заключающийся в невозможности извлечь характер (вид) нелинейных зависимостей. В связи с этим представляется перспективным использование непараметрических методов [13, 14] в интеллектуальном анализе данных, в частности алгоритма ACE (Alternating Conditional Expectation), так как появляется возможность понимания структуры нелинейностей. Алгоритм условных чередующихся математических ожиданий позволяет находить сложные зависимости входных переменных к выходным, определить структуру связей, вид которых первоначально неизвестен, и выявить нелинейные функциональные зависимости на основе преобразований используемых переменных [20–26].

В работах [27–30] приводятся результаты успешного применения данного метода для построения математических моделей для слабо формализованных объектов, но они ограничиваются выявлением структуры зависимостей выходных данных ко входным и построением модели с использованием этой информации. При этом применяются аппроксимирующие функции, параметры которых подбираются [31]. Других способов применения алгоритма ACE в предсказательном моделировании не встречалось.

В работе предлагается использовать алгоритм ACE для построения математических моделей на основе экспериментальных данных, для оценки показателей качества выходных продуктов ректификационных колонн в условиях структурной неопределенности модели технологического объекта.

1. Алгоритм построения математических моделей для оценки показателей качества выходных продуктов

Рассматривается объект, функционирующий в условиях неопределенности с p входными, контролируруемыми переменными x_1, \dots, x_p и выходом Y , связанными некоторой функциональной зависимостью:

$$Y = F(\mathbf{X}, \mathbf{B}) + \varepsilon, \quad (1)$$

где \mathbf{X} – вектор входных контролируемых переменных, $\mathbf{X} = (x_1, \dots, x_p)$; \mathbf{B} – вектор коэффициентов, $\mathbf{B} = (\beta_1, \dots, \beta_p)$; ε – действующее на выходе неконтролируемое возмущение.

В случае, когда вид истинной зависимости (структуры) F в формуле (1) неизвестен, приходится выбирать структуру, опираясь на имеющуюся априорную информацию об исследуемом объекте, и обращаться к регрессионным моделям вида

$$\hat{Y}_j = \beta_0 + \sum_{i=1}^p \beta_i f_{ij}, \quad j = 1, \dots, N, \quad (2)$$

где \hat{Y}_j – предсказанное значение, соответствующее j -му значению Y_j выходной переменной Y ; $f_{ij} = f_{ij}(x_1, \dots, x_p)$, $i = 1, \dots, p$, $j = 1, \dots, N$ – некоторые функции входных контролируемых переменных x_1, \dots, x_p ; $\beta_0, \beta_1, \dots, \beta_p$ – оценки коэффициентов из формулы (1).

При этом их возможное количество в общем случае может быть бесконечным, но заранее неизвестно, удастся ли выбрать среди них модель, адекватно описывающую зависимость (1) в условиях заданных ограничений на функционирование объекта.

При использовании алгоритма АСЕ переходим от модели вида (1) к выражению

$$\theta(Y) = \sum_{i=1}^p \Phi_i(x_i) + \varepsilon, \quad (3)$$

где θ – функция выходной переменной Y ; Φ_i – функции входов x_i , $i = 1, \dots, p$; ε – ошибка.

Исходя из прямой аналогии с базовым АСЕ алгоритмом, минимизируем выражение

$$e^2(\theta, \varphi_1, \dots, \varphi_p) = E \left[\theta(Y) - \sum_{j=1}^p \varphi_j(x_j) \right]^2, \quad (4)$$

сохраняя $E\theta^2 = 1$, $E\theta = E\varphi_1 = \dots = E\varphi_p = 0$ на протяжении серий минимизаций единичной функции, включающих двумерные условные ма-

тематические ожидания. Для заданного набора функций $\varphi_1(x_1), \dots, \varphi_p(x_p)$ минимизация выражения (4) в отношении к $\theta(Y)$

$$\theta'(Y) = E \left[\sum_{i=1}^p \varphi_i(x_i) | Y \right] / \left\| E \left[\sum_{i=1}^p \varphi_i(x_i) | Y \right] \right\|. \quad (5)$$

Следующий шаг – это минимизация выражения (4) в отношении к $\varphi_1(x_1), \dots, \varphi_p(x_p)$ при заданном $\theta(Y)$. Для этого минимизируем выражение (4) в отношении единичной функции $\varphi_k(x_k)$ при данном $\theta(Y)$ и определенном наборе $\varphi_1, \dots, \varphi_{k-1}, \varphi_{k+1}, \dots, \varphi_p$. Решением является

$$\varphi'_k(x_k) = E \left[\theta(Y) - \sum_{i \neq k} \varphi_i(x_i) | x_k \right]. \quad (6)$$

Алгоритм для построения математических моделей с целью оценки показателей качества выходных продуктов можно представить следующим образом:

input: присвоить $\theta(Y) = Y / \|Y\|$ и $\varphi_1(x_1), \dots, \varphi_p(x_p) = 0$;

while $e^2(\theta, \varphi_1, \dots, \varphi_p)$ не перестанет убывать;

while $e^2(\theta, \varphi_1, \dots, \varphi_p)$ не перестанет убывать;

for $k = 1 : p$

$$\varphi'_k(x_k) = E \left[\theta(Y) - \sum_{i \neq k} \varphi_i(x_i) | x_k \right];$$

$$\varphi_k(x_k) = \varphi'_k(x_k)$$

end

end

$$\theta'(Y) = E \left[\sum_{i=1}^p \varphi_i(x_i) | Y \right] / \left\| E \left[\sum_{i=1}^p \varphi_i(x_i) | Y \right] \right\|;$$

$$\theta(x) = \theta'(x)$$

end

output: $\theta, \varphi_1, \dots, \varphi_p$ будут решениями для $\theta^*, \varphi_1^*, \dots, \varphi_p^*$.

Цель алгоритма заключается в получении оптимальных преобразований $\theta(Y), \varphi_1(x_1), \dots, \varphi_p(x_p)$. Среднеквадратичная ошибка e^2 :

$$e^2(\theta, \varphi_1, \dots, \varphi_p) = \frac{1}{N} \sum_{k=1}^N \left[\theta(Y_k) - \sum_{j=1}^p \varphi_j(x_{kj}) \right]^2.$$

2. Вычислительный эксперимент и анализ результатов

Предложенная методика иллюстрируется синтетическим примером со сложной нелинейной структурой и применена для промышленного технологического объекта.

При построении ММОПК важно решить такие задачи, как выбор регрессионного метода, обеспечивающего наилучшую точность, и выбор существенных (с точки зрения влияния на выходные показатели) входных переменных. Для этого приходится учитывать, что рассматриваемая промышленная ректификационная колонна функционирует в режиме, близком к предельному (захлебывание), а это существенно усложняет процесс построения ММОПК из-за нелинейности физико-химической модели данного режима.

Синтетический пример. Пусть объект задается функциональной зависимостью вида:

$$Y = \sin x_1 + \sin 1,7x_2 + \cos 3,4x_3 + \cos 2,4x_4 + 0,01x_5. \quad (7)$$

По уравнению (7) и входным переменным $x_i, i = 1, \dots, 5$, на которые наложены ограничения $-2,5 \leq x_1, \dots, x_5 \leq 2,5$, формируется выборка объемом $K = 500$, представляющая матрицу размером $(K \times 5)$. Применяя алгоритм АСЕ к расширенной выборке, формируем базовый набор векторов оптимальных преобразований $\Phi_i(x_i)$, графически представленный на рис. 1 и свидетельствующий о достаточно точно найденной структуре модели.

По полученному представлению о парных отношениях (см. рис. 1) можно сделать вывод, что параметрические методы построения модели практически неэффективны. В свою очередь, применение алгоритма АСЕ позволяет восстановить соответствующий вид функции в уравнении (7).

В случае отсутствия нелинейных зависимостей алгоритм АСЕ предлагает линейные преобразования, что приводит к хорошим результатам.

Предлагаемый алгоритм построения математической модели для оценки показателей качества выходных продуктов сравнен с математическими моделями, полученными: нейронной сетью прямого рас-

пространения с одним скрытым слоем (НС-1), прямого распространения с пятью скрытыми слоями (НС-2) и обобщенно-регрессионной нейронной сетью (НС-3). В качестве обучающего алгоритма выбран алгоритм Левенберга – Марквардта [32].

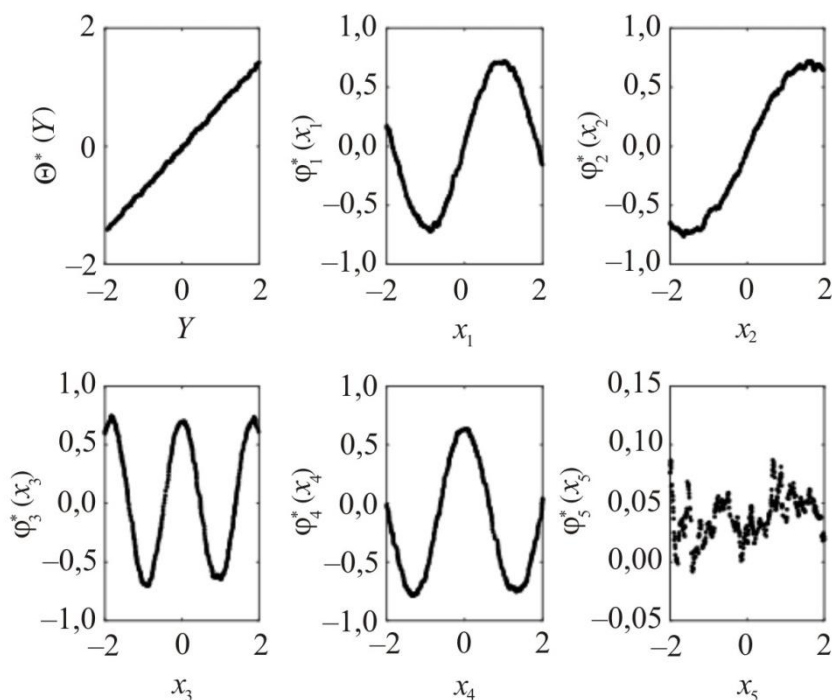


Рис. 1. Оптимальные преобразования алгоритма ACE к синтетическим данным

Сравнительный анализ проводили на основе критериев точности математической модели: коэффициента детерминации R^2 средней абсолютной ошибки MAE на обучающей и проверочной выборках (табл. 1).

По результатам, представленным в табл. 1, показано, что наиболее точно описывает исследуемый МТП непараметрическая модель, построенная на основе алгоритма ACE, по сравнению с другими методами. Результат применения непараметрической модели, построенной на основе алгоритма ACE, и данные модели (7) приведены на рис. 2.

Таблица 1

Значения R^2 и MAE для математических моделей, полученных различными методами

Метод получения ММОПК	Обучение		Проверка	
	R^2	MAE	R^2	MAE
НС-1	0,3168	0,9816	0,3139	0,9534
НС-2	0,5402	0,5739	0,5588	0,5928
НС-3	0,6936	0,6541	0,3860	0,9215
На основе алгоритма ACE	0,9996	0,0028	0,9709	0,1254

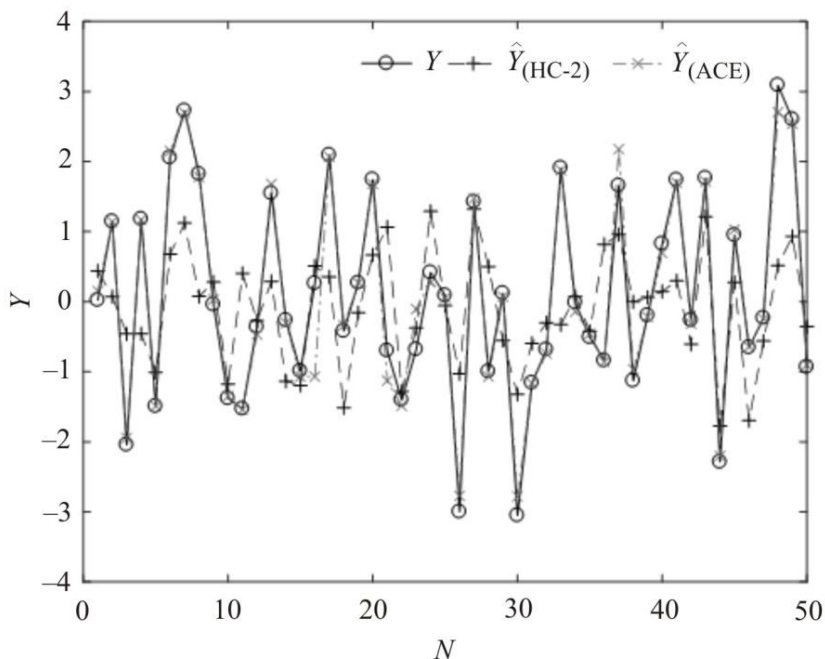


Рис. 2. Функционирование непараметрической модели на проверочной выборке

Из рис. 2 можно сделать вывод, что получена точная математическая модель, о чем свидетельствуют критерии точности, приведенные в табл. 1, например $R^2 = 0,9709$.

3. Построение математической модели для оценки показателей качества выходного продукта на основе экспериментальных данных

Эффективность предлагаемого подхода продемонстрирована на промышленном технологическом объекте – ректификационной колонне К-1 процесса производства высокооктановой добавки бензинов – метил-трет-бутилового эфира (МТБЭ). Схема РК К-1 и процесс массообмена в ней представлены на рис. 3.

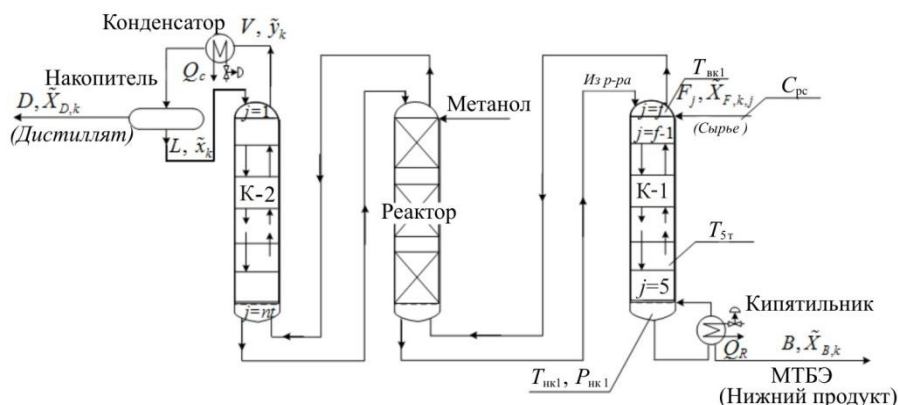


Рис. 3. Схематическое изображение процесса производства МТБЭ

При осуществлении синтеза МТБЭ особое внимание уделяется выбору соотношения изобутилен-метанол. Определение и поддержание соотношения компонентов осуществляется на основании опыта технолога. Неправильно заданное количество метанола приводит к перерасходу исходного сырья, недополучению готовой продукции и снижению ее качества. По окончании процесса выход МТБЭ должен составлять 98 %. Основным показателем качества процесса является содержание метанола в МТБЭ.

Для исследуемого объекта необходимо построить математическую модель с целью оценки концентрации метанола в нижнем (выходном) продукте при неизвестной структуре. Массовая доля метанола не должна превышать 1,5 % в товарном продукте МТБЭ по техническому регламенту.

Номинальные значения четырех технологических параметров (входных переменных) РК К-1, используемых при создании математической модели для оценки содержания метанола в МТБЭ, приведены в табл. 2.

Таблица 2

Наименование технологических переменных РК К-1 процесса
производства МТБЭ

№ п/п	Наименование переменных	Обозначение	Номинальные значения
x_1	Температура низа РК, °С	$T_{\text{НК1}}$	136,00
x_2	Давление низа РК, МПа _г	$P_{\text{НК1}}$	0,72
x_3	Температура на 5-й тарелке РК, °С	$T_{5\text{T}}$	125,00
x_4	Температура верха РК, °С	$T_{\text{ВК1}}$	60,00

Для построения математической модели с целью оценки концентрации метанола в нижнем продукте в качестве входных переменных модели использованы: x_1 – температура низа колонны К-1, °С ($T_{\text{НК1}}$) и x_2 – давление низа колонны К-1, МПа_г ($P_{\text{НК1}}$).

Применяя алгоритм АСЕ к экспериментальным данным, формируем базовый набор векторов оптимальных преобразований $\Phi_i(x_i)$, графически представленный на рис. 4.

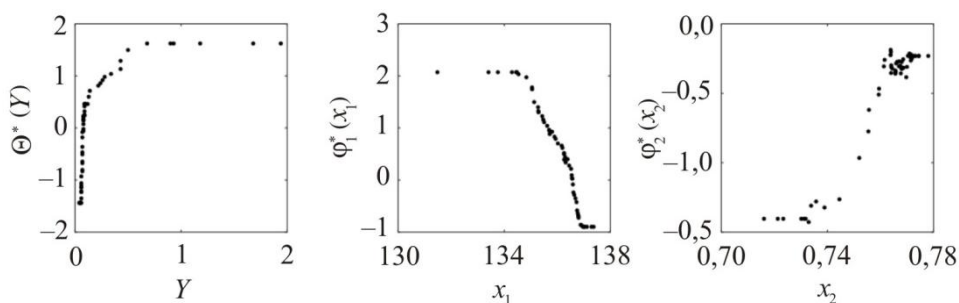


Рис. 4. Оптимальные преобразования алгоритма АСЕ
к экспериментальным данным технологического процесса

Следует отметить, что алгоритм АСЕ не будет работать с экспериментальными данными так же хорошо, как на синтетическом примере. Это может быть связано с тем, что: а) зависимая переменная имеет более низкую связь с предикторами; б) предикторы существенно коррелированы; в) данные содержат погрешности измерений, и др.

Из рис. 4 видно, что исследуемый объект нелинейный, и применение алгоритма АСЕ позволяет найти некоторый вид существующей функциональной зависимости.

В табл. 3 представлены R^2 и MAE математических моделей для оценки метанола в МТБЭ, полученных различными методами.

Таблица 3

Значения R^2 и MAE для математических моделей с целью оценки концентрации метанола в выходном продукте

Метод получения ММОПК	Обучение		Проверка	
	R^2	MAE	R^2	MAE
НС-1	0,6078	0,1725	0,7382	0,1548
НС-2	0,9875	0,0258	0,7439	0,0918
НС-3	0,6100	0,1298	0,6940	0,1102
На основе алгоритма ACE	0,9996	0,0032	0,9820	0,0301

По результатам, представленным в табл. 3, показано, что наиболее точную оценку содержания метанола в МТБЭ дает непараметрическая модель, построенная на основе алгоритма ACE, в сравнении с другими методами.

Из рис. 5 видно, что использование непараметрической модели, построенной на основе алгоритма ACE, для оценки содержания метанола в высокооктановой добавке к бензинам на проверочной выборке позволяет достаточно точно описать исследуемый процесс.

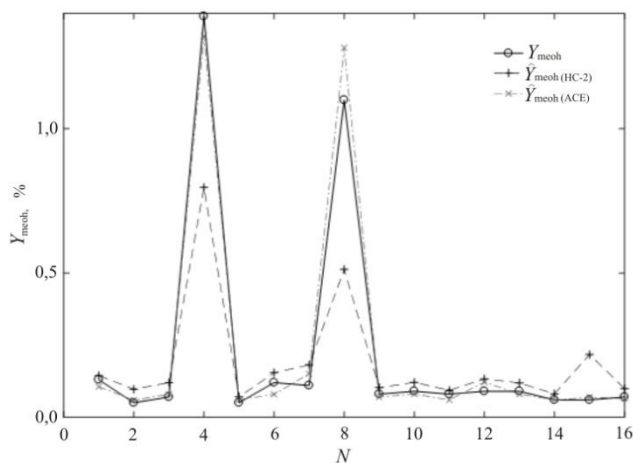


Рис. 5. Оценка концентрации метанола в МТБЭ на основе непараметрической модели

При использовании предложенного подхода результат работы математической модели для оценки показателей качества выходного продукта точнее на $(0,0918 - 0,0301 / 0,0918) \cdot 100 \% \approx 67,2 \%$ в сравнении с НС-2 по критерию МАЕ (см. табл. 3).

Заключение

Предложен подход для построения математических моделей с целью оценки качества выходных продуктов с использованием непараметрического алгоритма АСЕ, что позволило повысить точность математической модели в условиях структурной неопределенности.

Эффективность предложенного подхода продемонстрирована на синтетическом примере. Показан достаточно точный вид структуры модели с использованием векторов оптимальных преобразований.

Предложенный подход протестирован на технологическом процессе производства МТБЭ и позволил повысить точность математической модели для оценки концентрации метанола в выходном продукте до 67,2 % МАЕ в сравнении с моделью, полученной методом нейронной сети прямого распространения с пятью скрытыми слоями, на проверочной выборке.

Благодарность

Исследование выполнено при финансовой поддержке РФФИ и Государственного фонда естественных наук Китая в рамках научного проекта № 21-57-53005.

Список литературы

1. Greene W.H. On the asymptotic bias of the ordinary least squares estimator of the Tobit model // *Econometrica*. – 1981. – Vol. 49, no. 2. – P. 505–513.
2. Khatibisepehr S., Huang B. Dealing with irregular data in soft sensors: bayesian method and comparative study // *Ind. Eng. Chem. Res.* – 2008. – Vol. 47. – P. 8713–8723.
3. Berger B. Robust Gaussian process modelling for engine calibration / B. Berger, F. Rauscher // *IFAC Proceeding Volumes [Proceedings of the 7th Vienna International Conference on Mathematical Modelling, Vienna, Austria, 14–17 February, 2012]*. – Vienna, Austria, 2012. – Vol. 45, iss. 2. – P. 159–164. DOI: 10.3182/20120215-3-AT-3016/0028.

4. Cateni S., Colla V., Vannucci M. Outlier detection methods for industrial applications // *Advances in Robotics, Automation and Control* / Ed. by J. Arámburo, A.R. Trevino. – Vienna: I-Tech, 2008. – P. 265–282.

5. Marquardt D., Snee R. Ridge regression on practice // *The American statistician*. – 1975. – Vol. 29, no 1. – P. 37–41.

6. Khalaf G., Shukur G. Choosing ridge parameter for regression problems // *Communications in Statistics. – Theory and Methods*. – 2005. – Vol. 34. – P. 1177–1182.

7. Saunders C., Gammernan A., Vovk V. Ridge regression learning algorithm in dual variables // *Proceedings of the 15th International Conference on Machine Learning (ICML 1998)*, Madison, Wisconsin, USA, July 24–27, 1998. – San Francisco: Morgan Kaufmann Publishers Inc., 1998. – P. 515–521.

8. Hoerl A., Kennard R. Ridge regression: biased estimation for nonorthogonal problems // *Technometrics*. – 1970. – Vol. 12, no 1. – P. 55–67.

9. Yan X. Modified nonlinear generalized ridge regression and its application to develop naphtha cut point soft sensor // *Computers and Chemical Engineering*. – 2008. – Vol. 32. – P. 608–621.

10. Wold S., Sjostrom M., Eriksson L. PLS-regression: a basic tool of chemometrics // *Chemometrics and Intelligent Laboratory System*. – 2001. – Vol. 58. – P. 109–130.

11. De Jong S. SIMPLS: An Alternative Approach to Partial Least Squares Regression // *Chemometrics and Intelligent Laboratory Systems*. – 1993. – Vol. 18. – P. 251–263.

12. Souza F.A., Araujo R. Mixture of partial least squares experts and application in prediction settings with multiple operating modes // *Chemometrics and Intelligent Laboratory Systems*. – 2014. – Vol. 130. – P. 192–202.

13. Data-based identifiability analysis of non-linear dynamical models / S. Hengl, C. Kreutz, J. Timmer, T. Maiwald // *Bioinformatics*. – 2007. – Vol. 23, no. 19. – P. 2612–2618.

14. Di Ciaccio A., Montanari G.E. Non-parametric methods for data-mining applications // *Atti della XLI Riunione Scientifica Della Societa Italiana di Statistica. Sessioni Plenarie e Sessioni Specializzate*. CLEUP – Padova, 2002. – P. 339–348.

15. Tun M.S., Lakshminarayanan S., Emoto G. Data selection and regression method and its application to soft sensing using multi rate industrial data // *Chemical Engineering of Japan*. – 2008. – Vol. 41, no 5. – P. 374–383.

16. Kadlec P., Gabrys B., Strandt S. Data-driven soft sensors in the process industry // *Computers & Chemical Engineering*. – 2009. – Vol. 33. – P. 795–814.

17. Genetic fuzzy system for data-driven soft sensors design / J. Mendes, F. Souza, R. Araújo, N. Goncalves // *Applied Soft Computing*. – 2012. – Vol. 12. – P. 3237–3245.

18. Angelov P., Filev D. An approach to on-line identification of evolving Takagi-Sugeno models // *IEEE Trans. on Systems, Man and Cybernetics*. – 2004. – Vol. 34, no 1. – P. 484–498.
19. Sedoglavic A. A probabilistic algorithm to test local algebraic observability in polynomial time // *Journal of Symbolic Computation*. – 2002. – Vol. 33, no. 5. – P. 735–755.
20. Daş M.T., Dülger L.C. Off-line signature verification with PSO-NN algorithm // *2007 22nd International Symposium on Computer and Information Sciences, Ankara, Turkey, 7–9 November 2007*. – IEEE, 2007. – Art. 9940017. DOI: 10.1109/ISCIS.2007.4456842
21. Box G.E.P., Tidwell P.W. Transformation of the independent variables // *Technometrics*. – 1962. – Vol. 4, no 4. – P. 531–550.
22. Anscombe F.J., Tukey J.W. The examination and analysis of residuals // *Technometrics*. – 1963. – Vol. 5, no 2. – P. 141–160.
23. Box G.E.P., Cox D.R. An analysis of transformations // *Journal of the Royal Statistical Society. Series B (Methodological)*. – 1964. – Vol. 26, no. 2. – P. 211–252.
24. Kruskal J.B. Analysis of factorial experiments by estimating monotone transformations of the data // *Journal of the Royal Statistical Society. Series B (Methodological)*. – 1965. – Vol. 27, no 2. – P. 251–263.
25. Draper N.R., Cox D.R. On distributions and their transformation to normality // *Journal of the Royal Statistical Society. Series B (Methodological)*. – 1969. – Vol. 31, no 3. – P. 472–476.
26. Box G.E.P., Hill W.J. Correcting inhomogeneity of variance with power transformation weighting // *Technometrics*. – 1974. – Vol. 16, no 3. – P. 385–389.
27. Walter E., Pronzato L. On the identifiability and distinguishability of nonlinear parametric models // *Mathematics and Computers in Simulation*. – 1996. – Vol. 42. – P. 125–134.
28. Bellman R., Astrom K.J. On structural Identifiability // *Mathematical Biosciences*. – 1970. – Vol. 7, no 3/4. – P. 329–339.
29. Chappell M.J., Godfrey K.R. Structural Identifiability of the parameters of a nonlinear batch reactor model // *Mathematical Biosciences*. – 1992. – Vol. 108. – P. 241–251.
30. Cobelli C., Distefano J.J. Parameter and structural identifiability concepts and ambiguities: a critical review and analysis // *American Physiological Society*. – 1980. – Vol. 239, no 1. – P. 7–24.
31. Wang D., Murphy M. Estimating optimal transformations for multiple regression using the ACE algorithm // *Journal of Data Science*. – 2004. – Vol. 2. – P. 329–346.
32. Ranganathan A. The levenberg-marquardt algorithm // *Tutorial on LM Algorithm*. – 2004. – Vol. 11, no. 1. – P. 101–110.

References

1. Greene W.H. On the asymptotic bias of the ordinary least squares estimator of the Tobit model. *Econometrica*, 1981, vol. 49, no. 2, pp. 505–513.
2. Khatibisepehr S., Huang B. Dealing with irregular data in soft sensors: Bayesian method and comparative study. *Ind. Eng. Chem. Res.*, 2008, vol. 47, pp. 8713–8723.
3. Berger B., Rauscher F. Robust Gaussian process modelling for engine calibration. *IFAC Proceedings Volumes [Proceedings of the 7th Vienna International Conference on Mathematical Modelling]*, 2012, vol. 45, iss. 2, pp. 159–164. DOI: 10.3182/20/20215-3-AT-3016.00028.
4. Cateni S., Colla V., Vannucci M. Outlier detection methods for industrial applications. *Advances in Robotics. Automation and Control*. Ed by J. Arámburo, A.R. Treviño. Vienna, I-Tech, 2008, pp. 265–282.
5. Marquardt D., Snee R. Ridge regression on practice. *The American statistician*, 1975, vol. 29, no. 1, pp. 37–41.
6. Khalaf G., Shukur G. Choosing Ridge Parameter for Regression Problems. *Communications in statistics – theory and methods*, 2005, vol. 34, pp. 1177–1182.
7. Saunders C., Gammernan A., Vovk V. Ridge regression learning algorithm in dual variables. *Proceedings of the 15th International Conference on Machine Learning (ICML. 1998)* [Madison, Wisconsin, USA, July 24–27, 1998]. San Francisco, Morgan Kaufmann Publishers Inc., 1998. pp. 515–521.
8. Hoerl A., Kennard R. Ridge regression: biased estimation for nonorthogonal problems. *Technometrics*, 1970, vol. 12, no. 1, pp. 55–67.
9. Yan X. Modified nonlinear generalized ridge regression and its application to develop naphtha cut point soft sensor. *Computers and Chemical Engineering*, 2008, vol. 32, pp. 608–621.
10. Wold S., Sjostrom M., Eriksson L. PLS-regression: a Basic Tool of Chemometrics. *Chemometrics and Intelligent Laboratory System*, 2001, vol. 58, pp. 109–130.
11. De Jong S. SIMPLS: An Alternative Approach to Partial Least Squares Regression. *Chemometrics and Intelligent Laboratory Systems*, 1993, vol. 18, pp. 251–263.
12. Souza F.A., Araujo R. Mixture of partial least squares experts and application in prediction settings with multiple operating modes. *Chemometrics and Intelligent Laboratory Systems*, 2014, vol. 130, pp. 192–202.
13. Hengl S., Kreutz C., Timmer J., Maiwald T. Data-based identifiability analysis of non-linear dynamical models. *Bioinformatics*, 2007, vol. 23, no. 19, pp. 2612–2618.
14. Di Ciaccio A., Montanari G.E. Non-parametric methods for data-mining applications. *Atti della XLI Riunione scientifica della Societa italiana di Statistica. Sessioni plenarie e Sessioni specializzate*. CLEUP – Padova, 2002, pp. 339–348.

15. Tun M.S., Lakshminarayanan S., Emoto G. Data selection and regression method and its application to soft sensing using multi rate industrial data. *Chemical Engineering of Japan*, 2008, vol. 41, no 5, pp. 374–383.
16. Kadlec P., Gabrys B., Strandt S. Data-driven soft sensors in the process industry. *Computers & Chemical Engineering*, 2009, vol. 33, pp. 795–814.
17. Mendes J., Souza F., Araújo R., Goncalves N. Genetic fuzzy system for data-driven soft sensors design. *Applied Soft Computing*, 2012, vol. 12, pp. 3237–3245.
18. Angelov P., Filev D. An approach to on-line identification of evolving Takagi-Sugeno models. *IEEE Trans. on Systems, Man and Cybernetics*, 2004, vol. 34, no. 1, pp. 484–498.
19. Sedoglavic A. A probabilistic algorithm to test local algebraic observability in polynomial time. *Proceedings of the International Symposium on Symbolic and algebraic Computation*, 2001, pp. 309–317.
20. Daş M.T., Dülger L.C. Off-line signature verification with PSO-NN algorithm. *2007 22nd International Symposium on Computer and Information Sciences [Ankara, Turkey, 7–9 November 2007]. IEEE, 2007.* 6 p., art. 9940017. DOI: 10.1109/ISCIS.2007.4456842
21. Box G.E.P., Tidwell P.W. Transformation of the independent variables. *Technometrics*, 1962, vol. 4, no. 4, pp. 531–550.
22. Anscombe F. J., Tukey J. W. The examination and analysis of residuals. *Technometrics*, 1963, vol. 5, no. 2, pp. 141–160.
23. Box G.E.P., Cox D.R. An Analysis of transformations, *Journal of the Royal Statistical Society. Series B (Methodological)*, 1964, vol. 26, no. 2, pp. 211–252.
24. Kruskal J. B. Analysis of factorial experiments by estimating monotone transformations of the data. *Journal of the Royal Statistical Society. Series B (Methodological)*, 1965, vol. 27, no. 2, pp. 251–263.
25. Draper N.R., Cox D.R. On distributions and their transformation to normality. *Journal of the Royal Statistical Society. Series B (Methodological)*, 1969, vol. 31, no. 3, pp. 472–476.
26. Box G.E.P, Hill W.J. Correcting inhomogeneity of variance with power transformation weighting. *Technometrics*, 1974, vol. 16, no. 3, pp. 385–389.
27. Walter E., Pronzato L. On the identifiability and distinguishability of nonlinear parametric models. *Mathematics and Computers in Simulation*, 1996, vol. 42, pp. 125–134.
28. Bellman R., Astrom K.J. On structural Identifiability. *Mathematical Biosciences*, 1970, vol. 7, no³/₄, pp. 329–339.
29. Chappell M.J., Godfrey K.R. Structural Identifiability of the parameters of a nonlinear batch reactor model. *Mathematical Biosciences*, 1992, vol. 108, pp. 241–251.
30. Cobelli C., Distefano J.J. Parameter and structural identifiability concepts and ambiguities: a critical review and analysis. *American Physiological Society*, 1980, vol. 239, no. 1, pp. 7–24.

31. Wang D., Murphy M. Estimating optimal transformations for multiple regression using the ACE algorithm. *Journal of Data Science*, 2004, vol. 2, pp. 329–346.

32. Ranganathan A. The levenberg-marquardt algorithm. *Tutorial on LM algorithm*, 2004, vol. 11, no. 1, pp. 101–110.

Статья получена: 20.06.2021

Статья принята: 25.06.2021

Сведения об авторах

Можаровский Игорь Сергеевич (Владивосток, Россия) – старший преподаватель, кафедра «Информационные технологии и системы», Владивостокский государственный университет экономики и сервиса (690014, Владивосток, ул. Гоголя, 41, e-mail: studvvsu@gmail.com); младший научный сотрудник, лаборатория информационно-аналитических и управляющих систем и технологий, Институт автоматизации и процессов управления ДВО РАН (690041 Россия, Владивосток, ул. Радио, 5).

Самотылова Светлана Александровна (Владивосток, Россия) – кандидат технических наук, младший научный сотрудник, лаборатория систем управления технологическими процессами, Институт автоматизации и процессов управления ДВО РАН (690041, Владивосток, ул. Радио, 5, e-mail: samotylova@dvo.ru); ассистент, департамент компьютерно-интегрированных производственных систем Политехнического института (Школы), Дальневосточный федеральный университет (690922, Владивосток, о. Русский, п. Аякс, 10).

Баулин Евгений Сергеевич (Москва, Россия) – кандидат технических наук, директор по разработкам и исследовательской деятельности, ООО «Центр цифровых технологий» (121205, Москва, ул. Нобеля, 7, e-mail: Baulin.es@mipt.ru).

About the authors

Igor S. Mozharovskii (Vladivostok, Russian Federation) – Senior Lecturer, Department of Information Technologies and Systems, Vladivostok State University of Economics and service (41, Gogolya st., Vladivostok, 690014, e-mail: studvvsu@gmail.com); Junior Researcher, Laboratory of Information-Analytical and Control Systems and Technologies, Institute of Automation and Control (5, Radio st., Vladivostok, 690041).

Svetlana A. Samotylova (Vladivostok, Russian Federation) – Ph.D. in Engineering, Junior Researcher, Laboratory of Technological Process Control Systems, Institute of Automation and Control (5, Radio st., Vladivostok, 690041,

e-mail: samotylova@dvo.ru); Assistant, Department of Computer-Integrated Manufacturing Systems of the Polytechnic Institute (School), Far Eastern Federal University (10, Ajax, Russky Island, Vladivostok, 690922).

Evgenii S. Baulin (Moscow, Russian Federation) – Ph.D. in Engineering, Research and Development Director, LLL “Center of Digital Technology” (7, Nobelya st., Moscow, 121205, e-mail: Baulin.es@mipt.ru).

**Библиографическое описание статьи согласно
ГОСТ Р 7.0.100–2018:**

Можаровский, И.С. Использование алгоритма условных чередующихся математических ожиданий для построения математических моделей для оценки показателей качества выходных продуктов / И. С. Можаровский, С. А. Самотылова, Е. С. Баулин. – текст : непосредственный. – DOI: 10.15593/2499-9873/2021.2.04 // Прикладная математика и вопросы управления = Applied Mathematics and Control Sciences. – 2021. – № 2. – С. 53–70.

Цитирование статьи в изданиях РИНЦ:

Можаровский И.С., Самотылова С.А., Баулин Е.С. Использование алгоритма условных чередующихся математических ожиданий для построения математических моделей для оценки показателей качества выходных продуктов // Прикладная математика и вопросы управления. – 2021. – № 2. – С. 53–70. DOI: 10.15593/2499-9873/2021.2.04

Цитирование статьи в references и международных изданиях

Cite this article as:

Mozharovskii I.S., Samotylova S.A., Baulin E.S. Using the algorithm of conditional alternating mathematical expectations to mathematical models design to estimate the quality indicators of output products. Applied Mathematics and Control Sciences, 2021, no. 2, pp. 53–70. DOI: 10.15593/2499-9873/2021.2.04 (*in Russian*)