

# ИССЛЕДОВАНИЕ МЕТОДОВ ИДЕНТИФИКАЦИИ МОДЕЛЕЙ ВИРТУАЛЬНЫХ АНАЛИЗАТОРОВ ПОКАЗАТЕЛЕЙ КАЧЕСТВА РЕКТИФИКАЦИОННОЙ КОЛОННЫ

Г.Б. Диго

*Институт автоматизации и процессов управления ДВО РАН*

Россия, 690041, Владивосток, Радио 5

E-mail: [digo@iacp.dvo.ru](mailto:digo@iacp.dvo.ru)

Н.Б. Диго

*Институт автоматизации и процессов управления ДВО РАН*

Россия, 690041, Владивосток, Радио 5

E-mail: [bernatsk@iacp.dvo.ru](mailto:bernatsk@iacp.dvo.ru)

И.С. Можаровский

*Институт автоматизации и процессов управления ДВО РАН*

Россия, 690041, Владивосток, Радио 5

E-mail: [studvvsu@gmail.com](mailto:studvvsu@gmail.com)

А.Ю. Торгашов

*Институт автоматизации и процессов управления ДВО РАН*

Россия, 690041, Владивосток, Радио 5

E-mail: [torgashov@iacp.dvo.ru](mailto:torgashov@iacp.dvo.ru)

**Ключевые слова:** регрессионная модель, идентификация, математическое ожидание, виртуальный анализатор, ректификационная колонна

Излагается подход к построению виртуальных анализаторов качества продуктов промышленной РК на основе регрессионных моделей и алгоритма чередующихся условных математических ожиданий (АСЕ). Показывается, что полученные результаты не противоречат существующим методикам выбора регрессионных методов моделирования в области хемометрики и что алгоритм АСЕ следует применять, когда традиционные регрессионные методы дают плохую точность модели и не применять его, если результаты моделирования на порядок хуже традиционных регрессионных методов.

## Введение

В настоящее время при оценке качества выходных продуктов нефтеперерабатывающих и химических процессов и, в частности, ректификационных

колонн (РК), используются виртуальные анализаторы (ВА) [1]. Они предназначены для оценки не измеряемого непосредственно, но необходимого показателя качества продукта по измеряемым параметрам технологического процесса (температура, давление, расход), непрерывно контролируемым современными системами управления.

Принцип действия ВА основан на непрерывном определении показателя качества по математической модели, описывающей его взаимосвязь с текущими значениями измеряемых характеристик режима технологического процесса. Использование ВА обусловлено тем, что системы лабораторного контроля качества на предприятии не в состоянии выдавать своевременную непрерывную информацию о качестве продукции из-за длительных и трудоемких отбора, транспортировки и анализа производимых проб (иногда они могут осуществляться не более двух раз в смену). Это может привести либо к выпуску брака, либо к ведению процесса с большим запасом по качеству. Однако оба эти варианта существенно снижают эффективность производства.

Использование в этих целях автоматических анализаторов с прямым замером показателя качества в потоке не получило должного развития из-за своей сложности в эксплуатации и высокой стоимости. В отличие от них виртуальные анализаторы, практически не уступая в точности, значительно дешевле и надежнее.

В докладе предлагается методика построения ВА качества продуктов промышленной РК с применением регрессионных моделей и алгоритма чередующегося условного математического ожидания (АСЕ) [2].

## 1. Постановка задачи

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

Для построения ВА использовались метод наименьших квадратов (МНК) [3], робастная регрессия (РР) [4], метод проекции на латентные структуры (ПЛС) [5, 6]. Перечисленные выше методы применялись для построения линейных моделей ВА, а алгоритм АСЕ - для построения нелинейных моделей.

Ставится задача разработки метода построения ВА ректификационной колонны с помощью алгоритма АСЕ и анализа качества моделей, получаемых на основе нескольких регрессионных методов с последующим выбором наиболее адекватной из них.

## 2. Описание промышленной РК

Процесс массообмена в РК показан на рис. 1. Средние значения технологических параметров в определенных точках РК, на основе которых будут создаваться ВА, приведены ниже в табл. 1.

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

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

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

$K_{ij}$  - 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 газобразной фазе на  $j$ -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  $N=44, (j=1, 2, \dots, N).$

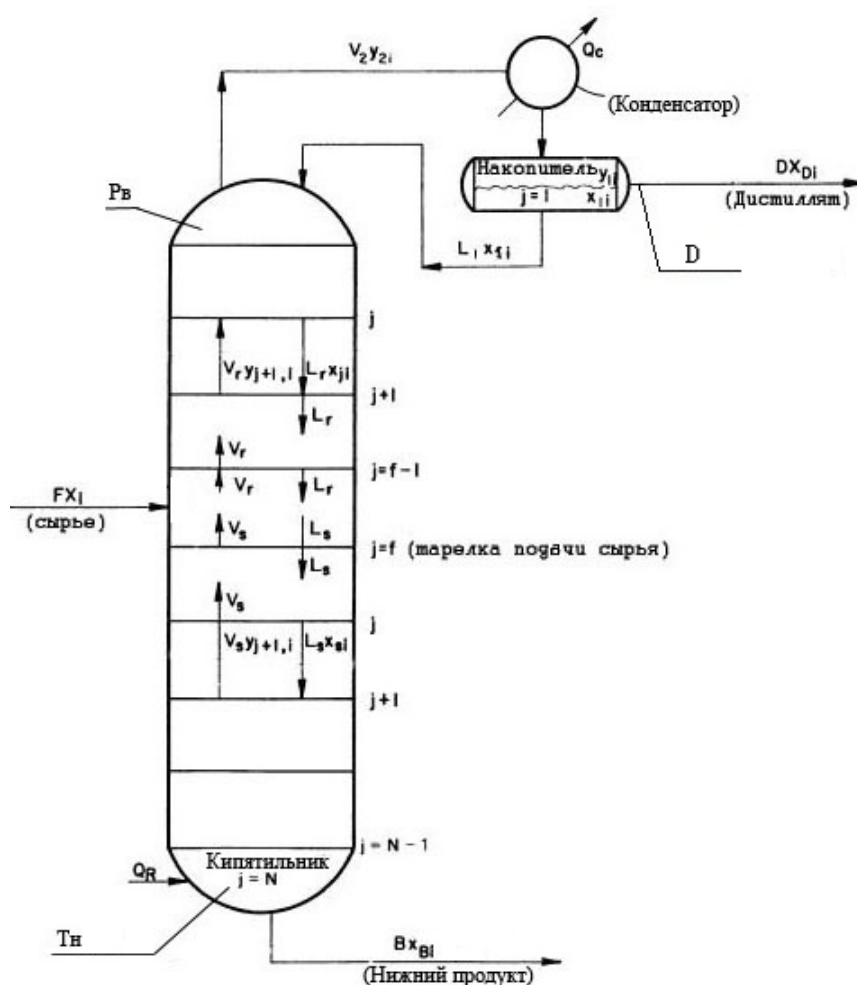


Рис. 1. Схематическое изображение РК

Таблица 1 Средние значения технологических параметров в заданных точках ректификационной колонны

Наименование датчика РК	Р <sub>в</sub> - давление вверху РК, кг/см <sup>3</sup>	D – расход дистиллята, т/ч	T <sub>н</sub> - температура внизу РК, °C
	<b>PIC643</b>	<b>FIC622</b>	<b>TIC663</b>
	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>
Показатель датчика	2	30	120

Уравнения материального баланса имеют вид:

$$\begin{aligned}
 V_{j+1}y_{j+1,i} &= L_jx_{ji} + DX_{Di} & (j = 1, 2, \dots, f - 2) \\
 V_fy_{fi} + V_Fy_{Fi} &= L_{f-1}x_{f-1,i} + DX_{Di} \\
 V_{j+1}y_{j+1,i} &= L_jx_{ji} - Bx_{Bi} & (j = f, f + 1, \dots, N - 1) \\
 FX_i &= DX_{Di} + Bx_{Bi}
 \end{aligned} \tag{2}$$

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  $f$ ,  
 $f$  - номер тарелки в РК, на которую подается сырье,  
 $DX_i$  - концентрация  $i$ -го компонента в дистилляте.

Уравнения энергетического баланса имеют вид:

$$\begin{aligned}
 V_{j+1}H_{j+1} &= L_jh_j + DH_D + Q_c & (j = 1, 2, \dots, f - 2) \\
 V_fH_f + V_FH_F &= L_{f-1}h_{f-1} + DH_D + Q_c \\
 V_{j+1}H_{j+1} &= L_jh_j - Bh_B + Q_R & (j = f, f + 1, \dots, N - 1) \\
 FH &= Bh_B + DH_D + Q_c - Q_R
 \end{aligned} \tag{3}$$

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 кипятивильника.

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 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.

### 3. Алгоритм АСЕ

Модель, получаемая по алгоритму АСЕ [2, 8], имеет следующую общую форму:

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

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 АСЕ 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

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

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

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

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

## 4. Исследование методов построения моделей виртуального анализатора

Строятся модели ВА, для построения которых необходимо сформировать данные из базы данных технологических переменных РК. Для моделирования они выбраны с учетом физико-химических процессов, влияющих на качество продуктов в РК (табл. 1). Вычисляются средние значения переменных за последний час перед лабораторным измерением. Это связано с тем, что ВА производит оценку показателей качества РК в статическом режиме. Полученные средние значения (матрица  $X$ ) сопоставляются с лабораторными измерениями качества продукта колонны (вектор  $Y$ ).

Для идентификации моделей ВА (содержание бензолобразующих веществ и концентрация изо-пентана в дистилляте) использованы 4 регрессионных метода и алгоритм АСЕ, 2 набора лабораторных данных (значения  $Y$ ), 3 технологических переменных (значения  $X$ ). Результаты представлены в табл. 2.

Таблица 2 Результаты идентификации модели ВА по содержанию бензолобразующих веществ в дистилляте

Используемый метод	Коэффициент мн. корреляции R	Среднеквадратичное отклонение	Коэффициент детерминации R <sup>2</sup>
МНК	0.6069	1.7067	0.3683
МЛР	0.4970	1.8636	0.2468
ПЛС 2	0.4970	1.8636	0.2468
РР	0.6067	1.7070	0.3679
ПЛС	0.6069	1.7067	0.3683
АСЕ	<b>0.8694</b>	<b>1.0735</b>	<b>0.7500</b>

При использовании в качестве критерия коэффициента детерминации  $R^2$  оказалось, что наиболее точной является модель, полученная по алгоритму АСЕ. Это подтверждает тот факт, что алгоритм АСЕ наиболее подходит для идентификации ВА качества выходных продуктов РК с высокой долей нелинейности.

Результаты построения моделей разными регрессионными методами по концентрации изо-пентана в выходном продукте РК представлены в табл. 3.

Таблица 3 Результаты идентификации модели ВА по концентрации изо-пентана в дистилляте

Используемый метод	Коэффициент мн. корреляции R	Среднеквадратичное отклонение	Коэффициент детерминации R <sup>2</sup>
МНК	0.9308	0.7851	0.8663
МЛР	0.9083	0.9009	0.8240
ПЛС 2	0.9083	0.9009	0,8240
РР	0.9307	0.7859	0.8661
ПЛС	0.9308	0.7851	0.8663
АСЕ	<b>0.8434</b>	<b>1.2622</b>	<b>0.4920</b>

Из табл. 3 следует, что при построении ВА по концентрации изо-пентана наилучшие результаты дают традиционные методы регрессионного модели-

рования ПЛС, МНК.

## 5. Заключение

Проведенные исследования промышленных данных ректификационной колонны регрессионными методами с целью идентификации ВА показали, что в случае явно выраженной нелинейности алгоритм АСЕ дает более точные результаты, чем регрессионные методы МНК, МЛР, ПЛС, ПЛС2, РР.

В процессе идентификации ВА по содержанию изо-пентана в дистилляте алгоритм АСЕ дал результат почти в два раза хуже, по сравнению с традиционными регрессионными методами. Поэтому его следует применять, когда традиционные регрессионные методы дают плохую точность модели из-за наличия существенно нелинейных взаимосвязей в структуре моделируемых данных и отказаться от его применения, если результаты моделирования оказываются на порядок хуже традиционных регрессионных методов. В связи с этим при создании ВА для промышленных ректификационных колонн представляется целесообразным для проверки описанных условий применять многометодные технологии вместо использования контроля разработчика.

Исследования показали, что из традиционных регрессионных методов моделирования наиболее подходящим, точным и удобным для интерпретации результатов является метод проекции на латентные структуры (ПЛС). Полученные результаты не противоречат существующим методикам выбора регрессионных методов моделирования в области хемометрики.

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".

## Список литературы

1. Бахтадзе Н.Н. Виртуальные анализаторы (идентификационный подход) // АиТ. 2004. № 11. С. 3-23.
2. Breiman L., Friedman J. Estimating optional transformations for multiple regression and correlation // Journal of the American Statistical Association. 1985. Vol. 80. P. 580-598.
3. 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.
4. 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.
5. 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.
6. 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.
7. Holland, C. D. Fundamentals of multicomponent distillation. New York: McGraw-Hill Book Company, 1981. 633P.
8. Wang D., Murphy M. Estimating optimal transformations for multiple regression using the ACE algorithm // Journal of Data Science. 2004. Vol. 2. P. 329-346.