МИНОБРНАУКИ РОССИИ

ФГБОУ ВПО «УРАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ ЛЕСОТЕХНИЧЕСКИЙ УНИВЕСИТЕТ»

Кафедра технической механики и оборудования целлюлозно-бумажных производств

В.П. Сиваков
В.И. Музыкантова

МОДЕЛИ И МЕТОДЫ ФАКТОРНЫХ ЭКСПЕРИМЕНТОВ

Методические указания к лабораторным и практическим занятиям по курсу «Основы научных исследований и физического эксперимента» для студентов очной и заочной формы обучения направлений 15.03.02, 15.04.02 «Технологические машины и оборудование»

Екатеринбург 2016

Печатается по решению кафедры технической механики и оборудования ЦБП, протокол № 5 от 27.01.2016 г.

Рецензент профессор, д.т.н. А.А. Санников

Редактор Черных Л.Д.

Подписано в печать		Поз. 103
Печать плоская	Формат 60x84 1/16	Тираж 50 экз.
Заказ	Печ.л. 1,16	Цена руб.

Редакционно-издательский отдел УГЛТУ

Отдел оперативной полиграфии УГЛТУ

СОДЕРЖАНИЕ

1. РЕГРЕССИОННЫЙ АНАЛИЗ ФАКТОРНЫХ

ЭКСПЕРИМЕНТОВ

Исследование закономерностей связи между явлениями (процессами), которые зависят от многих, иногда неизвестных факторов, называется **регрессионным анализом** [1,2,3].

Фактором называется измеряемая величина, принимаемая в некоторый момент времени. Факторы могут быть постоянными и переменными.

Часто между переменными *x* и *y* существует связь, но не вполне определенная, при которой одному значению *x* соответствует несколько значений *y*. Связь называют регрессионной (корреляционной), если каждому значению аргумента *x* соответствует статистический ряд распределения *y*. Регрессионные зависимости характеризуются вероятностными или стохастическими связями. Поэтому установление регрессионных зависимостей между величинами *x* и *y* возможно лишь тогда, когда выполнимы статистические измерения.

Статистические зависимости описываются математическими моделями процесса, т.е. регрессионными выражениями, связывающими независимые значения *x* (факторы) с зависимой переменной *y*. Модель по возможности должна быть простой и адекватной.

Регрессионный анализ сводится к установлению уравнения регрессии, т.е. вида кривой между случайными величинами (аргументами *x* и функцией *y*), оценке тесноты связей между ними, достоверности и адекватности результатов измерений.

Чтобы предварительно определить наличие такой связи между *x* и *y*, наносят точки на график и строят корреляционное поле (рис. 1). По тесноте группирования точек вокруг прямой или кривой линии, по наклону линии можно визуально судить о наличии корреляционной связи. Так из рис. 1,а видно, что экспериментальные данные имеют определенную связь между *x* и *y*, а измерения, приведенные на рис. 1,б, такой связи не показывают.

Корреляционное поле характеризует вид связи между *x* и *y*. По форме поля можно ориентировочно судить о форме графика, характеризующего прямолинейную или криволинейную зависимости. Даже для вполне выраженной формы корреляционного поля вследствие статистического характера связи исследуемого явления одно значение *x* может иметь несколько значений *y*. Если на корреляционном поле усреднить точки, т.е. для каждого значения x_i определить среднеарифметическое значение функции y_i и соединить точки $\bar{y}_i$, то можно будет получить ломаную линию, называемую экспериментальной регрессионной зависимостью (линией). Наличие ломаной линии объясняется погрешностями измерений, недостаточным количеством измерений, физической сущностью исследуемого явления и др. Если на корреляционном поле провести плавную линию между $\overline{y}_i$, которая равноудалена от усредненных, то получится новая теоретическая регрессионная зависимость – линия АБ (см. рис.1).

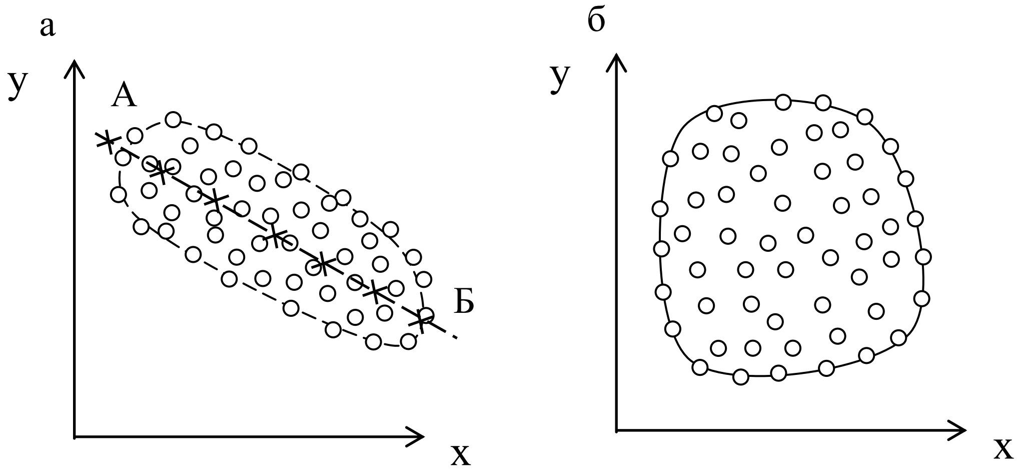

Рис. 1. Корреляционное поле:
а – при наличии коэффициента связи между переменными *x* и *y*;
б – при отсутствии корреляции

Различают однофакторные и многофакторные регрессионные зависимости. При *однофакторном эксперименте* опыты проводятся последовательно сериями, в которых меняется один из факторов, другие остаются постоянными. При *многофакторном эксперименте* опыты планируются сериями, однако значения всех факторов изменяются от опыта к опыту.

Однофакторный эксперимент имеет наглядность представления результатов, но требует большого количества опытов. Время и затраты труда на проведение многофакторного эксперимента значительно сокращаются, снижаются погрешности эксперимента.

При планировании эксперимента могут решаться следующие задачи:

- интерполяционная – построение поверхности отклика в пространстве изменения факторов, иначе – выявление степени влияния каждого фактора на функцию отклика;
- оптимизационная – определение наилучшего сочетания факторов, обеспечивающих экстремальное значение функции отклика.

Однофакторная регрессия при парной зависимости может быть аппроксимирована прямой линией, параболой, гиперболой, логарифмической, степенной или показательной функцией, полиномом и др.

Двухфакторное поле можно аппроксимировать плоскостью, параболоидом второго порядка, гиперболоидом. Для переменных факторов связь может быть установлена с помощью n-мерного пространства уравнениями второго порядка:

$$y = b_0 + \sum_{i}^{n} b_i x_i + \sum_{i}^{n} b_{ij} x_i x_j + \sum_{i}^{n} b_{ii} x_i^2 , \tag{1}$$

где y – функция многофакторных переменных;

x_i – независимые факторы;

b_i – коэффициенты регрессии, характеризующие влияние фактора x_i на функцию цели;

b_{ij} – коэффициенты, характеризующие двойное влияние факторов x_i и x_j на функцию цели.

2. РАСЧЕТ РЕГРЕССИОННОЙ ЗАВИСИМОСТИ ДЛЯ ОДНОФАКТОРНОГО ЭКСПЕРИМЕНТА

2.1. СТАТИСТИЧЕСКАЯ ОБРАБОТКА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

Обработка экспериментальных данных [4] имеет значение для исключения воздействия на функцию отклика внешних переменных (шумового поля). Точность результата зависит от числа n измеренных образцов и проведенных опытов. При статистической обработке экспериментальных данных, удовлетворяющих нормальному распределению, рекомендуется принимать $n \geq 27$.

Статистические выводы, определенные по свойствам ограниченного числа опытов n исследуемого процесса, распространяются на всю совокупность свойств исследуемого процесса. Среднее арифметическое $\bar{y}$ и среднее квадратическое отклонение S, измеренные для выборки из n опытов, рассматриваются как приближенные оценки параметров генеральной совокупности соответственно математического ожидания μ и дисперсии σ^2: $\mu \approx \bar{y}$, $\sigma^2 \approx S^2$.

Среднее арифметическое определяется по формуле

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i , \tag{2}$$

где n – число опытов в выборке.

Среднее квадратическое отклонение (дисперсия выборки) характеризует рассеивание (вариацию) случайных величин y_i ($i = \overline{1, n}$).

Величина S^2 определяется по формулам

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2 ; \tag{3}$$

$$S^2 = \frac{\sum_{i=1}^{n} y_i^2}{n-1} = \bar{y}^2 \frac{n}{n-1} . \qquad (4)$$

Для оценки изменчивости случайных величин используют вариационный коэффициент V, % :

$$V = 100\ S/\bar{y} . \qquad (5)$$

Ошибка среднего арифметического определяется по формуле

$$m = \pm S/\sqrt{n} . \qquad (6)$$

При экспериментальных измерениях значение $\bar{y}_i$ практически всегда будет находиться в интервале Δy, если удовлетворяется условие

$$\Delta y = y_i \pm 3m . \qquad (7)$$

Интервал Δy называют *доверительным*. Доверительный интервал для математического ожидания μ определяется по формуле

$$\bar{y} - t\,m \le \mu < \bar{y} + t\,m , \qquad (8)$$

где t – критическое значение критерия Стъюдента, принимается по табл.1.

Вероятность P нахождения математического ожидания μ генеральной совокупности в доверительном интервале (8) называют ***доверительной***. Пределы $\bar{y} \pm 3m$, соответствующие доверительной вероятности, называют *доверительными границами*.

Уровень значимости q = 1 – P называют *вероятностью ошибки,* которой можно пренебречь в данном экспериментальном исследовании. Обычно при исследованиях принимают P в пределах от 0,8 до 0,95. Доверительная вероятность P = 0,95 означает, что в интервале (8), определенного по выборке из n наблюдений (опытов), находится μ с вероятностью 95%, т.е. вероятность ошибки составляет 5%.

2.2. МЕТОД ИСКЛЮЧЕНИЯ ГРУБЫХ ОШИБОК ЭКСПЕРИМЕНТА

Грубые ошибки происходят при экспериментальных исследованиях из-за нарушений условий эксперимента ошибок при измерении и обработке результатов эксперимента [4]. Измеренное значение y_i^*, значительно отклоняющееся от $\bar{y}$, выбраковывают как «грубую ошибку» на основании сравнения.

Вначале рассчитывают характеристики выборки $\bar{y}$ и S^2. Определяют расчетный критерий Стъюдента как абсолютную величину разности, выраженную в долях S, по формуле

$$t_p = \left|\bar{y} - y_i^*\right| / S . \qquad (9)$$

Затем сравнивают значения критерия Стъюдента расчетного t_p и табличного t (см. табл.1). Если $t_p \geq t$, то с вероятностью вывода, соответствующей P табличному, можно считать, что отклоняющееся значение y_i^* является грубой ошибкой (содержит грубую ошибку). Это значение исключают из дальнейших расчетов и обработки результатов измерений. Если $t_p < t$, то достаточных оснований для исключения отклоняющегося значения y_i^* из выборки не имеется.

Таблица 1

Значения критерия Стъюдента t

Число степеней свободы	P				
	0,80	0,90	0,95	0,98	0,99
1	3,08	6,31	12,70	31,8	67,3
2	1,89	2,92	4,30	6,96	9,92
3	1,64	2,35	3,18	4,54	5,84
4	1,53	2,13	2,78	3,75	4,60
5	1,48	2,02	2,57	3,36	4,03
6	1,44	1,94	2,45	3,14	3,71
7	1,42	1,89	2,36	3,00	3,50
8	1,40	1,86	2,31	2,90	3,36
9	1,38	1,83	2,26	2,82	3,25
10	1,37	1,81	2,23	2,76	3,17
11	1,36	1,80	2,20	2,72	3,11
12	1,36	1,78	2,18	2,68	3,05
13	1,35	1,77	2,16	2,65	3,01
14	1,34	1,76	2,14	2,62	2,98
15	1,34	1,75	2,13	2,60	2,95
16	1,34	1,75	2,12	2,58	2,92
17	1,33	1,74	2,11	2,57	2,90
18	1,33	1,73	2,10	2,55	2,88
19	1,33	1,73	2,09	2,54	2,86
20	1,33	1,72	2,09	2,53	2,85
30	1,31	1,70	2,04	2,46	2,75
∞	1,28	1,64	1,96	2,33	2,58

2.3. ОЦЕНКА ВОСПРОИЗВОДИМОСТИ ЭКСПЕРИМЕНТАЛЬНЫХ ИЗМЕРЕНИЙ

При оценке воспроизводимости проверяют достоверность эксперимента. Если проверкой установлено, что эксперимент воспроизводим, производится установление функциональной зависимости функции отклика от факторов. Для экспериментов, достоверность которых не подтверждается при оценке воспроизводимости, принимается решение по изменению лабораторной схемы, применению других методов измерения или по прекращению исследований по этому направлению.

Рассмотрим порядок оценки воспроизводимости экспериментов. Проводят 2…4 серии параллельных опытов в принятой области изменения факторов. Результаты измерения параллельных опытов, проводимых при одинаковых факторах, сводят в табл. 2.

Таблица 2

Экспериментальные данные для оценки воспроизводимости эксперимента

Номер серии опыта	Результаты измерений параллельных опытов	Среднее арифметическое значение функции отклика $\bar{y}_j$	Дисперсия выборки S_j^2
1	y_{11}, y_{12}, …….. , y_{1k}	$\bar{y}_1$	S_1^2
2	y_{21}, y_{22}, …….. , y_{2k}	$\bar{y}_2$	S_2^2
.	……..…………….	.	.
j	y_{j1}, y_{j2}, …….. , y_{jk}	$\bar{y}_j$	S_j^2
.	…………………….	.	.
n	y_{n1}, y_{n2}, …….. , y_{nk}	$\bar{y}_n$	S_n^2

Число параллельных опытов n в однофакторных экспериментах определяют с учетом вариационного коэффициента по формуле

$$n = 100\, V\, t/q, \tag{10}$$

где V – вариационный коэффициент, определяется по формуле (5);

t – критерий Стъюдента, принимается по табл. 1;

q – уровень значимости, который принимают равным 5 или 10.

Обычно при расчетах воспроизводимости используют те же данные, получение которых требуется по эксперименту. Например, изучают влияние внутреннего давления (фактор *x*) на собственную частоту колебаний (функция отклика *y*) трубопровода. Номера серии опытов в этом случае соответствуют фиксированным уровням давления, при которых произведены измерения собственных частот колебаний трубопровода.

По найденным значениям дисперсии выборки рассчитывают критерий Кохрена по формуле

$$G_p = (S_j^2)_{max} \Big/ \sum_{j=1}^{n} S_j^2 , \qquad (11)$$

где S^2_{jmax} – максимальное значение из рассчитанных дисперсий в сериях опытов;

n – общее количество оценок дисперсии.

Затем производят сравнение рассчитанного значения критерия Кохрена G_p с табличным значением G (табл. 3). Если $G_p \le G$, то опыты считают воспроизводимыми, а оценки дисперсий однородными. При $G_p > G$ эксперимент следует прекратить или изменить условия его проведения.

Таблица 3

Значения критерия Кохрена G

Номер серии опыта n	Число степеней свободы f = n -1							
	1	2	3	4	5	6	7	8
p = 0,90								
2	0,9999	0,9950	0,9794	0,9586	0,9373	0,9172	0,8988	0,8823
3	0,9933	0,9423	0,8831	0,8335	0,7933	0,7606	0,7335	0,7107
4	0,9676	0,8643	0,7814	0,7212	0,7661	0,6410	0,6129	0,5897
5	0,9279	0,7885	0,6957	0,6329	0,5875	0,5531	0,5259	0,5037
6	0,8828	0,7218	0,6258	0,5635	0,5195	0,4866	0,4608	0,4401
7	0,8376	0,6644	0,5685	0,5080	0,4659	0,3932	0,4105	0,3911
8	0,7945	0,6152	0,5209	0,4627	0,4226	0,3592	0,3704	0,3522
9	0,7544	0,5727	0,4810	0,4251	0,3870	0,3308	0,3378	0,3207
10	0,7175	0,5358	0,4469	0,3934	0,3572	0,2861	0,3106	0,2945
12	0,6528	0,4751	0,3919	0,2882	0,3099	0,2386	0,2680	0,2535
15	0,5747	0,4069	0,3317	0,4069	0,2593	0,1877	0,2228	0,2104
20	0,4799	0,3297	0,2654	0,2288	0,2048	0,4347	0,1748	0,1646
p = 0,95								
2	0,999	0,975	0,939	0,906	0,877	0,853	0,833	0,816
3	0,967	0,871	0,798	0,746	0,707	0,677	0,653	0,633
4	0,907	0,768	0,684	0,629	0,590	0,560	0,637	0,518
5	0,841	0,684	0,598	0,544	0,507	0,478	0,456	0,439
6	0,781	0,616	0,532	0,480	0,445	0,418	0,398	0,382
7	0,727	0,561	0,480	0,431	0,397	0,373	0,354	0,338
8	0,680	0,516	0,438	0,391	0,360	0,336	0,319	0,304
9	0,639	0,478	0,403	0,358	0,329	0,307	0,290	0,277
10	0,602	0,445	0,373	0,331	0,303	0,382	0,267	0,254
12	0,541	0,392	0,326	0,288	0,262	0,244	0,230	0,219
15	0,471	0,335	0,276	0,242	0,220	0,203	0,191	0,182
20	0,389	0,271	0,221	0,192	0,174	0,160	0,150	0,142

2.4. РАСЧЕТ ЛИНЕЙНОЙ РЕГРЕССИВНОЙ ЗАВИСИМОСТИ ФУНКЦИИ ОТ ФАКТОРА

Расчет регрессивной зависимости функции *y* от фактора *x* рассмотрим на примере.

Исходные данные в виде статистического ряда парных измерений приведены в табл. 4. Из выборки исходных данных исключены грубые ошибки измерений в соответствии с подразд. 2.2. Исходные данные проверены по критерию Кохрена (см. подразд. 2.3) и удовлетворяют условиям воспроизводимости эксперимента.

Таблица 4

Статический ряд парных измерений

x	2	6	10	14	18	22	26	27
y	4,01	12,01	20,18	28,09	39,95	47,9	55,85	58,93
x	28	32	36	40	44	45	46	50
y	58,71	72,59	83,8	91,22	101,07	102,9	106,4	116,69
x	54	72	76	77	78	82	86	104
y	131,9	178,49	190,2	192,2	195,09	207,19	222	261

Производим промежуточные вычисления для уравнения линейной регрессии, результаты расчетов сводим в табл. 5.

Определяем среднее арифметическое значение для факторов x_i по формуле (2)

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{24}\cdot 1075 = 44{,}791.$$

Коэффициент линейной корреляции r может принимать значения от –1 до +1. При r = 0 линейная корреляционная связь между *x* и *y* отсутствует, при $r = \pm 1$ существует строгая функциональная линейная связь *x* и *y*. Если $|r| \geq 0{,}5$, то принято считать, что между фактором *x* и функцией отклика *y* имеется линейная корреляционная связь.

Вывод. При r = 0,998 между фактором *x* и функцией отклика *y* имеется линейная корреляционная связь, близкая к функциональной.

Определяем среднее арифметическое значение функции отклика по формуле (2)

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{24}\cdot 2578{,}37 = 107{,}43.$$

Определяем коэффициент линейной корреляции по формуле

$$r = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{\left[\left(n\sum x_i^2 - (\sum x_i)^2\right]\left[n\sum y_i^2 - (\sum y_i)^2\right]}} = \frac{24\cdot 163601{,}17 - 1075\cdot 2578{,}37}{1251903{,}14} = 0{,}998.$$

Таблица 5

Значения промежуточных величин для уравнения регрессии

x_i	y_i	x_i^2	y_i^2	$x_i\ y_i$
2	4,01	4	16,08	8,02
6	12,01	36	144,24	72,06
10	20,18	100	407,23	201,80
14	28,09	196	789,05	393,26
18	39,95	324	1596,00	719,10
22	47,90	484	2294,41	1053,80
26	55,85	676	3119,22	1452,10
27	58,93	729	3472,74	1591,11
28	58,71	784	3446,86	1643,88
32	72,59	1024	5269,31	2322,88
36	83,80	1296	7022,44	3016,8
40	91,22	1600	8321,09	3648,8
44	101,07	1936	10215,14	4447,08
45	102,90	2025	10588,41	4630,50
46	106,40	2116	11320,96	4894,40
50	116,69	2500	13616,56	5834,50
54	131,90	2916	17397,61	7122,60
72	178,49	5184	31858,68	12851,28
76	190,20	5776	36176,04	14455,20
77	192,20	5929	36940,84	14799,40
78	195,09	6084	38060,11	15217,02
82	207,19	6724	42927,70	16989,58
86	222	7396	49284	19092
104	261	10816	68121	27144
$\Sigma x_i = 1075$	$\Sigma y_i = 2578{,}37$	$\Sigma x_i^2 = 66655$	$\Sigma y_i^2 = 402405{,}727$	$\Sigma(x_i\ y_i) = 163601{,}17$

Определяем значение постоянных коэффициентов a, b уравнения линейной регрессии $y = a + bx$:

тангенс угла наклона линии регрессии

$$b = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2} = \frac{24\cdot 163601{,}17 - 163601{,}37}{24\cdot 66655 - 1075^2} = 8{,}473;$$

значение свободного члена

$$a = \bar{y} - b\bar{x} = \frac{\sum y_i}{n} - \frac{b\sum x_i}{n} = \frac{2578{,}37}{24} - \frac{8{,}47\cdot 1075}{24} = -272{,}089.$$

После подстановки значений коэффициентов a и b уравнение линейной регрессии примет вид: $y = 8{,}473\,x - 272{,}089.$

Графическая зависимость y от x приведена на рис. 2.

Рис. 2. График зависимости *y* от фактора *x*

Определяем коэффициент детерминации, который отражает влияние фактора x на функцию y. $К_Д = r^2 = 0{,}998^2 = 0{,}997$.

При $К_Д = 0{,}997$ функция y на 99,7 % зависит от фактора x и лишь на 0,03 % от других причин.

При $r < 0{,}5$ линейная корреляционная связь между фактором x и функцией отклика y считают неудовлетворительной. В этом случае степень влияния переменных друг на друга может изменяться непрямолинейно. Для определения связи между x и y в виде плавных кривых линий применяются методы подбора эмпирических формул.

3. ПОДБОР ЭМПИРИЧЕСКИХ ФОРМУЛ ДЛЯ ОДНОФАКТОРНОГО ЭКСПЕРИМЕНТА ПО МЕТОДУ СРЕДНИХ КВАДРАТОВ

При подборе эмпирических формул [3] применяют полиномы вида:

$$y = A_0 + A_1 x + A_2 x^2 + \ldots + A_n x^n, \tag{12}$$

где $A_0, A_1, \ldots, A_n$ – постоянные коэффициенты.

Полиномами можно аппроксимировать любые результаты измерений, если они графически выражаются непрерывными функциями. Значения постоянных коэффициентов A_i определяются по экспериментальным данным x и y однофакторного эксперимента.

Метод средних квадратов основан на следующем положении. По экспериментальным данным можно построить несколько плавных кривых. Наилучшей будет та кривая, у которой разностные отклонения Е оказываются наименьшими, т.е. $\Sigma E_i \approx 0$. Порядок расчета коэффициентов полинома сводится к следующему. Определяется число членов ряда (12). В принятое выражение последовательно подставляют координаты *x* и *y* нескольких n экспериментальных точек и получают систему из n уравнений. Каждое уравнение приравнивают соответствующему отклонению:

$$\begin{aligned} &A_0 + A_1 x_1 + A_2 x_1^2 + \ldots + A_n x_1^n - y_1 = E_1 ; \\ &A_0 + A_1 x_2 + A_2 x_2^2 + \ldots + A_n x_2^n - y_2 = E_2 ; \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &A_0 + A_1 x_n + A_2 x_n^2 + \ldots + A_n x_n^n - y_n = E_n . \end{aligned} \qquad (13)$$

Число точек, т.е. число уравнений, должно быть не меньше числа коэффициентов А.

Разбивают систему начальных уравнений (13) последовательно сверху вниз на группы, число которых должно быть равно количеству коэффициентов A_0. В каждой группе складывают уравнения и получают новую систему уравнений, равную количеству групп (обычно 2…3). Решая систему, вычисляют коэффициенты A_i.

Метод средних квадратов обладает высокой точностью, если число точек достаточно велико (не менее 3). Однако степень точности можно повысить, если начальные условия сгруппировать по 2…3 варианта и вычислить для каждого варианта эмпирическую формулу. Предпочтение следует отдать той формуле, у которой $\Sigma E^2 = \min$.

Пусть, например, выполнено восемь измерений (табл. 6).

Таблица 6

Экспериментальные данные

x	29	30	31	32	33	34	35	36
$y_э$	63,50	65,82	69,72	72,92	74,92	78,96	79,90	83,60

Для подбора эмпирической формулы используем полином

$$y = A_0 + A_1 x + A_2 x^2 . \qquad (14)$$

Полином решаем для трех коэффициентов. Составляем систему начальных уравнений для расчета коэффициентов полинома

$$\begin{cases} \begin{cases} A_0 + 29A_1 + 841A_2 - 63{,}50 = E_1 \\ A_0 + 30A_1 + 900A_2 - 65{,}82 = E_2 \\ A_0 + 31A_1 + 961A_2 - 69{,}72 = E_3 \end{cases} \\ \begin{cases} A_0 + 32A_1 + 1024A_2 - 72{,}59 = E_4 \\ A_0 + 33A_1 + 1089A_2 - 74{,}92 = E_5 \\ A_0 + 34A_1 + 1156A_2 - 78{,}96 = E_6 \end{cases} \\ \begin{cases} A_0 + 35A_2 + 1225A_2 - 79{,}90 = E_7 \\ A_0 + 36A_1 + 1296A_2 - 83{,}60 = E_8 \end{cases} \end{cases}, \tag{15}$$

где E_i – разность между экспериментальными и расчетными данными.

Производим группировку начальных уравнений по числу коэффициентов уравнения (14)

$$\begin{cases} 3A_0 + 90A_1 + 2702A_2 = 199{,}04 \\ 3A_0 + 99A_1 + 3269A_2 = 226{,}47 \\ 2A_0 + 71A_1 + 2521A_2 = 163{,}50 \end{cases}. \tag{16}$$

Определяем коэффициенты A_0, A_1, A_2 по уравнениям группировки (16)

$$A = \begin{pmatrix} 3 & 90 & 2702 \\ 3 & 99 & 3269 \\ 2 & 71 & 2521 \end{pmatrix}; \qquad A_0 = \begin{pmatrix} 99{,}04 & 90 & 2702 \\ 226{,}47 & 99 & 3269 \\ 163{,}50 & 71 & 2521 \end{pmatrix}$$

$$A_1 = \begin{pmatrix} 3 & 199{,}04 & 2702 \\ 3 & 226{,}47 & 3269 \\ 2 & 163{,}50 & 2521 \end{pmatrix}; \qquad A_2 = \begin{pmatrix} 3 & 90 & 199{,}04 \\ 3 & 99 & 226{,}47 \\ 2 & 71 & 163{,}50 \end{pmatrix};$$

$$|A| = 720; \quad |A_0| = -90690; \quad |A_1| = 6819; \quad |A_2| = -411{,}45;$$

$$A_0 = \frac{|A_0|}{|A|} = -125{,}96; \quad A_1 = \frac{|A_1|}{|A|} = 9{,}47; \quad A_2 = \frac{|A2|}{|A|} = -0{,}571.$$

После подстановки значений A_0, A_1, A_2 в уравнение (14) получим эмпирическую формулу в виде:

$$y = -125{,}957 + 9{,}471\,x - 0{,}571\,x^2. \tag{17}$$

Значения функции y_p, рассчитанные по уравнению (17) для исходных значений фактора *x*, приведены в табл. 7.

Таблица 7

Расчетные значения y_p

x	29	30	31	32	33	34	35	36
y_p	61,59	65,80	70,10	71,12	75,47	77,56	80,92	83,96

Строим графики зависимости функций от фактора по данным табл. 6 и 7 (рис. 3).

Определяем величины E_i (разностные отклонения), т.е. разность между экспериментальной и расчетной ординатами.

$E_1 = 63{,}5 - 61{,}59 = 1{,}91;$ $E_2 = 65{,}82 - 65{,}80 = 0{,}02;$
$E_3 = 69{,}72 - 70{,}10 = -0{,}38;$ $E_4 = 72{,}59 - 71{,}72 = 0{,}87;$
$E_5 = 72{,}92 - 75{,}47 = -0{,}55;$ $E_6 = 78{,}96 - 77{,}58 = 1{,}37;$
$E_7 = 77{,}90 - 80{,}92 = -3{,}02;$ $E_8 = 83{,}60 - 83{,}96 = -0{,}36.$

Определяем сумму разностных отклонений

$$\sum_{i=1}^{8} E_i = 1{,}91 + 0{,}02 - 0{,}38 + 0{,}87 - 0{,}55 + 1{,}37 - 3{,}02 - 0{,}36 = -0{,}14.$$

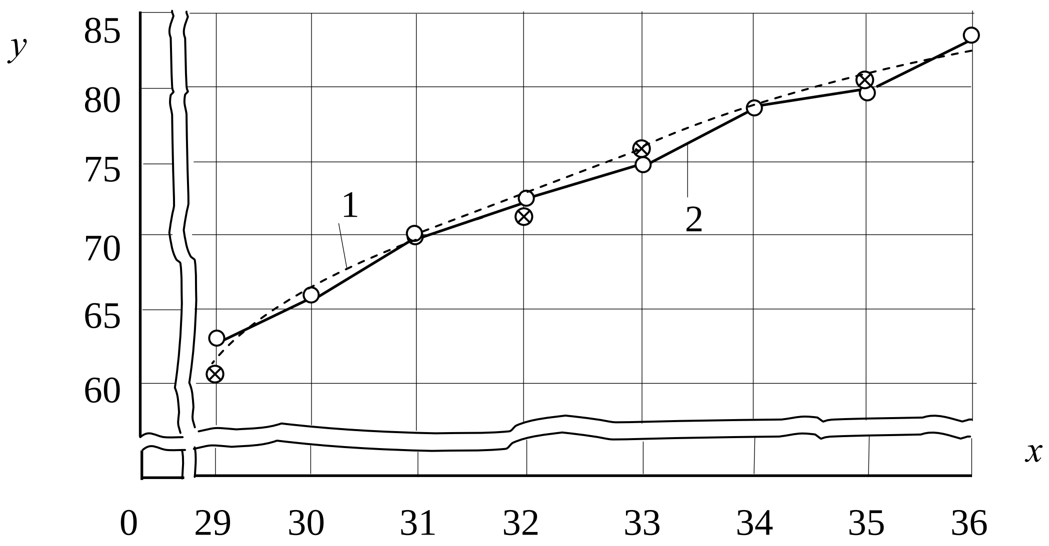

Рис. 3. Зависимость функции *y* от фактора *x*:
1 – экспериментальная; 2 – расчетная

Поскольку $\Sigma E_i \approx 0$, принимаем, что эмпирическая формула (17) хорошо аппроксимирует экспериментальные данные.

4. МОДЕЛЬ И МЕТОД ПОЛНОГО ФАКТОРНОГО ЭКСПЕРИМЕНТА

Полные факторные эксперименты (ПФЭ) используют в задачах, когда нужно оценить линейное влияние нескольких факторов и эффект совместного воздействия факторов на функцию. В отличие от классического однофакторного эксперимента в ПФЭ варьируют одновременно всеми факторами на двух уровнях.

К планированию ПФЭ приступают после того, когда доказана воспроизводимость опытов по методике, рассмотренной в разд. 2. Проверку воспроизводимости можно выполнить также при обработке данных факторного эксперимента.

Метод ПФЭ дает возможность получить экспериментально-статистическую модель для некоторой локальной области факторного пространства, лежащей в окрестности выбранной точки с координатами (z_1^o, z_2^o,…, z_k). Эта выбранная точка называется *центром плана*, иногда ее называют *основным уровнем*.

При планировании по схеме ПФЭ реализуются все возможные комбинации из k факторов на двух выбранных уровнях. Необходимое количество опытов при этом определяется по формуле

$$N = 2^k . \tag{18}$$

Если k =3, то мы имеем ПФЭ 2^3, или полный трехфакторный эксперимент. Уровни факторов представляют собой границы исследуемой области.

Математическую модель ПФЭ рассмотрим на примере [5]. Исследуем влияние на выход целлюлозы y, %, трех факторов: температура z_1 в диапазоне 140…180 ^{0}С; давление z_2 в диапазоне 0,8…1,2 МПа; время пребывания сырья в в реакционной зоне автоклава z_3 в диапазоне 30…90 мин.

Производим расчет центра плана эксперимента. Основной уровень z_j^o по оси j-го фактора определяем по формуле

$$z_j^o = \left(z_j^{max} + z_j^{min}\right)/2; \qquad j = \overline{1,\dots,k}, \tag{19}$$

где z_j^{max}, z_j^{min} – максимальный и минимальный уровни j-го фактора;

k = 3 – число факторов.

По формуле (19) для рассматриваемого примера имеем:

$$z_1^o = \left(z_1^{max} + z_1^{min}\right)/2 = (180 + 140)/2 = 160\ ^0\text{C};$$

$$z_2^o = \left(z_2^{max} + z_2^{min}\right)/2 = (1{,}2 + 0{,}8)/2 = 1\ \text{МПа};$$

$$z_3^o = \left(z_3^{max} + z_3^{min}\right)/2 = (90 + 30)/2 = 60\ \text{мин}.$$

Интервалы варьирования факторов в эксперименте определяем по формуле

$$\Delta z_j = \left(z_j^{max} - z_j^{min}\right)/2; \qquad j = \overline{1,\dots,k}. \tag{20}$$

Для рассматриваемого примера имеем:

$$\Delta z_1 = \left(z_1^{max} - z_1^{min}\right)/2 = (180 - 140)/2 = 20\ ^0\text{C};$$

$$\Delta z_2 = \left(z_2^{max} - z_2^{min}\right)/2 = (1{,}2 - 0{,}8)/2 = 0{,}2 \text{ МПа};$$

$$\Delta z_3 = \left(z_3^{max} - z_3^{min}\right)/2 = (90 - 30)/2 = 30 \text{ мин}.$$

Переходим к новой безразмерной системе координат по формуле

$$x_j^{max,min} = \left(z_j^{max,min} - z_j^o\right)/\Delta z_j; \qquad j = \overline{1,\dots,k}. \tag{21}$$

Для рассматриваемого примера имеем:

$$x_1^{max} = \left(z_1^{max} - z_1^o\right)/\Delta z_1 = (180 - 160)/20 = 1;$$

$$x_1^{min} = \left(z_1^{min} - z_1^o\right)/\Delta z_1 = (140 - 160)/20 = -1;$$

$$x_2^{max} = \left(z_2^{max} - z_2^o\right)/\Delta z_2 = (1{,}2 - 1{,}0)/0{,}2 = 1;$$

$$x_2^{min} = \left(z_2^{min} - z_2^o\right)/\Delta z_2 = (0{,}8 - 1{,}0)/0{,}2 = -1;$$

$$x_3^{max} = \left(z_3^{max} - z_1^o\right)/\Delta z_3 = (90 - 60)/30 = 1;$$

$$x_3^{min} = \left(z_3^{min} - z_3^o\right)/\Delta z_3 = (30 - 60)/30 = -1.$$

В безразмерной системе координат верхний уровень для всех факторов равен +1, нижний уровень равен –1, координаты центра плана равны нулю и совпадают с началом новой системы координат. В рассматриваемом примере k = 3, тогда по (18) $N = 2^k = 2^3 = 8$.

План проведения эксперимента (матрица планирования) записывается в виде таблицы для факторов в безразмерной системе координат (их называют кодированные переменные), а также с указанием факторов в натуральном масштабе (табл. 8). Здесь же в матрицу планирования введен столбец так называемой фиктивной переменной $x_o = +1$.

Кодированный план геометрически можно представит в виде куба, восемь вершин которого представляют собой восемь экспериментальных точек. Центр куба соответствует точке основного уровня и является началом новой системы координат.

Пользуясь планом табл. 8 математическое описание процесса будем искать в виде линейного уравнения регрессии

$$y = b_0 + x_1 b_1 + x_2 b_2 + x_3 b_3. \tag{22}$$

Следует заметить, что матрица полного факторного эксперимента ортогональна. Любой коэффициент уравнения регрессии b_j определяется скалярным произведением столбца y на соответствующий столбец x_j, деленным на число опытов в матрице планирования:

$$b_j = \frac{1}{N}\sum_{i=1}^{N}(x_{ji}\, y_i). \tag{23}$$

Таблица 8

Матрица планирования эксперимента

Номер опыта	x_0	x_1	x_2	x_3	z_1, ^{0}C	z_2, МПа	z_3, мин	y, %
1	1	–1	–1	–1	140	0,8	30	37
2	1	1	–1	–1	180	0,8	30	42
3	1	–1	1	–1	140	1,2	30	41
4	1	1	1	–1	180	1,2	30	32
5	1	–1	–1	1	140	0,8	90	46
6	1	1	–1	1	180	0,8	90	41
7	1	–1	1	1	140	1,2	90	39
8	1	1	1	1	180	1,2	90	40

Расчет скалярных произведений столбцов $x_0 x_i$ на столбец y (см. табл. 8) производим по формуле

$$b'_j = \sum_{i=1}^{N} (x_{ji}\, y_i); \qquad j = 0,1,2,3\,. \tag{24}$$

Для исследуемого примера имеем:

$$b'_0 = \begin{vmatrix} 1\\1\\1\\1\\1\\1\\1\\1 \end{vmatrix} \times \begin{vmatrix} 37\\42\\41\\32\\46\\41\\39\\40 \end{vmatrix} = 318;\; b'_1 = \begin{vmatrix} -1\\1\\-1\\1\\-1\\1\\-1\\1 \end{vmatrix} \times \begin{vmatrix} 37\\42\\41\\32\\46\\41\\39\\40 \end{vmatrix} = -8;\; b'_2 = \begin{vmatrix} -1\\-1\\1\\1\\-1\\-1\\1\\1 \end{vmatrix} \times \begin{vmatrix} 37\\42\\41\\32\\46\\41\\39\\40 \end{vmatrix} = -14;\; b'_3 = \begin{vmatrix} -1\\-1\\-1\\-1\\1\\1\\1\\1 \end{vmatrix} \times \begin{vmatrix} 37\\42\\41\\32\\46\\41\\39\\40 \end{vmatrix} = 14.$$

Цифровые значения коэффициентов b_j, определенные по (23) равны:

$$b_0 = \frac{318}{8} = 39{,}75\,; \quad b_1 = \frac{-8}{8} = -1; \quad b_2 = \frac{-14}{8} = -1{,}75; \quad b_3 = \frac{14}{8} = 1{,}75.$$

Уравнение линейной регрессии для рассматриваемого примера запишется в виде:

$$y = 39{,}75 - x_1 - 1{,}75x_2 + 1{,}75x_3. \tag{25}$$

Для определения влияния эффектов парных и тройного взаимодействий факторов используем более полное уравнение линейной регрессии

$$y = b_0 + x_1 b_1 + x_2 b_2 + x_3 b_3 + b_{12}\, x_1 x_2 + b_{13}\, x_1 x_3 + b_{23}\, x_2 x_3 + b_{123}\, x_1 x_2 x_3. \tag{26}$$

Составляем уравнение регрессии с учетом парных и тройного взаимодействий факторов. Расширенная матрица планирования приведена в табл. 9.

Таблица 9

Расширенная матрица планирования

Номер опыта	x_0	x_1	x_2	x_3	z_1, ^{0}C	z_2, МПа	z_3, мин	y, %	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$
1	1	–1	–1	–1	140	0,8	30	37	1	1	1	–1
2	1	1	–1	–1	180	0,8	30	42	–1	–1	1	1
3	1	–1	1	–1	140	1,2	30	41	–1	1	–1	1
4	1	1	1	–1	180	1,2	30	32	1	–1	–1	–1
5	1	–1	–1	1	140	0,8	90	46	1	–1	–1	1
6	1	1	–1	1	180	0,8	90	41	–1	1	–1	–1
7	1	–1	1	1	140	1,2	90	39	–1	–1	1	–1
8	1	1	1	1	180	1,2	90	40	1	1	1	1

Эффекты взаимодействия определяются аналогично линейным эффектам по формулам (25) и (26). Для рассматриваемого примера получим:

$$b'_{12} = \begin{array}{c} x_1x_2 \\ \left| \begin{array}{c} 1 \\ -1 \\ -1 \\ 1 \\ 1 \\ -1 \\ -1 \\ 1 \end{array} \right| \end{array} \times \begin{array}{c} y \\ \left| \begin{array}{c} 37 \\ 42 \\ 41 \\ 32 \\ 46 \\ 41 \\ 39 \\ 40 \end{array} \right| \end{array} = -8; \qquad b'_{13} = \begin{array}{c} x_1x_3 \\ \left| \begin{array}{c} 1 \\ -1 \\ 1 \\ -1 \\ -1 \\ 1 \\ -1 \\ 1 \end{array} \right| \end{array} \times \begin{array}{c} y \\ \left| \begin{array}{c} 37 \\ 42 \\ 41 \\ 32 \\ 46 \\ 41 \\ 39 \\ 40 \end{array} \right| \end{array} = -0;$$

$$b'_{23} = \begin{array}{c} x_2x_3 \\ \left| \begin{array}{c} 1 \\ 1 \\ -1 \\ -1 \\ -1 \\ -1 \\ 1 \\ 1 \end{array} \right| \end{array} \times \begin{array}{c} y \\ \left| \begin{array}{c} 37 \\ 42 \\ 41 \\ 32 \\ 46 \\ 41 \\ 39 \\ 40 \end{array} \right| \end{array} = -2; \qquad b'_{123} = \begin{array}{c} x_2x_3x_3 \\ \left| \begin{array}{c} -1 \\ 1 \\ 1 \\ -1 \\ 1 \\ -1 \\ -1 \\ 1 \end{array} \right| \end{array} \times \begin{array}{c} y \\ \left| \begin{array}{c} 37 \\ 42 \\ 41 \\ 32 \\ 46 \\ 41 \\ 39 \\ 40 \end{array} \right| \end{array} = 2;$$

$b_{12} = -8/8 = -1$; $b_{13} = 0/8 = 0$; $b_{23} = -2/8 = -0{,}25$; $b_{123} = 20/8 = 2{,}5$.

Тогда уравнение (26) запишется в виде:

$$y = 39{,}75 - x_1 - 1{,}75x_2 + 1{,}75x_3 - 1x_1x_2 - 0{,}25\,x_2x_3 + 2{,}5\,x_1x_2x_3.$$

Некоторые из коэффициентов регрессии могут оказаться пренебрежимо малыми – незначительными. Чтобы установить, значим коэффициент или нет, необходимо вычислить оценку дисперсии S_b:

$$S_b = \sqrt{S^2/N} = \sqrt{0{,}29/8} = 0{,}19. \qquad (27)$$

где $N = 8$ – число опытов;

$S^2 = 0{,}29$ – дисперсия воспроизводимости эксперимента.

Расчет дисперсии воспроизводимости рассмотрен в разд. 2.

Следует отметить, что с помощью полного факторного эксперимента все коэффициенты определяются с одинаковой погрешностью.

Принято считать, что коэффициент регрессии значим, если выполнено условие

$$\left|b_{ij}\right| \geq S_b\, t, \qquad (28)$$

где t – значение критерия Стъюдента.

В противном случае коэффициент регрессии незначим, и соответствующий член следует исключить из уравнения.

Из табл.1 при P = 0,90 и f = 2 значение критерия Стъюдента t = 2,92.

Производим сравнение коэффициентов b_{ij} уравнения (26) с допускаемым $S_b\, t = 0{,}19 \cdot 2{,}92 = 0{,}5548$. Из сравнения следует, что коэффициенты b_1, b_2, b_3, b_{12} и b_{123} значимы.

Окончательный вид уравнения регрессии без пренебрежимо малых коэффициентов

$$\mathbf{y = 39{,}75 - x_1 - 1{,}75x_2 + 1{,}75x_3 - x_1x_2 + 2{,}5\,x_1x_2x_3.}$$

Делаем вывод о влиянии факторов на технологический процесс. Наиболее влиятельными являются факторы x_2, x_3 и эффект сочетания трех факторов $x_1x_2x_3$.

Производим проверку адекватности уравнения линейной регрессии с целью определить способность уравнения регрессии удовлетворительно характеризовать поверхность функции отклика *y*.

Определяем оценку дисперсии адекватности по формуле

$$S^2_{ад} = \frac{1}{N-B}\sum_{i=1}^{N}\left(y_i^э - y_i^p\right)^2, \qquad (30)$$

где N – число экспериментов;

B – число значимых коэффициентов уравнения линейной регрессии, включая свободный член;

$y_i^э$, y_i^p – соответственно экспериментальное и расчетное значения функции отклика i-го эксперимента, %.

Для рассматриваемого примера:

$$y_1^p = 39{,}75 + 1 + 1{,}75 - 1{,}75 - 1 - 2{,}5 = 37{,}25;$$

$y_2^p = 39{,}75 - 1 + 1{,}75 - 1{,}75 + 1 + 2{,}5 = 42{,}25;$
$y_3^p = 39{,}75 + 1 - 1{,}75 - 1{,}75 + 1 + 2{,}5 = 40{,}75;$
$y_4^p = 39{,}75 - 1 - 1{,}75 - 1{,}75 - 1 - 2{,}5 = 31{,}75;$
$y_5^p = 39{,}75 + 1 + 1{,}75 + 1{,}75 - 1 + 2{,}5 = 45{,}75;$
$y_6^p = 39{,}75 - 1 + 1{,}75 +1{,}75 + 1 - 2{,}5 = 40{,}25;$
$y_7^p = 39{,}75 + 1 - 1{,}75 + 1{,}75 + 1 - 2{,}5 = 39{,}25;$
$y_8^p = 39{,}75 - 1 - 1{,}75 + 1{,}75 - 1 + 2{,}5 = 40{,}25.$

Оценку дисперсии адекватности определяем по формуле (30)

$$S^2_{ад} = \frac{1}{8-6}\Big[(37-37{,}25)^2+(42-42{,}25)^2+(41-40{,}75)^2+(32-31{,}75)^2+$$
$$+(46-45{,}75)^2+(41-40{,}25)^2+(39-39{,}25)^2+(40-40{,}25)^2\Big]= 0{,}47.$$

Определяем расчетный критерий Фишера по формуле

$$F_р = S^2_{ад} \big/ S^2_в = 0{,}47/0{,}29 = 1{,}61\,. \qquad (31)$$

Выбираем табличное значение Фишера при P = 0,9; $f_1 = 2$; $f_2 = 8$ $F_т = 3{,}113$ (табл. 10), где $f_1 = N - B = 8 - 6 = 2$; $f_2 = N(k-1) = 8(2 - 1) = 8$; k = 2 – число серий измерения функции *y* в каждом из N опытов.

Таблица 10

Значения критерия Фишера F

Число степеней свободы f_2	Число степеней свободы f_1 (для числителя)					
	1	2	3	4	5	6
p = 0,90						
1	39,864	45,500	53,593	55,833	57,241	58,204
2	8,526	9,000	9,162	9,243	9,263	9,326
3	5,538	5,462	5,391	5,343	5,309	5,285
4	4,545	4,325	4,191	4,107	4,501	4,010
5	4,060	3,780	3,620	3,520	3,453	3,405
6	3,776	3,463	3,289	3,181	3,108	3,055
7	3,589	3,257	3,074	3,074	2,883	2,827
8	3,458	3,113	2,924	2,961	2,727	2,668
9	3,360	3,007	2,813	2,806	2,611	2,551
10	3,285	2,925	2,728	2,693	2,522	2,461
11	3,225	2,860	2,660	2,605	2,451	2,389
12	3,177	2,807	2,606	2,536	2,394	2,331
13	3,136	2,763	2,560	2,480	2,347	2,283
14	3,102	3,727	2,522	2,434	2,307	2,243
15	3,073	3,695	2,490	2,361	2,273	2,208
p = 0,95						

1	161,45	199,50	215,71	224,58	230,16	233,99
2	18,51	19,00	19,16	19,25	19,30	19,33
3	10,13	9,55	9,28	9,12	9,01	8,94
4	7,71	6,94	6,59	6,39	6,26	6,16
5	6,61	5,79	5,41	5,19	5,05	4,95
6	5,99	5,14	4,76	4,53	4,39	4,28
7	5,59	4,74	4,35	4,12	3,97	3,87
8	5,32	4,46	4,07	3,84	3,69	3,58
9	5,12	4,26	3,86	3,63	3,48	3,37
10	4,97	4,10	3,71	3,48	3,33	3,22
11	4,84	3,98	3,59	3,36	3,20	3,10
12	4,75	3,89	3,49	3,26	3,11	3,00
13	4,67	3,81	3,41	3,18	3,03	2,92
14	4,60	3,74	3,34	3,11	2,96	2,85
15	4,54	3,68	3,29	3,06	2,90	2,79

Проверяем адекватность уравнения регрессии. Уравнение регрессии удовлетворительно характеризует поверхность функции отклика, если удовлетворяет условию

$$F_р \leq F_т .$$

В рассматриваемом примере 1,61 < 3,113, следовательно уравнение (30) адекватно определяет исследуемую функцию технологического процесса.

ЛИТЕРАТУРА

1. Саутин С.Н. Планирование эксперимента в химии и химической технологии. Л.: Химия, 1975. – 48 с.

2. Шенк Х. Теория инженерного эксперимента. М.: Мир, 1972. – 382 с.

3. Основы научных исследований / Под ред. В.И. Крутова, В.В. Попова. М.: Высшая школа, 1989. – 400 с.

4. Румшинский Л.З. Математическая обработка результатов эксперимента (справочное пособие). М.: Наука, 1971. – 192 с.

5. Леонович А.А. Основы научных исследований в химической переработке древесины. Л.: ЛТА, 1982. – 55 с.