

УДК 681.51

# РАЗРАБОТКА МОДЕЛЕЙ ПОКАЗАТЕЛЕЙ КАЧЕСТВА РЕКТИФИКАЦИОННЫХ КОЛОНН, ФУНКЦИОНИРУЮЩИХ В ПРЕДЕЛЬНЫХ РЕЖИМАХ\*

**А.Ю. Торгашов**

*Институт автоматике и процессов управления ДВО РАН*  
Россия, 690041, Владивосток, Радио ул., 5  
E-mail: [torgashov@iacp.dvo.ru](mailto:torgashov@iacp.dvo.ru)

**И.С. Можаровский**

*Институт автоматике и процессов управления ДВО РАН*  
Россия, 690041, Владивосток, Радио ул., 5  
E-mail: [studvvsu@gmail.com](mailto:studvvsu@gmail.com)

**Г.Б. Диго**

*Институт автоматике и процессов управления ДВО РАН*  
Россия, 690041, Владивосток, Радио ул., 5  
E-mail: [bernatsk@iacp.dvo.ru](mailto:bernatsk@iacp.dvo.ru)

**Н.Б. Диго**

*Институт автоматике и процессов управления ДВО РАН*  
Россия, 690041, Владивосток, Радио ул., 5  
E-mail: [digo@iacp.dvo.ru](mailto:digo@iacp.dvo.ru)

**Ключевые слова:** ректификационная колонна, показатель качества, идентификация, виртуальный анализатор.

Обсуждается один из подходов к построению моделей показателей качества массообменных процессов (в ректификационных колоннах), функционирующих в предельных режимах. Излагается методика определения оптимального количества входных переменных моделей показателей качества (виртуальных анализаторов) на примере промышленной ректификационной колонны, работающей в режиме, близком к состоянию захлебывания.

## DEVELOPING MODELS OF QUALITY INDEXES OF RECTIFYING COLUMNS, PERFORMING UNDER LIMIT MODES \*

**A.Yu. Torgashov**

*Institute of Automatics and Control Processes  
of the Far-Eastern Department of the Russian Academy of Sciences*

\*Работа выполнена при финансовой поддержке гранта РФФИ-ДВО РАН 11-08-98500-р\_восток\_а «Математическое моделирование предельных режимов функционирования массообменных технологических процессов для задач управления».

\*Работа выполнена при финансовой поддержке гранта РФФИ-ДВО РАН 11-08-98500-р\_восток\_а «Математическое моделирование предельных режимов функционирования массообменных технологических процессов для задач управления».

Russia, 690041, Vladivostok, Radio street, 5  
E-mail: [torgashov@iacp.dvo.ru](mailto:torgashov@iacp.dvo.ru)

**I.S. Mozharovsky**

*Institute of Automatics and Control Processes  
of the Far-Eastern Department of the Russian Academy of Sciences*  
Russia, 690041, Vladivostok, Radio street, 5  
E-mail: [studvvsu@gmail.com](mailto:studvvsu@gmail.com)

**G.B. Digo**

*Institute of Automatics and Control Processes  
of the Far-Eastern Department of the Russian Academy of Sciences*  
Russia, 690041, Vladivostok, Radio street, 5  
E-mail: [bernatsk@iacp.dvo.ru](mailto:bernatsk@iacp.dvo.ru)

**N.B. Digo**

*Institute of Automatics and Control Processes  
of the Far-Eastern Department of the Russian Academy of Sciences*  
Russia, 690041, Vladivostok, Radio street, 5  
E-mail: [digo@iacp.dvo.ru](mailto:digo@iacp.dvo.ru)

**Key words:** rectifying column, quality index, identification, soft sensor.

An approach to constructing models of quality indexes of mass-transfer processes (in rectifying columns) is discussed, performing under limit modes. A procedure of determining the optimal number of input variables of models of quality indexes (soft sensors) is presented by use of an example of a rectifying column performing under a mode being close to the flood status.

## 1. Введение

Одной из важнейших задач в нефтепереработке является задача контроля качества изготавливаемых нефтепродуктов. В настоящее время при оценке качества выходных продуктов нефтеперерабатывающих процессов и, в частности, на ректификационных установках, используются три способа контроля показателей качества: лабораторные анализы, данные поточных анализаторов, данные виртуальных анализаторов (ВА) [1]. Лабораторные анализы достаточно редки и не могут использоваться для управления качеством в реальном времени. Поточные анализаторы постоянно требуют калибровки, чрезвычайно дороги и потому далеко не всегда доступны. В отличие от них виртуальные анализаторы, практически не уступая в точности, значительно дешевле и надежнее. Принцип действия ВА основан на непрерывном определении показателя качества по математической модели, описывающей его взаимосвязь с текущими значениями измеряемых технологических переменных. Они позволяют оценивать не измеряемые непосредственно, но необходимые показатели качества продукта по таким измеряемым параметрам технологического процесса как температура, давление, расход, непрерывно контролируемым современными системами управления.

Для построения ВА обычно используются такие методы как наименьших квадратов (МНК) [2], робастная регрессия (РР) [3], проекции на латентные структуры (ПЛС) [4-5] (случай линейных моделей) и алгоритм АСЕ [6] (нелинейные модели). При этом уделяется недостаточное внимание анализу оптимального количества входов ВА.

В докладе предлагается методика определения оптимального количества входных переменных моделей ВА на примере промышленной ректификационной колонны, работающей в режиме, близком к состоянию захлебывания.

## 2. Постановка задачи

В реальных условиях качество выходного продукта РК анализируется лабораторным путем, это достаточно сложная, дорогостоящая и, что немаловажно, долговременная процедура. Альтернативой такому подходу является переход к виртуальным анализаторам, использующим промышленные данные (данные со встроенных измерительных датчиков) ректификационной колонны. При разработке ВА также применяются некоторые виды гибридных нейронных технологий, адаптивные алгоритмы автоматического управления для непрерывной настройки модели, регрессионные модели и т.д.

Ставится задача разработки моделей ВА с привлечением различных регрессионных методов (МНК, РР, ПЛС, АСЕ) и методики определения оптимального количества входных переменных при построении ВА ректификационной колонны по содержанию бензолобразующих веществ и концентрации изо-пентана в дистилляте. Следует отметить, что рассматриваемая промышленная РК функционирует в режиме, близком к предельному (захлебывание), поэтому существенно усложняется процесс построения ВА из-за нелинейности физико-химической модели данного режима.

## 3. Описание промышленной ректификационной колонны

Процесс массообмена в ректификационной колонне показан на рис. 1.

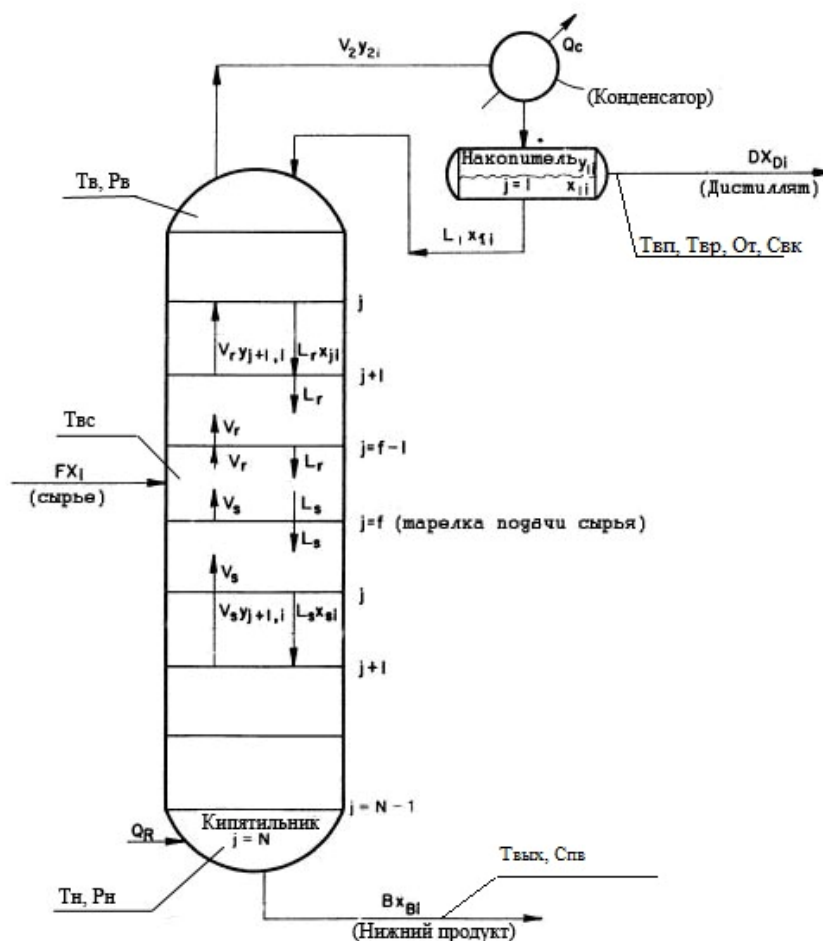


Рис. 1. Схематическое изображение РК.

Средние значения технологических переменных в определенных точках РК, на основе которых будут создаваться ВА, приведены ниже в таблице 1.

Равновесие фаз на каждой ступени разделения РК описывается следующими уравнениями [7]:

$$(1) \quad y_{ij} = K_{ij} x_{ij}$$

при условиях

$$\sum_{i=1}^C y_{ij} = 1, \quad \sum_{i=1}^C x_{ij} = 1,$$

$K_{ij}$  – константа физического равновесия на  $j$ -ой тарелке  $i$ -го компонента (находится по уравнению Антуана [7]),  $x_{ij}$  – концентрация в жидкой фазе на  $j$ -ой тарелке  $i$ -го компонента,  $y_{ij}$  – концентрация в газообразной фазе  $j$ -ой тарелке  $i$ -го компонента,  $C$  – количество компонентов в смеси.

В исследуемом случае количество компонентов  $C = 28$ , ( $i=1, 2, \dots, C$ ), количество тарелок в РК  $N = 44$ , ( $j=1, 2, \dots, N$ ).

Таблица 1. Технологические переменные РК.

Наименование датчика РК	Свк (поток сырья вверху РК) М <sup>3</sup> /ч	Тв (темп. вверху РК) град. С	Рв (давление вверху РК) кг/см <sup>2</sup>	Слк (поток сырья на входе в РК) М <sup>3</sup> /ч	Твс (темп. сырья на входе в РК) град. С	Тн (темп. внизу РК) град. С	Рн (давление внизу РК) кг/см <sup>2</sup>	От (кол-во производ. прод.) Т/ч	Твр (темп. входа прод.) град. С	Спв (скор. выхода продукта) М <sup>3</sup> /ч	Твых (темп. выхода продукта) С	Тыв (темп. верх. вых. продукта) С
	FIC605	TIC653	PI643	FIC606	TI649	TIC663	PI644	FIC622	TI666	FI607	TI654	TI667
Знач. датч.	46	88	2	72	85	120	3	30	73	22	194	77

Уравнения материального баланса имеют вид:

$$(2) \quad \begin{aligned} V_{j+1} y_{j+1i} &= L_j x_{ji} + DX_{ji} & (j = 1, 2, \dots, f-2) \\ V_f y_{fi} + V_F y_{Fi} &= L_{f-1} x_{f-1i} + DX_{Di} \\ V_{j+1} y_{j+1i} &= L_j x_{ji} - Bx_{Fi} \\ FX_i &= DX_{Di} + Bx_{Fi} & (j = f, f+1, \dots, N-1) \end{aligned}$$

где  $V_j$  – паровой поток, покидающий  $j$ -ую тарелку,  $L_j$  – поток жидкости, покидающий  $j$ -ую тарелку,  $D$  – дистиллят (верхний продукт РК),  $B$  – нижний продукт в РК,  $N$  – количество ступеней в РК,  $FX_i$  – количество  $i$ -го компонента в сырье, поступающем в РК на тарелку  $c$ ,  $j$  – номер тарелки в РК, на которую подается сырье,  $X_{Di}$  – концентрация  $i$ -го компонента в дистилляте.

Уравнения энергетического баланса имеют вид:

$$(3) \quad \begin{aligned} V_{j+1} H_{j+1} &= L_j h_j + DH_D & (j = 1, 2, \dots, f-2) \\ V_f H_f + V_F H_F &= L_{f-1} h_{f-1} + DH_D + Q_c \\ V_{j+1} H_{j+1} &= L_j h_j - B h_B + Q_R & (j = f, f+1, \dots, N-1) \\ FH &= B h_B + DH_D + Q_c - Q_R \end{aligned}$$

где  $H_j$  – энтальпия паров на  $j$ -ой тарелке,  $h_j$  – энтальпия жидкости на  $j$ -ой тарелке,  $Q_c$  – тепловая нагрузка конденсатора,  $Q_R$  – тепловая нагрузка кипятильника.

Решение системы уравнений (1)-(3) является сложным и трудоемким, что не позволяет проводить вычисления в режиме on-line (сложность и время вычисления увеличиваются в разы с каждым новым компонентом  $C$  в исследуемой системе) и использовать ее как рабочую модель ВА.

Создание ВА описывается на примере двух показателей качества: содержание бензолобразующих веществ и концентрация изо-пентана в дистилляте. Для этого имеются необходимые данные лабораторных исследований верхнего продукта (дистиллята) РК и данные датчиков РК в заданное время.

#### 4. Характеристика регрессионных методов, используемых при построении моделей ВА

Проведенные вычислительные эксперименты с промышленными данными РК для построения моделей ВА регрессионными методами показали, что в случае явно выраженной нелинейности алгоритм ACE дает более точные результаты, чем регрессионные методы МНК, МЛР, ПЛС, ПЛС2, РР [8]. Модель, получаемая по алгоритму ACE [6, 9], имеет следующую общую форму:

$$(4) \quad \theta(Y) = \alpha + \sum_{i=1}^p \phi_i(X_i) + \varepsilon,$$

где  $\theta$  – функция выходной переменной  $Y$  и  $\phi_i$  – функции входов  $X_i$ ,  $i = 1, \dots, p$ ,  $\varepsilon$  – ошибка, т.е. этот алгоритм позволяет переходить от классического уравнения регрессии, линейного по входным переменным, к уравнению типа (4).

Алгоритм ACE для заданного набора данных, состоящего из выходной переменной  $Y$  и входных переменных  $X_1, X_2, \dots, X_p$ , начинает свою работу с вычисления некоторых начальных преобразований  $\theta(Y), \phi_1(X_1), \dots, \phi_p(X_p)$ . Используя их, вычисляется

$$(5) \quad \varepsilon^2(\theta, \phi_1, \dots, \phi_p) = E\left\{[\theta(Y) - \sum_{i=1}^p \phi_i(X_i)]^2\right\}.$$

Минимизация  $\varepsilon^2$  из (5) по  $\phi_1(X_1), \dots, \phi_p(X_p)$  и  $\theta(Y)$  достигается посредством минимизации функций, заданных уравнениями:

$$\begin{aligned} \phi_i(X_i) &= E[\theta(Y) - \sum_{i=1}^p \phi_i(X_i) | X_i], \\ \theta(Y) &= E\left[\sum_{i=1}^p \phi_i(X_i) | Y\right] / \left\| \left[\sum_{i=1}^p \phi_i(X_i) | Y\right] \right\|. \end{aligned}$$

Процесс минимизации  $\varepsilon^2$  осуществляется последовательно по каждой из переменных, для которых вычисляются условные математические ожидания [5].

После минимизации  $\phi_i(X_i)$ ,  $i = 1, \dots, p$ , и  $\theta(Y)$  получаем оптимальные преобразования  $\phi_i^*(X_i)$ ,  $i = 1, \dots, p$  и  $\theta^*(Y)$ . В результате получаем уравнение:

$$\theta^*(Y) = \sum_{i=1}^p \phi_i^*(X_i) + \varepsilon^*,$$

где  $\varepsilon^*$  – ошибка, не устраняемая использованием алгоритма АСЕ. Минимум ошибки регрессии  $\varepsilon^*$  и максимум коэффициента множественной корреляции  $\rho^*$  связаны соотношением  $\varepsilon^{*2} = 1 - \rho^{*2}$ .

Оптимальные преобразования в алгоритме АСЕ не требуют априорных предположений о какой-либо конкретной функции, связывающей выход и входные переменные. Таким образом, алгоритм АСЕ оказывается мощным средством для исследования данных и их взаимосвязей.

Алгоритм АСЕ, основанный на нелинейных преобразованиях входных переменных, дает более точную модель при построении ее по данным с явной нелинейной зависимостью. Это относится и к данным РК, так как протекающие в ней процессы являются нелинейными. Следовательно, к данным РК можно применять алгоритм АСЕ, получая более точную, по сравнению с традиционными методами линейной регрессии, модель.

## 5. Выбор оптимального количества входных переменных при построении моделей виртуальных анализаторов

При построении моделей ВА показателей качества продукта промышленной РК возникает проблема выбора существенных (с точки зрения влияния на выходные показатели) входных переменных. В силу отсутствия теоретического обоснования такого выбора для технологических процессов ректификации приходится проводить вычислительные эксперименты. Для выбора оптимального (в смысле максимального влияния на величину коэффициента детерминации) количества входов предлагается методика, опробованная на доступных промышленных данных (см. таблицу 1), выбранных с учетом физико-химических процессов, влияющих на качество продуктов в РК.

Согласно этой методике имеющиеся исходные данные разбиваются на обучающую и обобщающую выборки размерами  $24 \times 12$  и  $18 \times 12$  соответственно.

Для идентификации моделей ВА (содержание бензолобразующих веществ и концентрация изо-пентана в дистилляте) использованы 4 регрессионных метода и алгоритм АСЕ, два набора лабораторных данных (значения  $Y$ ), 12 технологических переменных (значения  $X$ ). Коэффициенты детерминации, полученные для обучающей и обобщающей выборок, представлены в таблицах 2 и 3.

**Таблица 2.** Коэффициенты детерминации моделей построенных на разном количестве входов от 1 – 12, сформированных с точки зрения физического смысла для моделей качества промышленной РК по доле бензолобразующих компонентов.

Метод	Количество входов модели ВА по содержанию бензолобразующих											
	1	2	3	4	5	6	7	8	9	10	11	12
<b>П.ЛС</b>												
<b>R<sup>2</sup>_обуч.</b>	0,156	0,288	0,369	0,446	0,447	0,905	0,935	0,942	0,960	0,960	0,962	0,962
<b>R<sup>2</sup>_обоб.</b>	0,313	0,087	-0,636	0,023	0,086	0,811	0,577	0,556	0,714	0,708	0,725	0,710
<b>МНК</b>												
<b>R<sup>2</sup>_обуч.</b>	0,156	0,288	0,369	0,446	0,447	0,905	0,935	0,942	0,960	0,960	0,962	0,962
<b>R<sup>2</sup>_обоб.</b>	0,313	0,087	-0,636	0,023	0,086	0,811	0,577	0,556	0,714	0,708	0,725	0,710
<b>П.ЛС2</b>												
<b>R<sup>2</sup>_обуч.</b>	-0,24	-0,216	0,247	0,402	0,420	0,899	0,929	0,932	0,954	0,954	0,956	0,957
<b>R<sup>2</sup>_обоб.</b>	-0,30	-0,104	0,096	0,367	0,386	0,933	0,754	0,801	0,815	0,812	0,828	0,805

<b>МЛР</b>												
<b>R<sup>2</sup>_обуч.</b>	-0,24	-0,216	0,247	0,402	0,420	0,899	0,929	0,932	0,954	0,954	0,956	0,957
<b>R<sup>2</sup>_обоб.</b>	-0,30	-0,104	0,096	0,367	0,386	0,933	0,754	0,801	0,815	0,812	0,828	0,805
<b>PP</b>												
<b>R<sup>2</sup>_обуч.</b>	0,156	0,288	0,368	0,446	0,447	0,905	0,934	0,942	0,959	0,960	0,961	0,962
<b>R<sup>2</sup>_обоб.</b>	0,320	0,100	-0,698	-0,009	0,077	0,825	0,589	0,566	0,666	0,681	0,709	0,701
<b>ACE</b>												
<b>R<sup>2</sup>_24+18</b>	-3,44	0,479	0,709	0,893	0,801	0,944	0,932	0,929	0,927	0,928	0,936	0,971
<b>R<sup>2</sup>_18</b>	0,512	-0,067	0,762	0,976	0,981	0,974	0,968	0,949	0,949	0,937	0,937	0,925
<b>R<sup>2</sup>_24</b>	-0,17	-2,702	0,75	0,277	-87,714	0,795	0,974	0,971	0,978	0,964	0,965	0,972

**Таблица 3.** Коэффициенты детерминации моделей построенных на разном количестве входов от 1 – 12, сформированных с точки зрения физического смысла для моделей качества промышленной РК по доли изо-пентана

	Количество входов модели ВА по доли изо-пентана											
Метод	1	2	3	4	5	6	7	8	9	10	11	12
ПЛС												
R <sup>2</sup> _обуч.	0,003	0,430	0,446	0,459	0,499	0,886	0,892	0,950	0,952	0,953	0,954	0,955
R <sup>2</sup> _обоб.	-0,369	-1,785	-2,621	-1,909	-3,284	-0,340	-0,903	-4,015	-3,884	-3,666	-3,930	-3,86
МНК												
R <sup>2</sup> _обуч.	0,003	0,430	0,446	0,459	0,499	0,886	0,892	0,950	0,952	0,953	0,954	0,955
R <sup>2</sup> _обоб.	-0,369	-1,785	-2,621	-1,909	-3,284	-0,340	-0,903	-4,015	-3,884	-3,666	-3,930	-3,86
ПЛС2												
R <sup>2</sup> _обуч.	-0,030	0,328	0,422	0,449	0,465	0,875	0,881	0,917	0,921	0,921	0,921	0,926
R <sup>2</sup> _обоб.	-0,159	-1,619	-1,725	-1,384	-1,358	0,643	0,242	-0,910	-1,356	-1,234	-1,204	-2,57
МЛР												
R <sup>2</sup> _обуч.	-0,030	0,328	0,422	0,449	0,465	0,875	0,881	0,917	0,921	0,921	0,921	0,926
R <sup>2</sup> _обоб.	-0,159	-1,619	-1,725	-1,384	-1,358	0,643	0,242	-0,910	-1,356	-1,234	-1,204	-2,57
РР												
R <sup>2</sup> _обуч.	0,001	0,430	0,446	0,458	0,499	0,877	0,891	0,946	0,950	0,953	0,953	0,954
R <sup>2</sup> _обоб.	-0,279	-1,742	-2,501	-1,818	-3,369	0,173	-0,583	-3,016	-3,234	-3,347	-3,683	-3,68
АСЕ												
R <sup>2</sup> _24+18	-4,088	-0,326	0,619	0,745	0,832	0,947	0,894	0,986	0,984	0,987	0,986	0,979
R <sup>2</sup> _18	0,717	0,850	0,562	0,987	0,988	0,984	0,986	0,968	0,963	0,950	0,950	0,950
R <sup>2</sup> _24	-1,745	0,572	0,346	-3,414	0,911	0,959	0,953	0,958	0,960	0,962	0,966	0,946

В них под номерами 1-12 во вторых строках подразумевается количество входных переменных, включенных в модель согласно порядку следования в таблице 1.

Из таблиц 2, 3 следует, что для обеих выборок наибольший коэффициент детерминации  $R^2$  получен для шести входных переменных, и его максимальное значение достигается при использовании метода ПЛС.

Для нахождения оптимального количества входных переменных также используется матрица парной корреляции между выходом модели ВА и 12 входными переменными (таблицы 4, 5).

**Таблица 4.** Коэффициенты парной корреляции.

у	1	2	3	4	5	6	7	8	9	10	11	12
<b>Имя входа</b>	<b>TIC653</b>	<b>Ti667</b>	<b>Ti666</b>	<b>FIC606</b>	<b>Ti654</b>	<b>FIC622</b>	<b>FI607</b>	<b>Ti649</b>	<b>PIC643</b>	<b>FIC605</b>	<b>PI644</b>	<b>TIC663</b>
1,000	0,181	0,239	0,015	0,513	0,211	-0,884	0,085	0,197	-0,290	-0,582	-0,835	-0,086
0,181	1,000	-0,558	0,129	-0,134	0,316	-0,121	0,365	0,323	-0,075	0,422	0,012	0,474
0,239	-0,558	1,000	-0,093	0,472	-0,429	-0,085	-0,221	-0,656	0,097	-0,609	-0,223	-0,798
0,015	0,129	-0,093	1,000	0,055	0,023	0,030	0,541	0,050	0,045	0,009	0,010	0,195
0,513	-0,134	0,472	0,055	1,000	0,479	-0,531	-0,300	-0,185	0,254	-0,411	-0,431	-0,611
0,211	0,316	-0,429	0,023	0,479	1,000	-0,317	-0,341	0,307	-0,030	0,193	-0,085	-0,036
-0,884	-0,121	-0,085	0,030	-0,531	-0,317	1,000	-0,092	-0,338	0,235	0,599	0,925	-0,028
0,085	0,365	-0,221	0,541	-0,300	-0,341	-0,092	1,000	0,074	0,184	-0,028	-0,050	0,488
0,197	0,323	-0,656	0,050	-0,185	0,307	-0,338	0,074	1,000	-0,394	0,266	-0,347	0,796
-0,290	-0,075	0,097	0,045	0,254	-0,030	0,235	0,184	-0,394	1,000	0,139	0,394	-0,243
-0,582	0,422	-0,609	0,009	-0,411	0,193	0,599	-0,028	0,266	0,139	1,000	0,662	0,441
-0,835	0,012	-0,223	0,010	-0,431	-0,085	0,925	-0,050	-0,347	0,394	0,662	1,000	-0,031
-0,086	0,474	-0,798	0,195	-0,611	-0,036	-0,028	0,488	0,796	-0,243	0,441	-0,031	1,000



**Таблица 5.** Входные переменные, упорядоченные по убыванию коэффициентов парной корреляции.

Наименование датчика РК	Тв (темпл.верху РК) град. С	Твр (темпл.верха вых.продукта) С	Твр (темпл.входа прол.) град. С	Спк (поток сырья на входа в РК) М <sup>3</sup> /ч	Твых (темпл.выхода про-дукта) С	От (кол-во производ. прол.) Т/ч	Спв (скор. выхода про-дукта) М <sup>3</sup> /ч	Твс (темпл. сырья на входе в РК) град. С	Рв (давление (давлениеверху РК) кг/см <sup>2</sup> )	Свк (поток сырьяверху РК) М <sup>3</sup> /ч	Рн (давление внизу РК) кг/см <sup>2</sup> )	Тн (темпл.внизу РК) град. С
	TIС653	TI667	TI666	FIС606	TI654	FIС622	FI607	TI649	PIС643	FIС605	PI644	TIС663
Знач. датч.	Х <sub>6</sub>	Х <sub>11</sub>	Х <sub>10</sub>	Х <sub>4</sub>	Х <sub>9</sub>	Х <sub>2</sub>	Х <sub>5</sub>	Х <sub>8</sub>	Х <sub>1</sub>	Х <sub>12</sub>	Х <sub>7</sub>	Х <sub>3</sub>
	88	77	73	72	194	30	22	85	2	46	3	120

Использование коэффициентов парной корреляции не всегда позволяет построить модель с высоким значением  $R^2$  на обобщающей выборке. Это происходит в связи с тем, что наиболее важные с точки зрения физического смысла входы, такие как давление в колонне и температура верха, имеют малый разброс относительно номинальных значений на обучающей выборке. Поэтому практически невозможно выявить наличие статистических связей на уровне коэффициентов парной корреляции, но в то же время небольшие отклонения данных входов от значений, не учтенных в обучающей выборке, негативно влияют на невязку модели ВА.

В таблицах 6, 7 представлены коэффициенты детерминации моделей, построенных на разном количестве входов от 1–12, сформированных с точки зрения возрастания коэффициентов парной корреляции для моделей ВА по доли бензолобразующих компонентов и по доли изо-пентана в верхнем продукте промышленной РК.

Кроме этого, в ходе исследования было перебрано достаточно большое количество комбинаций входов с целью учета физического смысла переменных в моделях ВА.

**Таблица 6.** Коэффициенты детерминации моделей, построенных на разном количестве входов от 1–12, сформированных по возрастанию коэффициентов парной корреляции для моделей качества промышленной РК по доли бензолобразующих компонентов.

метод	Количество входов для модели ВА по содержанию бензолобразующих											
	1	2	3	4	5	6	7	8	9	10	11	12
ПЛС												
R <sup>2</sup> <sub>обуч.</sub>	0,6026	0,7694	0,8029	0,8574	0,8616	0,8623	0,8626	0,8628	0,955	0,9557	0,9578	0,9619
R <sup>2</sup> <sub>обоб.</sub>	0,6432	0,678	0,8665	0,7781	0,8078	0,8267	0,8389	0,8324	0,8023	0,8308	0,802	0,709
МНК												
R <sup>2</sup> <sub>обуч.</sub>	0,6026	0,7694	0,8029	0,8574	0,8616	0,8623	0,8626	0,8628	0,955	0,9557	0,9578	0,9619
R <sup>2</sup> <sub>обоб.</sub>	0,6432	0,678	0,8665	0,7781	0,8078	0,8267	0,8389	0,8324	0,8023	0,8308	0,802	0,709
MLR												
R <sup>2</sup> <sub>обуч.</sub>	0,252	0,2521	0,298	0,3588	0,6126	0,6195	0,6798	0,7108	0,954	0,954	0,9565	0,9565
R <sup>2</sup> <sub>обоб.</sub>	0,1957	0,1909	-0,171	0,2088	0,5845	0,7218	0,6307	0,6796	0,8351	0,8298	0,8043	0,8046
MLR												
R <sup>2</sup> <sub>обуч.</sub>	0,252	0,2521	0,298	0,3588	0,6126	0,6195	0,6798	0,7108	0,954	0,954	0,9565	0,9565
R <sup>2</sup> <sub>обоб.</sub>	0,1957	0,1909	-0,171	0,2088	0,5845	0,7218	0,6307	0,6796	0,8351	0,8298	0,8043	0,8046
PP												
R <sup>2</sup> <sub>обуч.</sub>	0,5981	0,7573	0,8021	0,8571	0,8523	0,8038	0,8607	0,7225	0,9549	0,9556	0,9577	0,9619
R <sup>2</sup> <sub>обоб.</sub>	0,6333	0,6712	0,8571	0,7809	0,8252	0,7058	0,8331	0,7893	0,7977	0,8315	0,803	0,7014
ACE												
R <sup>2</sup> <sub>24</sub>	0,4015	0,8879	0,901	0,8746	0,8228	0,9085	0,9384	0,8612	0,8219	0,889	0,9655	0,9682
R <sup>2</sup> <sub>18</sub>	0,809	33,638	0,9261	0,9032	0,968	0,9415	0,9388	0,9184	0,9283	0,912	0,9269	0,9325

**Таблица 7.** Коэффициенты детерминации моделей построенных на разном количестве входов от 1-12, сформированных с точки зрения возрастания коэффициентов парной корреляции для моделей качества промышленной РК по доли изо-пентана.

Метод	Количество входов для модели ВА по доле изо-пентана											
	1	2	3	4	5	6	7	8	9	10	11	12
ПЛС												
R <sup>2</sup> _обуч.	0,7742	0,7792	0,7887	0,7887	0,8062	0,8367	0,8485	0,8682	0,9202	0,9202	0,9213	0,9547
R <sup>2</sup> _обоб.	0,7196	0,7171	0,4574	0,457	0,4274	0,7404	0,4983	-0,3922	-1,8584	-2,017	-1,9979	-3,8533
МНК												
R <sup>2</sup> _обуч.	0,7742	0,7792	0,7887	0,7887	0,8062	0,8367	0,8485	0,8682	0,9202	0,9202	0,9213	0,9547
R <sup>2</sup> _обоб.	0,7196	0,7171	0,4574	0,457	0,4274	0,7404	0,4983	-0,3922	-1,8584	-2,017	-1,9979	-3,8533
MLR												
R <sup>2</sup> _обуч.	-0,1826	0,1802	0,2799	0,4661	0,5358	0,5491	0,6899	0,7797	0,9197	0,9201	0,9212	0,9263
R <sup>2</sup> _обоб.	-0,556	0,1915	-0,7085	0,3466	0,264	-0,0788	-0,2738	-2,2201	-1,6222	-2,1086	-2,0472	-2,5614
MLR												
R <sup>2</sup> _обуч.	-0,1826	0,1802	0,2799	0,4661	0,5358	0,5491	0,6899	0,7797	0,9197	0,9201	0,9212	0,9263
R <sup>2</sup> _обоб.	-0,556	0,1915	-0,7085	0,3466	0,264	-0,0788	-0,2738	-2,2201	-1,6222	-2,1086	-2,0472	-2,5614
PP												
R <sup>2</sup> _обуч.	0,7742	0,7792	0,7887	0,7886	0,8061	0,8365	0,8484	0,8676	0,8841	0,9103	0,9194	0,9545
R <sup>2</sup> _обоб.	0,7231	0,7204	0,4312	0,4291	0,3906	0,7495	0,505	-0,3803	-0,9917	-1,8124	-2,1996	-3,6761
ACE												
R <sup>2</sup> _24	0,6717	0,4431	-16,115	0,8075	0,6971	0,9304	0,9813	0,9566	0,9727	0,967	0,951	0,969
R <sup>2</sup> _18	0,7311	0,7863	0,8444	0,8908	0,9261	0,9187	0,9164	0,9129	0,9355	0,9357	0,9464	0,948

Для каждой модели ВА выполнялась проверка на адекватность по критерию Фишера [10]. Результаты проверки статистической значимости параметров моделей ВА (уравнений регрессии) по критерию Стьюдента [10] приведены в таблице 8.

**Таблица 8.** Значимость параметров моделей ВА.

Параметр	b <sub>0</sub>	b <sub>1</sub>	b <sub>2</sub>	b <sub>3</sub>	b <sub>4</sub>	b <sub>5</sub>	b <sub>6</sub>
	Для модели ВА изо-пентана (входы выбирались по физич. смыслу)						
Значимость	Не знач.	Значим	Значим	Значим	Значим	Значим	Значим
Значения	1,271<1,74	6,612>1,74	3,812>1,74	1,891>1,74	3,005>1,74	2,890>1,74	7,544>1,74
	Для модели ВА бензолообразующих компонентов (входы выбирались по физич. смыслу)						
Значимость	Не знач.	Значим	Не знач.	Не знач.	Не знач.	Не знач.	Значим
Значения	0,902>1,74	4,710>1,74	0,288<1,74	1,176<1,74	0,213<1,74	0,079<1,74	7,953>1,74
	Для модели ВА изо-пентана (входы выбирались по коэфф. парной корреляции)						
Значимость	Значим	Значим	Не знач.	Не знач.	Не знач.	Не знач.	Значим
Значения	33,647>1,74	3,847>1,74	1,209<1,74	0,593<1,74	1,248<1,74	1,570<1,74	1,976>1,74
	Для модели ВА бензолообразующих (входы выбирались по коэфф. парной корреляции)						
Значимость	Значим	Значим	Не знач.	Не знач.	Не знач.	Не знач.	Не знач.
Значения	5,271>1,74	5,316>1,74	1,617<1,74	1,471<1,74	1,724<1,74	0,707<1,74	0,295<1,74

Как показали расчеты, коэффициенты детерминации моделей регрессии при исключении незначимых параметров уменьшались, но лишь незначительно. В модели ВА по содержанию бензолообразующих компонентов, в случае выбора входных переменных с точки зрения физического смысла, количество значимых входов сокращается до двух. Для модели ВА по доле изо-пентана количество входов (оказывающих влияние на выход с точки зрения физического смысла) уменьшается с 12 до 6, а в моделях, постро-

енных на основе выбора входов по коэффициентам парной корреляции до двух (таблица 9).

**Таблица 9.** Количество входов в модели и коэффициенты детерминации.

<b>Входы выбирались на основе физического смысла</b>		
<b>МНК</b>	<b>Бензолообразующих</b>	<b>Изо-пентан</b>
<b>Кол-во входов</b>	2	6
<b>R<sup>2</sup>_обуч.</b>	0.8663	0.8746
<b>R<sup>2</sup>_обоб.</b>	0.8980	0.6433
<b>Входы выбирались по коэффициентам парной корреляции</b>		
<b>МНК</b>	<b>Бензолообразующих</b>	<b>Изо-пентан</b>
<b>Кол-во входов</b>	2	2
<b>R<sup>2</sup>_обуч.</b>	0.7396	0.7927
<b>R<sup>2</sup>_обоб.</b>	0.8164	0.7995

## 6. Заключение

Предложенная методика определения оптимального количества входных переменных моделей ВА опробована на примере промышленной ректификационной колонны, работающей в режиме, близком к состоянию захлебывания. Проведенные вычислительные эксперименты позволяют утверждать, что при выборе оптимального количества входных переменных целесообразно применять различные методы построения регрессионных моделей ВА, опираясь также на информативность входных переменных с точки зрения физического смысла.

## Список литературы

1. Бахтадзе Н.Н. Виртуальные анализаторы (идентификационный подход) // АиТ. 2004. № 11. С. 3-23.
2. Lee C., Choi S.W., Lee, I-B. Sensor fault identification based on time-lagged PCA in dynamic processes // Chemometrics and Intelligent Laboratory Systems. 2004. Vol. 70, No. 2. P. 165-178.
3. Street J. O., R. J. Carroll, Ruppert D. A Note on Computing Robust Regression Estimates via Iteratively Reweighted Least Squares // The American Statistician. 1988. Vol.42. P. 152-154.
4. Zamprogna E., Barolo M., Seborg D. E. Development of a soft sensor for a bath distillation column using liner and nonlinear PLS regression techniques // Control Engineering Practice. 2004. Vol. 12, No. 7. P. 917-929.
5. Mejdell T., Skogestad S. Estimation of Distillation Compositions from Multiple Temperature Measurements Using Partial least squares Regression // Ind. Eng. Chem. Res. 1991. Vol. 30. P. 2543-2555.
6. Breiman L., Friedman J. Estimating optional transformations for multiple regression and correlation // Journal of the American Statistical Association. 1985. Vol. 80. P. 580-598.
7. Holland C. D. Fundamentals of multicomponent distillation. New York: McGraw-Hill Book Company, 1981. 633P.
8. Диго Г.Б., Диго Н.Б., Можаровский И.С., Торгашов А.Ю. Исследование методов идентификации моделей виртуальных анализаторов показателей качества ректификационной колонны // Всероссийская научная конференция «Фундаментальные и прикладные вопросы механики и процессов управления». Сборник докладов. Владивосток: ИАПУ ДВО РАН, 2011.
9. Wang D., Murphy M. Estimating optimal transformations for multiple regression using the ACE algorithm // Journal of Data Science. 2004. Vol. 2. P. 329-346.
10. Вентцель Е.С. Теория вероятностей. М.: Физматгиз, 1962. 560 с.