

## РЕГРЕССИОННО-КОРРЕЛЯЦИОННЫХ АНАЛИЗ ТЕХНОЛОГИЧЕСКИХ ПЕРЕМЕННЫХ ПРИ ПРОИЗВОДСТВЕ АЛЮМИНИЯ

Клим Е. В.

Научный руководитель доктор технических наук Пискажова Т.В.  
*Сибирский федеральный университет*

Многофакторный анализ является важным инструментом исследования технико-экономических показателей (ТЭП) в металлургических процессах. При известных зависимостях ТЭП от технологических переменных возможно отыскание оптимальных областей работы металлургических аппаратов, а также разработка интеллектуальных систем управления производственным процессом.

Имеется следующая выборка: 37 технологических переменных и 51 измерение среднемесячных данных по алюминиевому заводу.

На данном этапе исследовалась зависимость выхода угольной пены от других технологических переменных. Для анализа был выбран метод пошагового отбора переменных. Этот метод заключается в том, что на каждом шаге вводят новые переменные, и с полученной выборкой делается регрессионный анализ и корреляционный анализ.

В первой выборке рассматриваем переменные, имеющие наибольший коэффициент корреляции с выходом пены. Обозначим  $y$  – выход пены (кг/т),  $x_1$  – высота настыли (см);  $x_2$  – частота анодного эффекта;  $x_3$  – криолитовое отношение. В программе Microsoft Excel получим уравнение множественной линейной регрессии на эти переменные:

$$\hat{y}_1(x) = 114,52 - 1,847 \cdot x_1 + 3,148 \cdot x_2 - 3,98 \cdot x_3. \quad (1)$$

Коэффициент детерминации  $R_2 = 0,512$ ;

Скорректированный коэффициент детерминации  $R_2 = 0,481$ ;

F-критерий Фишера  $F_1 = 16,437$ ;

Множественный коэффициент корреляции = 0,715.

Уравнение является статистически значимым, так как число Фишера больше критического значения. В таблице 1 приведены коэффициенты корреляции по первой выборке. Видно, что кроме корреляционной связи выходного показателя  $Y$  с входными  $x_i$ , имеется связь входных переменных между собой, что осложняет задачу исследования.

Таблица 1 – Корреляционная матрица по выборке 1.

	Y - Пена	X1 - Настыль	X2 - АЭ	X3 - КО
Y - Пена	1	-0,7115348	-0,2958093	-0,4923365
X1 - Настыль	-0,7115348	1	0,5065824	0,6880066
X2 - АЭ	-0,2958093	0,5065824	1	0,4000261
X3 - КО	-0,4923365	0,6880066	0,4000261	1

Поэтому нами были найдены коэффициенты частной корреляции, которые дают более точную характеристику тесноты связи двух признаков, чем коэффициенты парной корреляции, т.к. очищают парную зависимость от взаимодействия данной пары признаков с другими признаками, представленными в модели. Ниже представлено сравнение коэффициентов парной и частной корреляции.

$R_{yx1}$	-0,71153	$R_{yx1/x2x3}$	-0,58807
$R_{yx2}$	-0,29581	$R_{yx2/x1x3}$	-0,10752

$R_{yx3}$	-0,49234	$R_{yx3/x1x2}$	-0,01440
-----------	----------	----------------	----------

Видно, что коэффициент корреляции «выход пены - криолитовое отношение» значительно изменился, значит связь не подтвердилась, а значение первого коэффициента изменилось мало, значит связь «выход пены – высота настыли» подтвердилась.

Рассмотрим вторую выборку с добавлением переменных.  $Y$  – выход пены (кг/т),  $x_1$  – высота настыли (см);  $x_2$  – частота анодного эффекта;  $x_3$  – криолитовое отношение;  $x_4$  – напряжение АЭ (В);  $x_5$  – баланс пека в анодной массе (%). Получим следующее уравнение регрессии:

$$\hat{y}(x) = -22,673 - 0,056 \cdot x_1 + 0,00034 \cdot x_2 + 0,826 \cdot x_3 + 3,212 \cdot x_4 + 0,575 \cdot x_5. \quad (2)$$

Коэффициент детерминации  $R_2 = 0,799$

Скорректированный коэффициент детерминации  $R_2 = 0,777$

F-критерий Фишера  $F_2 = 35,891$

Множественный коэффициент корреляции = 0,894

Видно, что по сравнению с уравнением (1) в уравнении (2) значительно вырос коэффициент детерминации (на 28%) и в целом выборка 2 объясняет 80% вариаций выхода пены. Уравнение регрессии является статистически значимым, так как значение F – критерия Фишера превышает критическое значение. В таблице 2 приведена корреляционная матрица по выборке 2.

Таблица 2 – Корреляционная матрица по выборке 2.

	Пена	Высота настыли	Частота АЭ	КО	Напряжение АЭ	Баланс пека
Пена	1	-0,712	-0,296	-0,492	-0,845	0,659
Высота настыли	-0,712	1	0,507	0,688	0,617	-0,620
Частота АЭ	-0,296	0,507	1	0,400	0,134	-0,564
КО	-0,492	0,688	0,400	1	0,365	-0,576
Напряжение АЭ	-0,845	0,617	0,134	0,365	1	-0,524
Баланс пека	0,659	-0,620	-0,564	-0,576	-0,524	1

Ниже приведено сравнение парных и частных коэффициентов корреляции выхода пены с высотой настыли, с напряжением анодного эффекта, с величиной баланса пека.

$R_{yx1}$	-0,712	$R_{yx1/x4x5}$	-0,244
$R_{yx4}$	-0,845	$R_{yx4/x1x5}$	-0,681
$R_{yx5}$	0,659	$R_{yx5/x1x4}$	0,303

Видно, что связь «выход пены – напряжение анодного эффекта» подтвердилась, две другие связи сохранились, но значительно ослабли. В литературе нами не было выявлено применение методов очистки связей в металлургии. Полученные предварительные результаты говорят о перспективности метода.

Рассмотрим третью выборку с добавлением переменных.  $Y$  – выход пены (кг/т),  $x_1$  – высота настыли (см);  $x_2$  – частота анодного эффекта;  $x_3$  – криолитовое отношение;  $x_4$  – напряжение АЭ (В);  $x_5$  – баланс пека в анодной массе (%);  $x_6$  – содержание серы в анодной массе(%). Получим следующее уравнение регрессии:

$$\hat{y}(x) = 86,067 - 0,588 \cdot x_1 + 3,25 \cdot x_2 - 5,851 \cdot x_3 - 2,363 \cdot x_4 + 2,207 \cdot x_5 + 7,920 \cdot x_6. \quad (3)$$

Коэффициент детерминации  $R_2 = 0,812$

Скорректированный коэффициент детерминации  $R_2 = 0,787$

F-критерий Фишера  $F_2 = 31,83$

Множественный коэффициент корреляции = 0,902

Видно, что в выборке 3 множественный коэффициент корреляции вырос незначительно по сравнению с выборкой 2 и коэффициент детерминации увеличился только на 2%. Хотя добавленная переменная – содержание серы имела коэффициент корреляции с  $Y$  средней силы 0,52. Добавление нескольких следующих переменных не увеличило коэффициент детерминации в последующих уравнениях регрессии, поэтому процедуру пошагового отбора решено было остановить на выборке 3.