Д. т. н. Ю. А. ДОЛГОВ, С. Г. ФЕДОРЧЕНКО

Молдова, г. Тирасполь, Приднестровский госуниверситет
им. Т. Г. Шевченко

Дата поступления в редакцию
01.02 2000 г.
Оппонент *д. т. н. Р. А. ВОРОБЕЛЬ*

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ТЕХНОЛОГИЧЕСКОГО ПРОЦЕССА ПО РЕЗУЛЬТАТАМ ПАССИВНОГО ЭКСПЕРИМЕНТА

Приведены результаты исследования устойчивости метода построения модели к виду закона распределения факторов и выходной величины.

Для эффективного управления технологическим процессом (ТП) необходимо знать его математическую модель (ММ). Поскольку реальный ТП постоянно меняется, то возникает необходимость регулярно строить новую ММ, отражающую его текущее состояние. Исходя из этого необходимо найти такую процедуру построения ММ, которая, с одной стороны, была бы достаточно проста в реализации, с другой стороны, позволяла строить достаточно точную ММ.

Для построения ММ исследуемых объектов используются методы активного и пассивного эксперимента. При активном эксперименте исследователь управляет факторами, меняя их значения в соответствии с планом эксперимента [1, с. 114]. При пассивном эксперименте исследователь может только регистрировать значения факторов и выходной величины (целевой функции). Результаты пассивного эксперимента сводятся в таблицу исходных данных, каждая строка которой представляет собой числовое значение целевой функции при некоторых условиях (в определенный момент времени, или для определенной партии изделий, или при прохождении определенной технологической операции и т. п.) и числовые значения исследуемых факторов при тех же условиях. Такая таблица, как правило, является результатом длительных контрольных измерений выходного показателя качества однородной продукции и сопутствующих ему факторов, например, режимов технологических операций или параметров самого изделия на предшествующих операциях.

МОДИФИЦИРОВАННЫЙ МЕТОД СЛУЧАЙНОГО БАЛАНСА

Рассмотрим более подробно модифицированный метод случайного баланса (ММСБ), позволяющий получить математическую модель по результатам пассивного эксперимента.

Как известно, общими требованиями всех факторных планов являются некоррелированность (слабая коррелированность) факторов, нормальность закона распределения целевой функции Y, ортогональность факторов, гомоскедастичность, т. е. равенство выборочных дисперсий выходной величины во всех точках факторного пространства [1, с. 76].

Поскольку в случае пассивного эксперимента никакого искусственного варьирования факторов нет, то имеет место лишь естественное производственное варьирование — как правило, в пределах допуска на фактор, т. е. со статистической точки зрения незначительное. Это означает, что изменение целевой функции также будет небольшим, и чтобы отличить его от шумовых флуктуаций, необходимо иметь достаточно длинную таблицу, в которой возможный эффект воздействия конкретного параметра на целевую функцию проявился бы в достаточной мере. Опытным путем установлено, что таблица исходных данных (ТИД) будет достаточно длинной, если на каждый исследуемый в ней фактор приходится 10—15 строк, но не более 350 строк всего. Последнее ограничение связано с групповым характером технологического процесса, где всегда присутствует хотя бы малая корреляция параметров и, следовательно, нарушается один из постулатов теории вероятностей о независимости наблюдений, что приводит к нарушению требования о состоятельности оценки параметра по выборке определенного объема в математической статистике.

При пассивном эксперименте [2] мы имеем дело с контрольно-измерительной информацией, представленной в виде таблицы, которую можно рассматривать как таблицу координат факторов в абсолютных единицах и соответствующих им величин целевой функции. Естественно, что все такие координаты не могут находиться в вершинах гиперкуба (т. е. быть ортогональными) или хотя бы на гиперсфере одного радиуса, однако с допустимой погрешностью можно выбрать некоторые из них, приблизительно удовлетворяющие указанному условию. Для этого достаточно перейти к новой системе координат в относительных единицах. Правило перехода (кодирования факторов) следующее:

все значения $X_{k,i} \le X_{k,N_1-1}$ должны кодироваться как $x_{k,i}=-1$;

все значения $X_{k,i} \ge X_{k,N-N_1-1}$ должны кодироваться как $x_{k,i}=+1$;

остальные значения $X_{k,i}$ — как $x_k = 0$.

Здесь и в дальнейшем символом X_k обозначаются значения k-го фактора в абсолютных единицах, а x_k — в относительных.

Поясним правило кодирования факторов: весь диапазон изменения k-го фактора X_k $X_{k\max}-X_{k\min}$ следует разбить на три области, которые будем в дальнейшем называть опорными областями. Пусть ТИД содержит N строк. Расположим значения фактора X_k в порядке возрастания и выберем справа и слева из полученной последовательности N_1 рядом стоящих чисел. Те числа, которые занимают N_1+1 и $N-N_1-1$ места, и будут границами опорных областей. Вероятности попадания случайной величины X в каждую из крайних областей при таком способе разбиения будут одинаковы и независимы от вида закона распределения фактора.

Поскольку N в каждом конкретном случае будет различным, введем вспомогательный коэффициент K, характеризующий долю чисел от исходного объема выборки, участвующих в формировании опорных областей $N_1=KN$.

Как установлено в результате проведенных исследований [3], величина коэффициента K, при которой достигается наибольшая точность вычисления коэффициентов модели, является функцией меры крутости фактора τ:

$$K=0{,}42+0{,}1\exp(-\tau). \qquad (1)$$

В результате перехода от X_k к x_k исходная таблица с контрольно-измерительной информацией превращается в план квазиактивного эксперимента. При переходе к относительным координатам толщина оболочки гиперсферы увеличивается (или, иначе, вместо точечных вершин гиперкуба появляются некоторые "вершинные" области). Поэтому требуемая точность выделения значимых факторов и определения оценок коэффициентов регрессии может быть обеспечена только за счет увеличения числа опытов (уже упоминавшееся требование 10—15 строк ТИД, приходящихся на каждый исследуемый фактор).

Поскольку ТИД достаточно длинная, неизбежны совпадения некоторых строк плана, у каждой из которых, тем не менее, имеется свое значение выходной величины. Такие совпадающие строки плана следует совместить, то есть представить в конечном плане в виде одной строки с несколькими значениями выходной величины, которые необходимо рассматривать как выборку. Пример плана эксперимента приведен в **табл. 1**.

Таблица 1

Пример плана и результатов эксперимента

j	План эксперимента			Результаты эксперимента	m_j	$\bar{Y}_j$	S_j^2
	x_1	x_2	x_1x_2				
1	–	–	+	1,82; 1,75; 1,73; 1,79; 1,81; ...	7	1,85	0,13
2	–	0	0	1,81; 1,96; 1,80; 1,80; 1,92;	6	1,87	0,07
3	–	+	–	1,68; 1,63; 1,74	3	1,69	0,21
4	0	–	0	1,76; 1,75; 1,72; 1,68; 1,74;	14	1,72	0,18
5	0	0	0	1,68; 1,64; 1,66; 1,70; 1,59;	11	1,66	0,22
6	0	+	0	1,56; 1,49; 1,70; 1,62; 1,61;	7	1,61	0,10
7	+	–	–	1,71; 1,75; 1,72; 1,82; 1,68;	17	1,71	0,17
8	+	0	0	1,52; 1,48; 1,58; 1,54; 1,50;	10	1,56	0,20
9	+	+	+	1,57; 1,52; 1,47; 1,64	4	1,55	0,16

Экспериментальные данные, особенно полученные в условиях реального производства, как правило, содержат некоторое количество так называемых “грубых промахов”, не присущих исследуемому объекту (технологическому процессу). К сожалению, большинство этих “грубых промахов” не видны на общем фоне, однако с расслоением общей выборки на частные по строкам плана появляется возможность проверить каждую строчную выборку на однородность (отсутствие аномальных измерений). Обнаруженные “грубые промахи” должны быть удалены из таблицы и не должны учитываться в дальнейших расчетах.

Из плана эксперимента можно извлечь дополнительную информацию о влиянии парных (как правило, не выше) взаимодействий, которые иногда могут иметь большее синергетическое воздействие на выходную величину, чем влияние каждого фактора в отдельности. С этой целью в план эксперимента включаются столбцы парных, иногда тройных, взаимодействий, каждая координата которых получается простым перемножением кодов координат исходных факторов.

Наконец, последнее общее требование факторных планов — гомоскедастичность — в квазиактивном плане ММСБ нарушается, поэтому для расчетов оценок коэффициентов регрессии b_k и их дисперсии D_k следует использовать специальные выражения, учитывающие поправки на это нарушение гомоскедастичности (гетероскедастичность) и являющиеся в этих условиях более эффективными, чем другие оценки [2]:

$$b_k=\frac{\left(\dfrac{D_{2k}}{N_{2k}}+2m_k^2\right)\mu_{1k}-\left(\dfrac{D_{1k}}{N_{1k}}+2m_k^2\right)\mu_{2k}}{\dfrac{D_{1k}}{N_{1k}}+\dfrac{D_{2k}}{N_{2k}}+4m_k^2}; \qquad (2)$$

$$D_k=\frac{\dfrac{D_{1k}}{N_{1k}}\dfrac{D_{2k}}{N_{2k}}+\left(\dfrac{D_{1k}}{N_{1k}}+\dfrac{D_{2k}}{N_{2k}}\right)m_k^2}{\dfrac{D_{1k}}{N_{1k}}+\dfrac{D_{2k}}{N_{2k}}+4m_k^2}, \qquad (3)$$

где D_k– дисперсии выходной величины при положительных и отрицательных значениях фактора x_k, соответственно —

$$D_{1k}=\frac{1}{N_{1k}-1}\sum_{j=1}^{N_{1k}}\left(Y_j^{(1k)}-\mu_{1k}\right)^2,$$

$$D_{2k}=\frac{1}{N_{2k}-1}\sum_{j=1}^{N_{2k}}\left(Y_j^{(2k)}-\mu_{2k}\right)^2;$$

N_{1k}, N_{2k} — объем соответствующих подмножеств, причем $N_k=N_{1k}+N_{2k}$ — общий объем выборки для k-го фактора;

$\left\{Y_j^{(1k)}\right\}_{N_{1k}}$ и $\left\{Y_j^{(2k)}\right\}_{N_{2k}}$ — подмножества элементов выходной величины из общей выборки, для которых x_{kj} имеет соответственно положительный или отрицательный знак;

$$\mu_{1k} = \frac{1}{N_{1k}} \sum_{j=1}^{N_{1k}} Y_j^{(1k)},\ \mu_{2k} = \frac{1}{N_{2k}} \sum_{j=1}^{N_{2k}} Y_j^{(2k)};$$

$m_k = \frac{1}{N_k}(\mu_{1k}N_{1k} + \mu_{2k}N_{2k})$ — общая оценка математического ожидания.

При вычислении оценок коэффициентов модели мы использовали только те значения факторов, которые принадлежат областям, кодируемым как +1 и −1. Нетрудно видеть, что при таком подходе мы не можем оценить влияние квадрата фактора, т. к. x_k^2 всегда будет равен 1.

Попытаемся вычислить коэффициенты модели при x_k^2 опираясь на области, закодированные как +1 и −1, с одной стороны, и как 0 — с другой. Введем следующие обозначения:

N_{1k}, N_{0k} — объем выборок выходной величины, для которых x_k^2 имеет значения 1 и 0, соответственно;

μ_{1k}, μ_{0k} — средние значения частных выборок выходной величины Y, для которых x_k^2 имеет значения 1 и 0, соответственно;

D_{1k}, D_{0k} — дисперсии частных выборок выходной величины, для которых x_k^2 имеет значения 1 и 0, соответственно.

Тогда выражение для вычисления оценки коэффициентов модели при квадратах факторов примет вид

$$b_{kk} = 2\frac{\left(\frac{D_{0k}}{N_{0k}} + 2m_k^2\right)\mu_{1k} - \left(\frac{D_{1k}}{N_{1k}} + 2m_k^2\right)\mu_{0k}}{\frac{D_{1k}}{N_{1k}} + \frac{D_{0k}}{N_{0k}} + 4m_k^2}. \qquad (4)$$

Сравнивая это выражение с аналогичным выражением, полученным для одиночных факторов, можно видеть, что они совпадают с точностью до постоянного множителя, равного 2. Выражение для дисперсии оценки коэффициента модели примет в этом случае вид

$$D[b_{kk}] = \qquad (5)$$

$$= \frac{\left(4m_k^2 + 2\frac{D_{0k}}{N_{0k}}\right)^2 \frac{D_{1k}}{N_{1k}} + \left(4m_k^2 + 2\frac{D_{1k}}{N_{1k}}\right)^2 + \left(\frac{D_{0k}}{N_{0k}} - \frac{D_{1k}}{N_{1k}}\right)^2 4m_k^2}{\left(\frac{D_{1k}}{N_{1k}} + \frac{D_{0k}}{N_{0k}} + 4m_k^2\right)^2}.$$

После определения оценок коэффициентов регрессии необходимо проверить гипотезу о значимости коэффициентов b_k при помощи t — критерия Стьюдента; лучше всего это сделать в виде нуль-гипотезы, т. е. гипотезы о равенстве $b_k = 0$:

$$t_k = \frac{|b_k|}{\sqrt{D_k}} \geq t_{\text{таб}}(q; \nu_k). \qquad (6)$$

Если вычисленная величина параметра t_k не превышает табличного значения $t_{\text{таб}}$, найденного для q%-ного уровня значимости и $\nu_k = N(m-1)$ числа степеней свободы, то гипотеза о равенстве нулю оценки коэффициента b_k принимается, он считается незначимым и его следует отбросить, не включая в искомую модель.

Проверка взаимодействий факторов на значимость ничем не отличается от соответствующих процедур для каждого из основных факторов.

Если среди незначимых оказался хотя бы один основной фактор x_k, то вследствие его исключения таблица плана эксперимента должна быть преобразована (свернута) и вся работа по определению оценок коэффициентов и их значимости проделана заново.

Для оставшихся после устранения “грубых промахов” данных плана по каждой строке вычисляются средние арифметические и дисперсии (см. табл. 1). Полученные результаты можно использовать при проверке на воспроизводимость опытов, которую, в силу выборок неодинакового объема, следует проводить по критерию Бартлетта

$$\chi^2 = \frac{1}{C}\left[\nu \ln S_p^2 - \sum_{j=1}^{N} \nu_j \ln S_j^2\right] \leq \chi^2_{\text{таб}}(q; \nu_N = N-1), \qquad (7)$$

$$C = 1 + \frac{1}{3(N-1)}\left[\sum_{j=1}^{N} \frac{1}{\nu_j} - \frac{1}{\nu}\right],$$

где C — поправочный коэффициент;
N — число строк плана;
— общее число степеней свободы плана;
— число степеней свободы j-й строки плана, объем выборки которой равен m_j;
— средневзвешенная дисперсия;
S_j^2 — дисперсия j-й строки плана.

В случае выполнения условия (7) результаты опытов правильно отражают реальную картину и могут быть использованы для дальнейших расчетов. При этом средняя дисперсия опытов может быть принята равной средневзвешенной: $S^2\{Y\} = S_p^2$. После этого математическая модель объекта составляется в виде уравнения связи выходного параметра Y и переменных x_k, включающего только значимые коэффициенты.

Чтобы проверить гипотезу об адекватности представления результатов эксперимента найденному уравнению связи (иными словами, чтобы проверить, насколько найденное уравнение соответствует экспериментальным результатам), достаточно оценить отклонение выходной величины $\hat{Y}_j$, предсказанное уравнением регрессии, от результатов экспериментов Y_j в точках факторного пространства.

Рассеяние результатов эксперимента вблизи уравнения связи, аппроксимирующего искомую функциональную зависимость, можно охарактеризовать с помощью дисперсии неадекватности σ_a^2, оценка которой $S_{\text{ад}}^2$ находится по формуле,

$$S_{\text{ад}}^2 = \frac{1}{N-d}\sum_{j=1}^{K}\sum_{i=1}^{m_j}(Y_{ij}-\hat{Y}_j)^2\ , \qquad (8)$$

где N — количество экспериментальных значений выходной величины;
d — число значимых членов модели;
K — количество строк плана эксперимента;
Y_{ij} — экспериментальное значение выходного параметра, полученное в j-й точке факторного пространства при i-м измерении;
— значение выходного параметра, найденное по уравнению регрессии в тех же точках.

Проверка адекватности состоит в выяснении соотношения между дисперсией неадекватности $S_{\text{ад}}^2$ и дисперсией воспроизводимости $S^2\{Y\}$ и проводится с использованием F — критерия Фишера, который позволяет проверить гипотезу о равенстве двух генеральных дисперсий:

$$F = \frac{S_{\text{ад}}^2}{S^2\{Y\}} \le F_{\text{таб}}\ (q;\nu_1;\nu_2)\ . \qquad (9)$$

Если вычисленное значение F меньше табличного $F_{\text{таб}}$, найденного для q%-ного уровня значимости, $\nu_1 = K-d$ числа степеней свободы числителя и $\nu_2 = \sum_{j=1}^{K}(m_j-1)$ числа степеней свободы знаменателя, то нуль-гипотеза, состоящая в утверждении адекватности модели, принимается. В противном случае она отвергается, и модель признается неадекватной.

В ходе работы может возникнуть ситуация, когда выборочная дисперсия неадекватности не превосходит оценки дисперсии воспроизводимости. Тогда неравенство $F<F_{\text{таб}}$ выполняется для любого числа степеней свободы ν_1 и ν_2, то есть нуль-гипотеза не противоречит выборочным данным, и математическая модель адекватно представляет объект.

ММСБ ПРИ НАРУШЕНИИ НОРМАЛЬНОСТИ ЗАКОНА РАСПРЕДЕЛЕНИЯ ВЫХОДНОЙ ВЕЛИЧИНЫ

При обработке реальных данных достаточно часто встречается ситуация, когда требование нормальности выходной величины не выполняется. В этом случае исследователь должен преобразовать выходную величину к нормальному виду и относительно преобразованной величины строить модель. После построения адекватной модели он должен, применив обратное преобразование, получить математическую модель относительно первоначального вида выходной величины. Такая процедура, с одной стороны, достаточно трудоемка, а с другой — полученная в этом случае модель сложна для восприятия. Вместе с тем опыт применения модифицированного метода случайного баланса показал, что возможно получение адекватной модели при отклонении закона распределения выходной величины от нормального в некоторых пределах без ее преобразования к нормальному виду. Возникает необходимость выявления границ допустимого отклонения и появляющуюся при этом погрешность оценки коэффициентов модели.

В результате проведенных исследований было установлено, что ММСБ позволяет строить адекватные ММ при выходной величине, не подчиняющейся нормальному закону (рассматривались только унимодальные законы распределения). Наибольшее влияние на точность оценивания коэффициентов модели оказывала мера кососити выходной величины α. При этом верхняя граница допустимого изменения α зависит от максимально приемлемой погрешности оценки коэффициентов модели $\Delta b_{\max}$% (см. **табл. 2**). Здесь в качестве меры погрешности нахождения коэффициентов модели используется величина

$$\Delta b_i\% = \frac{\left|b_i-\hat{b}_i\right|}{b_i}100\%\ ,$$

где — оценка влияния i-го фактора, полученная с помощью ММСБ;
b_i — истинное значение коэффициента влияния i-го фактора.

Таблица 2

Допустимые максимальные значения α *для различных* Δb

$\Delta b_{\max}$, %	Асимметрия выходной величины $\alpha_{y\max}$
20	0,3
30	0,35
40	1,3
50	1,4

Исследование, результаты которого приведены в табл. 2, проведены для коэффициента влияния доминирующего фактора (коэффициент влияния этого фактора является наибольшим), точность нахождения которого наивысшая. Точность же нахождения менее значимых коэффициентов модели, оценивающих вклад других линейных членов, будет ниже.

Как правило, модели, полученные при анализе реальных объектов, содержат не только линейные члены, но и взаимодействия факторов. Точность определения численных значений коэффициентов модели, характеризующих вклад взаимодействий факторов, обычно значительно ниже, чем точность оценки линейных членов. Проведенные исследования показали, что точность нахождения оценок коэффициентов модели при парных взаимодействиях не обнаруживает ярко выраженной зависимости от величин параметров распределений факторов и выходной величины; она колеблется в районе 50%.

Вывод — ММСБ позволяет строить модели, достаточно хорошо описывающие исследуемый объект при отклонении закона распределения выходной величины от нормального вида. При этом точность вычисления линейных коэффициентов модели несколько ухудшается, точность же определения коэффициентов при парных взаимодействиях существенного изменения не претерпевает.

ПРОВЕРКА АДЕКВАТНОСТИ МОДЕЛИ

Классическая методика проверки адекватности модели, используемая в активном эксперименте, требует, чтобы в каждой точке факторного пространства, координаты которой задаются содержанием соответствующей строки плана эксперимента, было проведено m экспериментов ($m \geq 3$). При выполнении этого требования можно найти дисперсию выходной величины при неизменных значениях факторов как дисперсию выходной величины, соответствующую каждой строке плана эксперимента, и усреднив полученные результаты, получить таким образом оценку величины ошибки эксперимента [1, с. 75]. Обозначим в дальнейшем эту величину как $S^2\{Y\}$. В качестве меры отклонения предсказанных моделью значений выходной величины от соответствующих им экспериментальных значений используется так называемая дисперсия адекватности $S_{\text{ад}}^2$ [1]. Если $S_{\text{ад}}^2$ меньше $S^2\{Y\}$ или статистически неотличима от нее, то модель признается адекватной.

Вышеизложенная методика была заимствована из теории активного эксперимента и применена для оценки адекватности моделей, полученных при обработке результатов пассивного эксперимента с помощью ММСБ. При этом, как правило, лишь часть строк плана эксперимента содержит повторяющиеся опыты. Те строки плана, которым соответствуют 1 или 2 эксперимента, из рассмотрения исключаются.

Обычно исследователь поступает следующим образом: он последовательно объявляет незначимыми факторы (взаимодействия факторов), попавшие в так называемую "серую зону значимости", фиксируя при этом величину меры адекватности модели. Верхняя граница этой зоны, установленная опытным путем, для t-статистики коэффициентов модели при одиночных факторах составляет в среднем 2,5, а для коэффициентов при парных взаимодействиях — 3,0. Исследователь старается подобрать такой набор коэффициентов модели из "серой зоны", чтобы обеспечить наибольшую адекватность модели. Опыт применения этой методики показал, что из первоначально признанных значимыми 10—15 факторов и их взаимодействий в модель, признанную адекватной, обычно попадают 3—5 факторов.

Рассмотрим этот процесс более подробно на примере данных, полученных в ходе контроля за технологическим процессом на одном из предприятий региона. Обрабатываемые данные характеризовались следующими показателями:

а) выходная величина, относительно которой строилась ММ, подчиняется нормальному закону, следовательно, для проверки адекватности модели можно воспользоваться классической мерой;

б) количество факторов (столбцов ТИД) — 13;

в) количество строк ТИД — 175.

Поступим следующим образом: из коэффициентов модели, признанных значимыми ММСБ, будем выбирать коэффициент, обладающий минимальным значением t-статистики, и объявлять его незначимым. Для каждого коэффициента, объявленного незначимым, будем фиксировать его индексы i, j, а также величины F, $S_{\text{ад}}^2$, $S^2\{Y\}$ и значение t-статистики t_{ij}.

Результаты вычислений сведены в **табл. 3**. Строка таблицы, выделенная жирным шрифтом, обозначает начало области адекватных моделей, определенной по классическому критерию.

Таблица 3

Поведение $S_{\text{ад}}^2$, $S^2\{Y\}$, F при последовательном объявлении факторов незначимыми

i	j	t_{ij}	$S^2_{\text{ад}}$	$S^2\{Y\}$	F	f^2	R	t_R
10	11	1,98	136,8	0	∞	1,1	–	–
6	0	2,05	138,9	22,98	6,30	1,06	–	–
10	12	2,06	130,6	22,98	5,68	0,999	0,03	0,4
2	10	2,09	126,2	22,98	5,49	0,966	***0,18***	***2,4***
2	0	2,12	108,9	63,43	1,72	0,879	0,35	5,1
6	7	2,22	108,1	47,34	2,28	0,875	0,35	5,1
6	***8***	***2,23***	***101,6***	***142,9***	***0,71***	***0,853***	0,38	5,8
9	9	2,25	99,08	87,91	1,13	0,881	0,35	5,0
1	8	2,43	98,02	87,91	1,12	0,871	0,36	5,3
4	8	2,50	80,41	77,16	1,04	0,879	0,35	5,1
7	0	2,51	72,30	73,51	0,98	0,839	0,40	6,2
1	7	2,54	69,85	71,38	0,98	0,812	0,43	6,9
6	10	2,63	50,85	111,1	0,46	0,769	0,48	8,1
3	13	2,78	29,36	109,3	0,27	0,642	0,6	12,2

Обращает на себя внимание поведение величины F, которая, если представить ее в виде кривой, является немонотонной. Вначале модель признается неадекватной (F = ∞, 6,3, 5,68, 5,49, 1,72, 2,28), затем резкий скачок до 0,71, — и все остальные строки табл. 3 соответствуют адекватным моделям. Можно предположить, что существует конкретный коэффициент модели (в нашем случае при x_6x_8), который и оказывает наибольшее влияние на адекватность модели. Однако при дальнейшем анализе этот вывод не подтвердился. Если объявлять незначимыми коэффициенты модели в другом порядке, то резкий скачок значений F, после которого модели признаются значимыми, наступит после объявления незначимым уже другого коэффициента.

Проанализируем поведение составляющих критерия F – $S^2_{\text{ад}}$ и $S^2\{Y\}$. Значение $S^2_{\text{ад}}$ изменяется достаточно плавно, без резких скачков, а вот $S^2\{Y\}$ — нет. Именно эта величина претерпевает скачок, после которого все остальные модели признаются значимыми. Объяснение этого скачка состоит в том, что после удаления из рассмотрения (так как они были объявлены незначимыми) ряда коэффициентов модели план эксперимента меняется, количество строк плана становится меньше и все большее число строк плана имеет повторяющиеся опыты. Следовательно, привлекается для анализа адекватности модели большее количество опытных данных, что и привело к резкому изменению величины $S^2\{Y\}$.

Из всего вышеизложенного можно сделать вывод, что предлагаемая методика оценки адекватности модели не подходит для случая пассивного эксперимента, когда число факторов велико, а количество строк ТИД, как правило, не зависит от исследователя. Для одного и того же набора данных можно построить целое семейство адекватных моделей, используя различные комбинации коэффициентов модели, которые можно объявлять незначимыми. Хотелось бы отметить, что величины F для всех моделей из этого семейства будут весьма близки друг к другу. Таким образом, применение классической меры адекватности модели, заимствованной из теории активного эксперимента, фактически приводит к тому, что исследователь случайным образом выбирает коэффициенты, попавшие в модель, что неприемлемо. Источник этой неопределенности заключается в том, что мы всегда вынуждены исключать из рассмотрения значительную часть строк плана эксперимента, в которых нет повторяющихся опытов. Первая строка табл. 3 представляет крайний случай, когда в плане нет ни одной строки с повторяющимися опытами и невозможно найти величину $S^2\{Y\}$, а следовательно, нельзя оценить адекватность модели.

Возникает необходимость в поиске иной меры адекватности модели, которая была бы свободна от указанного недостатка.

Предлагается следующая мера адекватности модели:

$$f^2 = \frac{\sum_{j=1}^{k} (\overline{Y}_j - \hat{Y}_j)^2 m_j}{D[y](N-d)}. \qquad (10)$$

Здесь использованы следующие обозначения:

- $\overline{Y}_j$ — среднее значение экспериментальных значений выходной величины Y в j-й строке плана;
- $\hat{Y}_j$ — значение Y, вычисленное в j-й строке плана в соответствии с моделью;
- m_j — количество повторных опытов в j-й строке плана;
- $D[y]$ — выборочная дисперсия, найденная по всей выборке значений выходной величины Y;
- N — количество строк таблицы исходных данных;
- d — количество значимых коэффициентов модели.

Предлагаемая мера адекватности модели свободна от недостатка классической меры — она не требует повторных опытов в каждой исследуемой точке факторного пространства. Она не зависит от выбранной единицы измерения, т. к. нормирована на величину дисперсии выходной величины. Предлагаемая мера есть не что иное, как средневзвешенное значение квадратов невязок, нормированное на величину $D[y]$.

Пример использования новой меры для оценки адекватности модели приведен в табл. 3. В столбце, озаглавленном как f^2, приведены значения новой меры, вычисленной для соответствующих строк таблицы. Видно, что эта мера не испытывает таких резких скачков, как величина F, и оказывается работоспособной и в том случае, когда F вычислить не удается.

Предварительные результаты применения новой меры позволяют предположить, что областью адекватности моделей для нее можно считать $f^2<1$.

Как известно [4, с. 70], в качестве меры адекватности модели, наряду с дисперсией адекватности, используется коэффициент множественной корреляции, вычисляемый по формуле

$$R = \sqrt{1 - \frac{N-1}{N-d-1}(1-R')^2}\,, \qquad (11)$$

где $R' = \sqrt{1 - \dfrac{\sum_{i=1}^{N}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{N}(Y_i - \overline{Y}_i)}}$.

Сопоставляя формулы (10) и (11), можно видеть, что коэффициент множественной корреляции может быть выражен через величину f^2 следующим образом:

$$R = \sqrt{1 - f^2}\,. \qquad (12)$$

Значимость R может быть определена путем проверки нуль-гипотезы о равенстве R нулю с помощью t-распределения Стьюдента. Так, если справедливо неравенство

$$t_R = \frac{R}{S_R} > t(q, N-d-1), \text{ где } S_R = \frac{1-R^2}{\sqrt{N-D-1}},$$

то нуль-гипотеза отклоняется и R признается значимым.

Последние две колонки табл. 3 содержат значения R и t_R. Видно, что для случая $f^2>1$ R не определено, так как подкоренное выражение принимает отрицательное значение. Начиная же с четвертой строки таблицы все значения множественного коэффициента корреляции признаются значимыми, что позволяет говорить об адекватности полученных математических моделей.

Вывод – для проверки адекватности ММ, полученной по результатам пассивного эксперимента с помощью ММСБ, необходимо использовать множественный коэффициент корреляции, а не дисперсию адекватности.

ЗАКЛЮЧЕНИЕ

Рассмотренный в статье модифицированный метод случайного баланса позволяет построить математическую модель исследуемого объекта по результатам пассивного эксперимента. Область применимости ММСБ достаточно широка, т. к. он не требует выполнения условия гомоскедастичности выходной величины — в отличие от других методов, например метода наименьших квадратов. Нарушение требования нормальности распределения выходной величины в допустимых пределах (см. табл. 2) также не является препятствием к использованию ММСБ. ТИД при использовании ММСБ подвергается предварительному кодированию, что позволяет перейти от пассивного к псевдоактивному эксперименту. С одной стороны, это приводит к потере части информации, т. к. производится свертка многомерной выборки, с другой стороны, позволяет выделить многомерные грубые промахи, а также обеспечивает устойчивость метода к нарушению требования нормальности закона выходной величины. Так как при кодировании каждого фактора учитывается его закон распределения, то это позволяет обрабатывать таблицы исходных данных, содержащие факторы, подчиняющиеся произвольным законам распределения (ограничение состоит в унимодальности закона распределения фактора).

Таким образом, предлагаемый метод является достаточно универсальным и может быть рекомендован для обработки результатов пассивного эксперимента в различных областях науки и техники.

ИСПОЛЬЗОВАННЫЕ ИСТОЧНИИ

1. Планирование эксперимента в исследовании технологических процессов / К. Хартман, Э. Лецкий, В. Шефер и др. — М. : Мир, 1977.
2. Долгов Ю. А. Модифицированный метод случайного баланса // Электрон. моделирование. — 1987. — Т. 9, № 4. — С. 79—84.
3. Федорченко С. Г. Построение модели по результатам пассивного эксперимента // Радиоэлектроника, информатика, управление (Запорожский гос. техн. ун-т). — 1999. — № 1. — С. 106—108.
4. Львовский Е. Н. Статистические методы построения эмпирических формул. — М. : Высшая школа, 1988.

Н. Н. ПРОШКИН

Украина, г. Одесса, НИИ «Шторм»

Дата поступления в редакцию 02.03 2000 г.
Оппонент *д. т. н. А. Л. ВАЙНЕР*

ТЕХНОЛОГИЧЕСКИЕ МЕТОДЫ ПОВЫШЕНИЯ НАДЕЖНОСТИ ТЕРМОЭЛЕКТРИЧЕСКИХ МОДУЛЕЙ

Исследованы технологические операции электрохимического травления и антидиффузионного покрытия ветвей термоэлементов, способствующие росту показателей надежности модулей.

Проблема повышения надежности термоэлектрических модулей охлаждения—нагрева, учитывающей различные факторы [1], остается актуальной.

Рост показателей надежности достигается, с одной стороны, схемно-конструктивными решениями, с другой —выбором технологии производства и поддержанием достигнутого уровня надежности при эксплуатации.

В настоящее время в производстве термоэлектрических модулей (ТЭМ) широко используется следующая технологическая схема:

— получение направленных термоэлектрических кристаллов из сплавов или исходных полупроводниковых материалов;

— разрезка кристаллов одним из бездефектных методов (электроискровой с проволочным электродом или механический с использованием алмазного диска или струны);

— обработка (подготовка) поверхности и лужение ветвей термоэлементов;

— получение рисунка металлизации на керамических теплопереходах;

— одновременная коммутация ТЭМ в многоместном приспособлении (кондукторе).

Технологический процесс *изготовления полупроводниковых термоэлектрических кристаллов* достаточно отработан, и качество здесь обеспечивается прежде всего высоким уровнем технологического оборудования и квалификацией рабочих и ИТР.

Современное производство ТЭМ выдвигает на первый план проблему экспрессности и достоверности контроля основных параметров кристаллов. Существующие методики и измерительные средства контроля предусматривают раздельное измерение параметров кристаллов и необходимость разрезки их на части размером до 20 мм, что приводит к снижению процента выхода годного материала. Методы измерения, в свою очередь, отличаются разнообразием и, как правило, невысокой точностью и низкой производительностью. Такое положение обусловлено многими причинами, и в частности тем, что в отличие от других, почти все производители ТЭМ пользуются установками контроля индивидуального изготовления.

Следующим существенным моментом в технологической цепочке является *разрезка термоэлектрических кристаллов* на ветви термоэлементов и их обработка, т. е. подготовка рабочих поверхностей к залуживанию. Использование электроискрового оборудования позволило значительно ускорить процесс разрезки кристаллов, а разрезка с применением ЧПУ позволит исключить ошибки оператора и автоматизировать процесс. Для ветвей термоэлементов с малыми габаритами перспективным представляется способ механической разрезки на установке струнной резки.

Обработка (подготовка) поверхности полученных после разрезки ветвей термоэлементов осуществляется в основном механическим и химическим способами. Каждый из этих способов выбирается в зависимости от метода обработки ветвей (например,