

хастических аппроксимаций (14), подробно изложенных в [3].

Основной итог статьи состоит в моделировании процессов «трубчатой» структуры, которая имеет место всегда, если компоненты вектора входных переменных исследуемого процесса стохастически зависимы. В этом случае традиционно используемые модели статических систем с запаздыванием неприменимы или, в лучшем случае, могут приводить к значительным ошибкам. Наиболее интересным является тот факт, что мы приходим к необходимости введения пространства дробной размерности. Интересен вопрос: будет ли это пространство Хаусдорфа–Безиковича? Проведенные численные исследования подтверждают эффективность использования Н-моделей вместо общепринятых.

A. V. Medvedev

H-MODELS FOR NON-INERTIA SYSTEMS WITH A DELAY

The author considers a problem of the new class of processes with a «tubular» structure in the space of “input-output” variables. Modeling of these class processes differs highly from the generally accepted parametric models being a surface of the same space. A question of multidimensional systems modeling having small volume of learning samples is analyzed specially. To develop the trainable parametric models of «tubular» processes, the corresponding non-parametric indicator is introduced. A new class of trainable parametric models, and some results of the numerical investigation are presented.

Keywords: identification, non parametric algorithms, parametric model, a priori information, stochastic processes, H-models.

© Медведев А. В. , 2012

УДК 669:51-7

Т. В. Пискажова, П. В. Поляков, Н. А. Шарыпов

ИСПОЛЬЗОВАНИЕ МЕТОДОВ МНОГОМЕРНОЙ ОБРАБОТКИ ДАННЫХ ДЛЯ АНАЛИЗА РАСХОДА ЭЛЕКТРОЭНЕРГИИ АЛЮМИНИЕВЫХ ЭЛЕКТРОЛИЗЕРОВ

Выполнен обзор применения методов многомерной статистической обработки базы данных алюминиевого производства для оценки эффективности работы аппаратов и качества получаемой продукции. Рассмотрено применение метода главных компонент для анализа расхода электроэнергии двух корпусов алюминиевого завода.

Ключевые слова: многомерный статистический анализ, метод главных компонент, алюминиевый электролизер, расход электроэнергии.

Аппарат для получения алюминия электролизом – алюминиевый электролизер – сложная диссипативная система. Работа аппарата оценивается технико-экономическими показателями (ТЭП). Входными параметрами для электролизера (рис. 1) являются качественные характеристики сырья (пек, кокс, анодная масса, глинозем, фтористые соли).

Возмущающие воздействия на электролизер вызывают колебания тока серии, загрузка анодной массы, перестановка штырей, перетяжка анодной рамы, технологические обработки, выборка металла. Пара-

Библиографические ссылки

1. Эйххофф П. Основы идентификации систем управления. М. : Мир, 1975.
2. Цыпкин Я. З. Основы информационной теории идентификации. М. : Наука, 1984.
3. Медведев А. В. Теория непараметрических систем. Моделирование // Вестник СибГАУ. 2010. Вып. 4. С. 4–9.
4. Медведев А. В. Анализ данных в задаче идентификации // Компьютер. анализ данных моделирования : сб. науч. ст. Т. 2. Минск : Изд-во БГУ, 1995. С. 201–206.
5. Медведев А. В. Непараметрические системы адаптации. Новосибирск : Наука, 1983.
6. Мондельброт Б. Фрактальная геометрия природы. М. : Ижевск : Ижев. ин-т компьютер. исслед. ; НИЦ «Регулярная и хаотическая динамика», 2010.

метры процесса используются для технологического управления. В автоматическом режиме осуществляется управление напряжением электролизера, концентрацией глинозема, криолитовым отношением.

Анализ ТЭП алюминиевого электролизера ведется по факту произведенного металла (удельные расходные коэффициенты).

Прогноз и управление ТЭПами на основании существующих знаний об электролизе и информации, имеющейся в базе данных алюминиевого завода, представляется важной задачей.



Рис. 1. Электролизер как объект управления

Существенно важным в современных экономических условиях для алюминиевых заводов является расход электроэнергии (до 30 % в себестоимости получаемого первичного алюминия). В литературных источниках [1–3] имеются отдельные сведения о зависимости расхода электроэнергии от некоторых параметров сырья и работы аппарата. Для повышения эффективности работы электролизера и снижения себестоимости металла важно знать одновременное влияние параметров процесса на расход электроэнергии. Наиболее простым методом построения и исследования подобных зависимостей по статистическим выборкам является множественный регрессионный анализ [4; 5]. Но в случае большого количества взаимозависимых (мультиколлинеарных) исходных параметров, что и наблюдается в такой сложной диссипативной системе, как алюминиевый электролизер, проявляется «эффект деградации» (неустойчивости) регрессионных коэффициентов, обнаруженный Фришом [6]. Для исключения мультиколлинеарности переменных при построении множественной регрессии существует несколько различных методов, например исключение взаимозависимых переменных из выборки, использование различных вариаций МНК. На наш взгляд, наиболее верным и математически обоснованным является метод главных компонент (МГК) [4]. МГК широко применяется для решения задач классификации, снижения размерности исходного массива, выделения новых латентных переменных в структуре данных, распознавания образов.

Блок-схема статистического анализа данных с использованием МГК представлена на рис. 2.

В алюминиевой промышленности использовались следующие возможности этого метода:

- выявление нескольких главных переменных из большого количества анализируемых переменных, повлиявших на ход процесса за рассматриваемое время [7];
- определение характера зависимости (прямо пропорциональная или обратно пропорциональная) между ТЭПами и выделенными главными переменными [8];
- диагностика нарушений или изменения ТЭПов [7; 9; 10];
- графический анализ для оценки концентрации веществ в определенной зоне [11];

- оценка производительности аппарата [11];
- онлайн-мониторинг состояния аппарата [11].

В данной работе показано применение МГК в качестве инструмента для многомерного анализа технологических параметров алюминиевых электролизеров.

Выборка данных. Выборка для анализа расхода электроэнергии содержала восемь переменных (табл. 1) и 46 наблюдений (данные с января 2007 г. по октябрь 2010 г.). Анализ проводился для корпусов, работающих с технологией бокового токоподвода, корпус 1 и корпус 2 Надвоицкого алюминиевого завода.

Целесообразность использования среднемесячных параметров в статистическом анализе исходит из технологических соображений и обосновывается подробно в работе Тесье с соавторами [12]. Выбор переменных для анализа осуществлялся на основе работ [1–3].

Анализ. Многомерный статистический анализ проводился с использованием ПП Statistica и MS Excel. Объяснение дисперсии для выборки для расхода энергии представлено в табл. 2. В этом случае для объяснения всей дисперсии требуется всего восемь ГК, тогда как четыре ГК объясняют уже $\approx 90\%$ исходных данных.

Вклады параметров в первые пять ГК представлены в табл. 3. Обычно рассматривают вклад в первые несколько ГК, так как они описывают значительную часть изменения в исходных параметрах (для данных выборок две ГК описывают больше половины дисперсии данных).

Первое направление (ГК1) для обоих корпусов характеризуется тем, что в него внесли вклад следующие параметры: сила тока, срок службы, температура электролита и уровень металла (в первом корпусе дополнительно вклад внесли изменения, связанные с выходом по току).

Второе направление (ГК2) характеризуется следующими параметрами, которые внесли большой вклад: удельный выход пены и криолитовое отношение. В первом корпусе дополнительный вклад имеет ЧАЭ, во втором корпусе значительный вклад имеет выход по току. Отметим, что независимо от корпуса, вклад в первые две компоненты внесли одни и те же переменные.

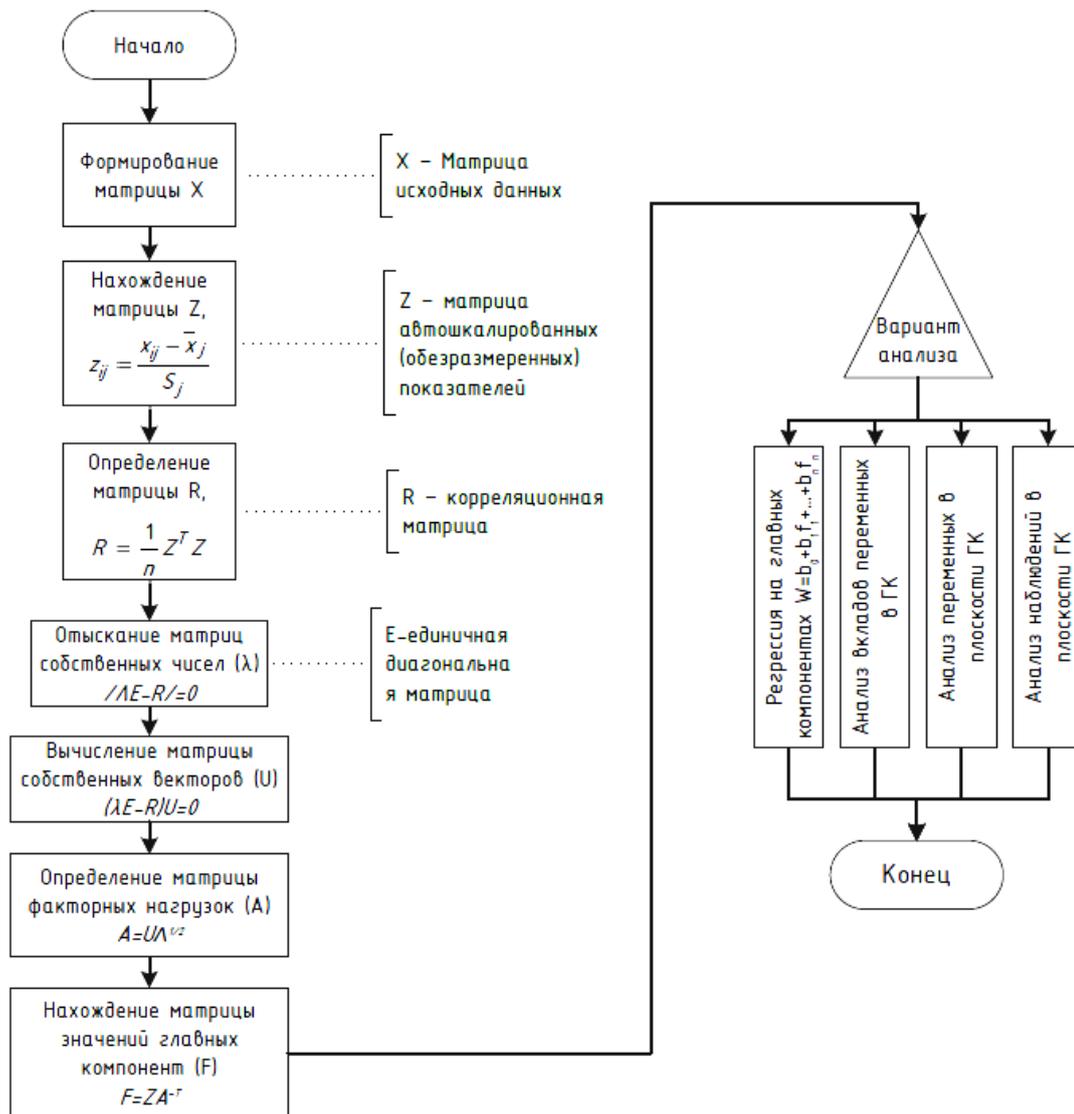


Рис. 2. Алгоритм анализа данных с применением МГК

Таблица 1

Параметры, включенные в выборку для анализа расхода электроэнергии

Параметр	Условное обозначение	Единицы измерения	Корпус 1		Корпус 2	
			Среднее значение	Стандартное отклонение	Среднее значение	Стандартное отклонение
Сила тока	CS	кА	81,38	1,15	81,38	1,15
Выход по току (ВПТ)	CE	%	90,07	1,70	90,07	1,76
Частота анодных эффектов (ЧАЭ)	AEF	шт./сут.	0,68	0,25	0,68	0,26
Срок службы	PL	мес.	41,75	9,67	43,64	10,03
Удельный выход угольной пены	Dust	кг/т Al	10,56	4,44	10,85	4,56
Температура электролита	BT	°C	960,82	3,19	961,34	3,27
Криолитовое отношение (КО)	CR	доли ед.	2,68	0,07	2,69	0,07
Уровень металла	ML	см	34,82	3,38	34,72	3,31
Расход электроэнергии*	EC	кВт·ч/т	15165,10	308,39	15 228,89	349,93

* Вспомогательный параметр не участвует в расчете коэффициентов при переходе к главным компонентам (ГК), но отображается в плоскости ГК для анализа характера зависимостей.

Таблица 2

Объяснение дисперсии исходных данных главными компонентами

Номера ГК	Корпус 1		Корпус 2	
	Объясненный процент общей дисперсии	Накопительный объясненный процент	Объясненный процент общей дисперсии	Накопительный объясненный процент
1	47,499	47,499	48,810	48,810
2	15,662	63,161	18,192	67,002
3	14,856	78,017	14,321	81,323
4	11,334	89,351	8,523	89,846
5	5,084	94,435	5,074	94,920
6	3,870	98,305	3,076	97,996
7	1,246	99,552	1,620	99,616
8	0,448	100	0,384	100

Таблица 3

Вклад переменных в первые две ГК

Параметр	Корпус 1		Корпус 2	
	ГК1	ГК2	ГК1	ГК2
Сила тока	0,198	0,006	0,184	0,000
ВПГ	0,102	0,011	0,099	0,138
ЧАЭ	0,015	0,330	0,008	0,001
Срок службы	0,206	0,008	0,210	0,003
Выход пены	0,002	0,203	0,001	0,529
Температура электролита	0,178	0,023	0,179	0,008
КО	0,062	0,416	0,076	0,315
Уровень металла	0,237	0,003	0,242	0,005

В ходе анализа получены проекции переменных в плоскостях ГК (рис. 3, 4). Плоскости ГК – это срезы (проекции) многомерного пространства. На основании этих проекций делают вывод о характере зависимости между анализируемыми переменными. По правилам метода [8], прямо пропорциональную зависимость между собой имеют переменные, находящиеся в одном квадранте, обратно пропорциональную зависимость имеют переменные, находящиеся в противоположных квадрантах. Характер зависимости между переменными, находящимися в соседних квадрантах, не указывается (в геометрической интерпретации они независимы, так как в многомерном пространстве угол между ними приближается к прямому).

Анализ проекций переменных в ГК (см. рис. 3, 4) показывает, что для обоих корпусов прямо пропорциональную зависимость между собой имеют следующие переменные: выход пены, срок службы, расход электроэнергии; ЧАЭ, сила тока и уровень металла. Обратно пропорциональную зависимость имеют следующие переменные: срок службы, выход по току; выход пены, выход по току; расход электроэнергии и выход по току. Такой характер зависимости подтверждается в литературе. Например, чем больше пены в электролизере, тем выше температура электролита и, как следствие, больше электроэнергии необходимо затратить на производство алюминия; с увеличением

выхода по току снижается расход электроэнергии, оба эти параметра входят в уравнение для расчета выхода по току и имеют обратно пропорциональную зависимость.

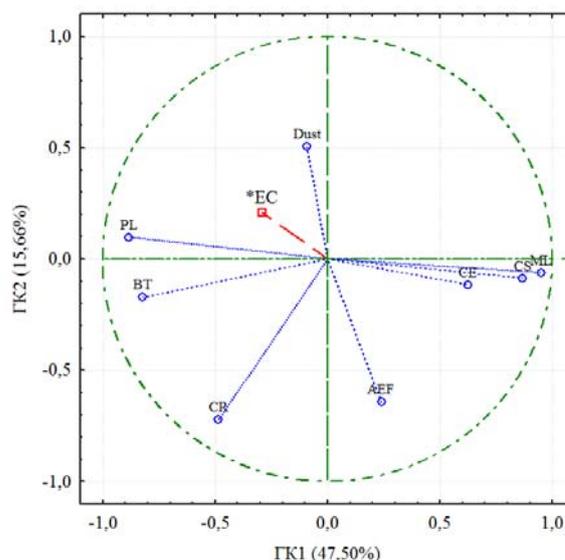


Рис. 3. Проекция переменных в плоскости ГК1 и ГК2 для КЭБТ № 1: □ – вспомогательные переменные (EC)

Следующим этапом анализа было построение МГК-моделей. МГК-модель – это регрессия на главных компонентах, а откликом выступает интересующий исследователей параметр.

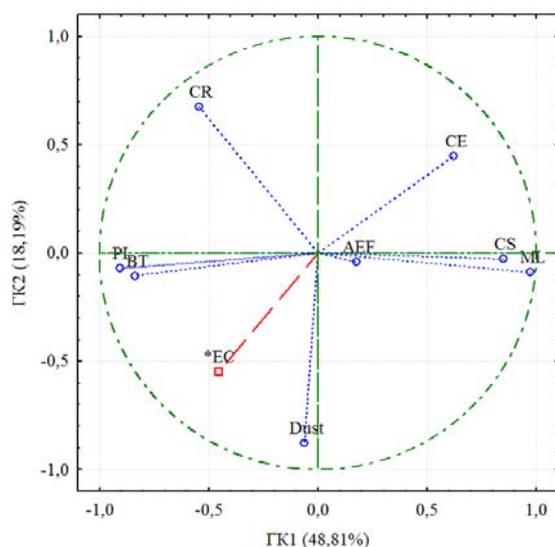


Рис. 4. Проекция переменных в плоскости GK1 и GK2 для КЭБТ № 2
□ – вспомогательные переменные (EC)

В нашем случае в качестве такого параметра был выбран расход электроэнергии. Следующие уравнения и представляют собой МГК-модели для анализа расхода электроэнергии корпуса 1 и корпуса 2 соответственно:

$$EC1 = 15\,165,1 - 90,3 \cdot GK1 + 64,6 \cdot GK2 - 91,4 \cdot GK3 + 245,6 \cdot GK4 - 31,1 \cdot GK5 - 79,5 \cdot GK6 - 7,6 \cdot GK7 - 0,6 \cdot GK8, \quad R^2 = 0,929; \quad (1)$$

$$EC2 = 15\,228,9 - 158,5 \cdot GK1 - 191,4 \cdot GK2 + 17,2 \cdot GK3 + 207,1 \cdot GK4 - 62 \cdot GK5 + 46,7 \cdot GK6 + 13,5 \cdot GK7 + 0,7 \cdot GK8, \quad R^2 = 0,907; \quad (2)$$

где EC1 и EC2 – расход электроэнергии для КЭБТ № 1 и КЭБТ № 2 соответственно.

Как видно из значений коэффициентов детерминации, уравнения хорошо описывают исходные выборки. Статистическое качество уравнений (1) и (2) подтверждено по критерию Фишера. Данные МГК-модели можно использовать для предсказания среднемесячных значений на уровне корпуса.

Рассмотрим зависимости остатков (Q) моделей (1)–(2) от расчетных значений результирующего признака – расхода электроэнергии (рис. 5, 6).

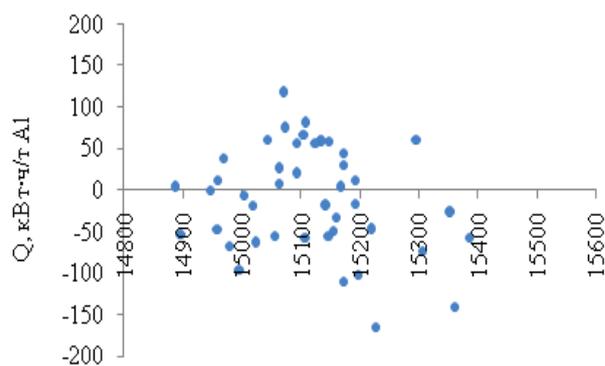
Анализ рисунков показывает, что остатки представляют собой случайные величины [6]. Расчетные значения результирующего признака хорошо аппроксимируют фактические значения.

Следует отметить, что данные уравнения неудобны для использования персоналом предприятия, по-

этому был выполнен переход к уравнениям в централизованных технологических переменных:

$$EC1 = 15\,165,1 + 130,36 \cdot CS - 529,29 \cdot CE + 59,10 \cdot AEF - 62,19 \cdot PL + 29,33 \cdot Dust + 29,76 \cdot BT - 64,68 \cdot CR + 98,81 \cdot ML; \quad (3)$$

$$EC2 = 15\,228,9 + 31,98 \cdot CS - 331,93 \cdot CE + 18,32 \cdot AEF - 1,85 \cdot PL + 8,2 \cdot Dust + 27,52 \cdot BT - 43,89 \cdot CR + 15,82 \cdot ML. \quad (4)$$



Расход электроэнергии по модели, кВт·ч/т Al

Рис. 5. Исследование остатков модели для КЭБТ № 1



Расход электроэнергии по модели, кВт·ч/т Al

Рис. 6. Исследование остатков модели для КЭБТ № 2

Расшифровка обозначений приведена в табл. 1. Графическое представление моделей и реального расхода электроэнергии приведено на рис. 7, 8.

Преобразовав централизованные параметры в истинные, получим влияние каждого отдельного параметра на расход электроэнергии (табл. 4). В таблице показано, как снижение каждого параметра на три величины его среднеквадратичного отклонения (3σ) влияет на расход электроэнергии в соответствии с уравнениями (3) и (4) для разных корпусов.

Влияние параметра в соответствии с литературными источниками [1–3] приведено в последнем столбце табл. 4. Как видно, характер влияния в полученных моделях для расхода электроэнергии согласуется с литературными источниками.

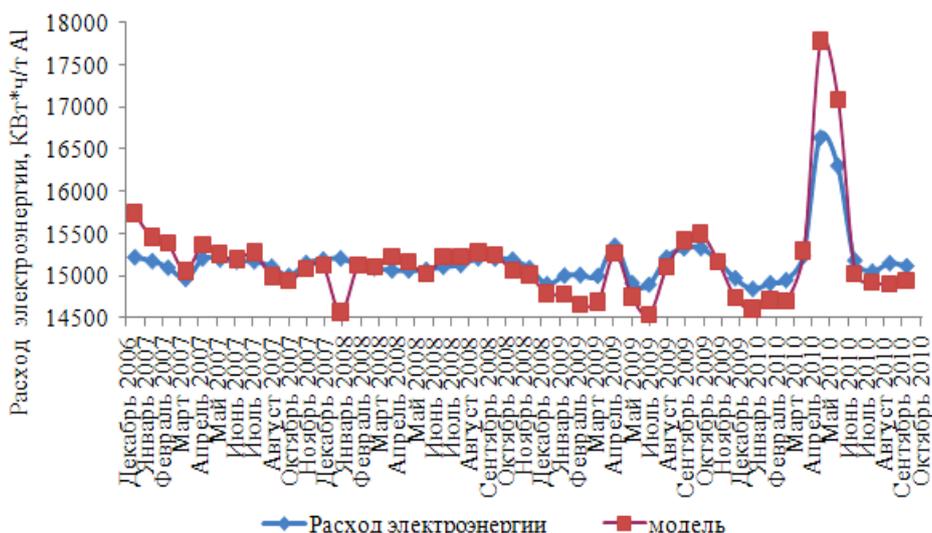


Рис. 7. Сравнение реального расхода электроэнергии и расхода, полученного по модели, для КЭБТ № 1

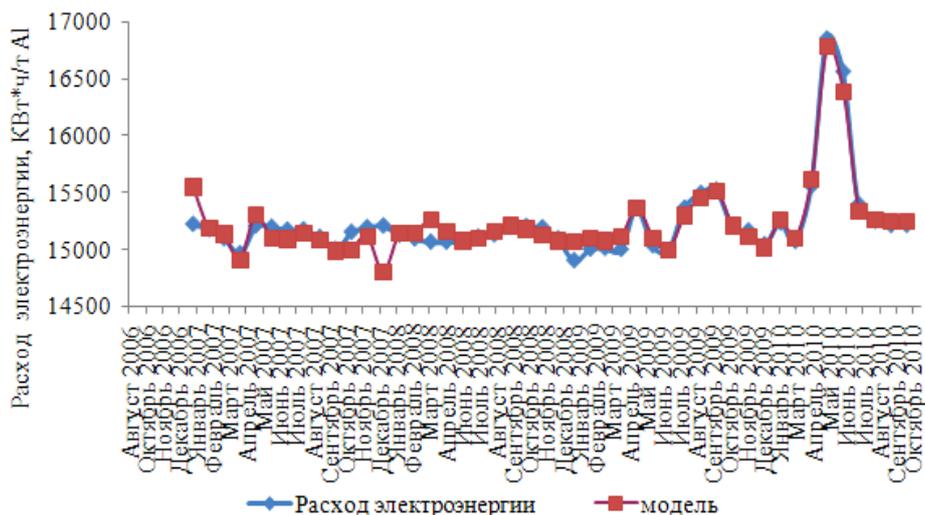


Рис. 8. Сравнение реального расхода электроэнергии и расхода, полученного по модели, для КЭБТ № 2

Таблица 4

Изменение расхода электроэнергии в зависимости от изменения параметров, взятых в анализ

Параметр	Единицы измерения	Корпус 1		Корпус 2		Характер зависимости по литературным данным
		Снижение параметра на 3 СКО	Изменение расхода электроэнергии, кВт·ч/т Al	Снижение параметра на 3 СКО	Изменение расхода электроэнергии, кВт·ч/т Al	
Сила тока	кА	-3,436	-391,08	-3,436	-95,95	Прямая
Выход по току	%	-5,105	1782,88	-5,288	995,79	Обратная
ЧАЭ	шт./сут.	-0,749	-177,29	-0,778	-54,95	Прямая
Срок службы	мес.	-29,324	186,58	-30,079	5,55	Обратная
Выход пены	кг/т Al	-15,540	-87,99	-13,686	-24,61	Прямая
Температура электролита	°С	-9,572	-89,28	-9,805	-82,55	Прямая
КО	доли ед.	-0,201	194,04	-0,212	131,66	Обратная
Уровень металла	см	-10,145	-296,42	-9,924	-47,46	Прямая

Библиографические ссылки

1. Introduction to aluminium electrolysis Understanding the Hall-Heroult process / K. Grjotheim, H. Kvande (ed.). Dusseldorf : Aluminium-Verlag, 1993.
2. Минцис М. Я., Поляков П. В., Сиразутдинов Г. А. Электрометаллургия алюминия. Новосибирск : Наука, 2001.
3. Grjotheim K., Welch B. J. Aluminium Smelter Technology. Dusseldorf : Aluminium-Verl. GmbH, 1980.
4. Светуных С. Г., Светуных И. С. Методы социально-экономического прогнозирования : учебник для вузов : в 2 т. Т. 1. СПб. : Изд-во СПбГУЭФ, 2009.
5. Ивченко Г. И., Медведев Ю. И. Введение в математическую статистику : учебник. М. : Изд-во ЛКИ, 2010.
6. Эконометрика : учебник / И. И. Елисеева, С. В. Курышева, Т. В. Костеева и др. ; под ред. И. И. Елисеевой. М. : Финансы и статистика, 2002.
7. Analysis of a potroom performance drift, from a multivariate point of view / J. Tessier, C. Duchesne, G. P. Tarcy et al. // Light Metals. 2008. P. 319–324.
8. Multivariate statistical process monitoring of reduction cell / J. Tessier, T. G. Zwirz, G. P. Tarcy, R. A. Manzini // Light Metals. 2009. P. 305–310.
9. Detecting abnormalities in aluminium reduction cells based on process events using multi-way principal component analysis (MPCA) / N. A. A. Majid, B. R. Young, M. P. Taylor, J. J. J. Chen // Light Metals. 2009. P. 589–593.
10. Multivariate monitoring of the prebaked anode manufacturing process and anode quality / J. Lauzon-Gauthier, C. Duchesne, J. Tessier et al. // Light Metals. 2011. P. 967–972.
11. A Taxonomy for Aluminium Process Fault Detection and Diagnosis / N. A. A. Majid, M. P. Taylor, J. J. J. Chen, B. R. Young // Proc. of the 10th Australasian Aluminium Smelting Technology Conf. Launceston. 2011.
12. Multiblock monitoring of aluminium reduction cells performance / J. Tessier, C. Duchesne, G. P. Tarcy // Light Metals. 2011. P. 407–412.

T. V. Piskazhova, P. V. Polyakov, N. A. Sharypov

USAGE OF METHODS OF MULTI-DIMENSIONAL DATA PROCESSING FOR ANALYSIS OF ENERGY CONSUMPTION OF ALUMINUM REDUCTION CELLS

The authors present a review of methods of application of multi-dimensional statistical analysis of database of aluminum industry. Estimation of effectiveness of devices and quality of products output is made. Application of major components for the analysis of energy consumption of two potrooms of an aluminum smelter is considered.

Keywords: multi-dimensional statistical analysis, major components analysis, aluminum reduction cell, energy consumption.

© Пискажова Т. В., Поляков П. В., Шарыпов Н. А., 2012

УДК 519.6

Е. А. Попов, М. Е. Семенкина, Л. В. Липинский

ПРИНЯТИЕ РЕШЕНИЙ КОЛЛЕКТИВОМ ИНТЕЛЛЕКТУАЛЬНЫХ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ*

Предлагается алгоритм генетического программирования для автоматического генерирования искусственных нейронных сетей. Предложена процедура принятия решений коллективом нейронных сетей, основанная на оценке компетентности членов коллектива. Подход апробирован на тестовых и реальных практических задачах.

Ключевые слова: интеллектуальные информационные технологии, алгоритм генетического программирования, коллективное принятие решений.

При математическом моделировании сложных систем и процессов часто используются так называемые интеллектуальные информационные технологии (ИИТ) – генетические алгоритмы (ГА) как метод оптимизации [1], искусственные нейронные сети (НС) как метод аппроксимации или классификации [2], генетическое программирование (ГП) как метод ав-

томатического генерирования математических выражений [3].

Проектирование интеллектуальных информационных технологий само по себе является достаточно сложной интеллектуальной процедурой, включающей, кроме всего прочего, выбор их эффективных структур.

* Работа выполнена при поддержке Министерства образования и науки Российской Федерации (ГК № 11.519.11.4002 и ГК № 16.740.11.0742).