УДК 631.362

О. В. СИНЯЕВА аспирант, М. М. АБДУЕВ, канд. техн. наук
Харьковский национальный технический университет сельского хозяйства им. П. Василенко,
А. А. ЖУРАВСКИЙ, доктор философии
Украинский научно-исследовательский углехимический институт (УХИН), г. Харьков

МЕТОД ПОЛУЧЕНИЯ МАТЕМАТИЧЕСКОЙ МОДЕЛИ ПРОИЗВОЛЬНОГО ОБЪЕКТА

При проведении научного исследования не всегда возможно уложиться в рамки требований проведения планированного эксперимента. Авторами была разработана новая методика, позволяющая построить математическую модель изучаемого объекта, избежав при этом ограничений, накладываемых классическим методом планирования эксперимента.

Для облегчения работы авторами была разработана компьютерная программа, позволяющая существенно облегчить обработку результатов научных исследований и построение математической модели изучаемого объекта.

При проведенні наукового дослідження не завжди можливо вкластися у рамки вимог проведення планованого експерименту. Співавторами була розроблена нова методика, яка дозволяє побудувати математичну модель вивчаємого об'єкту, виключаючи при цьому обмеження, які накладаються класичним методом планування експерименту.

Для полегшення роботи авторами була розроблена комп'ютерна програма, яка дозволяє суттєво полегшити обробку результатів наукових досліджень та побудова математичної моделі вивчає мого об'єкту.

Введение

При проведении научных экспериментов для получения эмпирических зависимостей используются, как правило, метод планируемого эксперимента. Однако большая часть способов построения многофакторных моделей используют различные модификации планируемого эксперимента, которые не всегда могут быть применимы в условиях действующего производства, особенно тогда, когда наблюдаемые значения не сводятся в определенные разряды. Поэтому при нахождении уравнений, описывающих изменение какого-либо показателя, приходится в виде общего правила иметь дело с неравноотстоящими значениями исследуемых величин. Из-за этого обычные способы нахождения нужных зависимостей оказываются непригодными, и возникает необходимость в разработке таких приемов, которые бы оказались эффективными в данных условиях.

Основная часть

Разработанная методика [1] основывается на том, что сначала весь массив данных мы рассматриваем как функциональную зависимость какого-либо одного переменного, допустим, X_1, а остальные считаем постоянными. В итоге получаем зависимость типа

$$Y(X) = f(X_1)\ \Big|_{X_2, X_3, X_4, \ldots, X_n - const} \tag{1}$$

Затем проводим подобную операцию для переменных X_2, X_3,….X_n, в результате чего получаем целую систему уравнений:

$$\begin{cases} Y(X) = f(X_2)\Big|_{X_1, X_3, .X_4 .., X_n - const} \\ Y(X) = f(X_3)\Big|_{X_1, X_2, X_4, \ldots, X_n - const} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ Y(X) = f(X_n)\Big|_{X_1, X_2, X_3, \ldots, X_{n-1} - const} \end{cases} \tag{2}$$

При перемножении этих строк друг на друга получим итоговое уравнение:

$$Y^n(X) = f(X_1) * f(X_2) * f(X_3) * ... * f(X_n) \tag{3}$$

или

$$Y(X) = \left[f(X_1) * f(X_2) * f(X_3) * ... * f(X_n)\right]^{1/n} \tag{4}$$

В качестве функциональных зависимостей *f(X)* используется набор элементарных функций, коэффициенты которых определяются по методу наименьших квадратов. Для каждой переменной X_i подбирают функциональную зависимость по значению коэффициента корреляции, т.е. из всего многообразия элементарных функций, которые описывают изменение параметра $Y(X)=f(X_i)$, выбирается та функциональная зависимость, у которой коэффициент корреляции наибольший.

В принципе, уравнением (4) можно было бы и ограничиться для описания изменения параметра *Y* от переменных X_i. Однако для получения более точного описания мы провели еще одно уточнение полученной зависимости. Для этого проводится получение корреляционной зависимости между фактическими $Y_{факт}$ и полученными расчетными $Y_{расч}$ значениями параметра *Y*. Функциональная зависимость определяется в виде линейной функции вида:

$$Y_{факт} = E*Y_{расч} + D \tag{5}$$

Полученный при этом коэффициент корреляции будет являться общим коэффициентом корреляции и указывает на тесноту связи между фактическими и расчетными значениями параметра *Y*. В идеальном случае, когда коэффициент корреляции между фактическими и расчетными значениями параметра *Y* равен 1, коэффициент *E*=1, а *D*=0, В противном случае, отличие этих коэффициентов от указанных значений говорит о том, что не все факторы, влияющие на изменение искомого параметра *Y*, учтены [3]. Коэффициенты корреляции, полученные при расчете уравнений (1-2), являются частными коэффициентами и говорят о том, насколько изменение данной переменной X_i влияло на изменение параметра *Y* в исследуемом интервале, но никоим образом не говорит о влиянии самой переменной на параметр *Y*. Например, при определении зависимости показателя чистоты разделения зерновой смеси от расхода воздуха, угла наклона рассеивающей плоскости, частоты и амплитуды вибрации расход воздуха был постоянным. Естественно, что при обработке полученных результатов, коэффициент корреляции уравнения, описывающего изменение чистоты разделения зерновой смеси от расхода воздуха будет очень мал либо вообще равен нулю. Можем ли мы после этого утверждать, что насыпная чистота разделения зерновой смеси не зависит от расхода воздуха? Естественно, нет. Малое значение частного коэффициента корреляции говорит лишь о том, что изменение расхода воздуха в указанных пределах (а он, в данном случае, не изменялся) не оказало существенного влияния на изменение показателя чистоты разделения зерновой смеси.

На основе вышеприведенных рассуждений была разработана программа для компьютера, которая позволяла автоматически рассчитывать корреляционные зависимости и составлять статистическую модель работы исследуемого объекта и, опираясь на полученные зависимости, получать прогноз поведения объекта при изменении условий его работы. Однако при подготовке исходных данных для расчета статистической модели следует быть крайне осторожным и учитывать несколько моментов, которые могут существенным образом повлиять на конечный результат, а именно:

1) Необходимо тщательно проверять, чтобы вводимые значения были однозначны. Это означает, что одним и тем же параметрам работы объекта должны соответствовать одинаковые значения откликов. Несоблюдение этого правила приводит к резкому падению

точности полученных зависимостей. Поэтому при подготовке исходных данных необходимо тщательно их проверять на однозначность.

2) Необходимо синхронизировать вводимые значения. Это означает, что параметры, вводимые для построения искомых зависимостей, должны быть согласованы между собой во времени. В противном случае, у нас может появиться неоднозначность, о которой говорилось выше.

Для того, чтобы определить место разработанной методики в ряду уже известных, были проведены сопоставительные расчеты проведенного эксперимента, рассчитанного по одной из известных методик и предлагаемой. Полученные данные говорили о том, что сравнение фактических и расчетных данных, полученных разными способами примерно одинаковы, что говорит об эффективности вышеописанной методики.

Описанная выше методика выгодно отличается от уже известных тем, что для нее не требуется построение матрицы проведения эксперимента, как это имеет место при проведении планируемого эксперимента и ее можно использовать при оценке работы действующего промышленного объекта. Впрочем, это отнюдь не означает, что планируемый эксперимент можно снимать с вооружения исследователей, наоборот, его использование позволит учесть все мыслимые режимы объекта, что только повысит адекватность полученной модели. И только лишь в тех случаях, когда невозможно это сделать, можно прибегать к описанной методике, но при этом следует быть готовым к тому, что точность полученных зависимостей не будет абсолютной. Использование подобной методики позволит:

– Получить математическую модель работы объекта и определить направление влияния каждого из факторов на качество зерновой смеси и сыпучих материалов;

– Оценить степень значимости влияния каждого из факторов на изменение качества зернового и сыпучего материалов по частным коэффициентам корреляции;

– Составить прогноз изменения качества зерновой смеси и сыпучих материалов в зависимости от изменения показателей качества делимой смеси.

Используя полученные теоретические рассуждения, была разработана компьютерная программа, которая позволяет на основе экспериментальных данных получать математическую модель изучаемого объекта.

При этом стал вопрос выбора операционной среды для составления компьютерных программ. Выбор остановился на электронных таблицах Microsoft Excel и этот выбор объясняется многими причинами:

1) Электронные таблицы Microsoft Excel входят в пакет прикладных программ Microsoft Office и установлены практически на всех компьютерах.

2) Изучение основ работы с электронными таблицами Microsoft Excel в настоящее время проходят во всех учебных заведениях – начиная со школы и заканчивая средними и высшими учебными заведениями, поэтому можно надеяться на более-менее тесное знакомство большинства пользователей разрабатываемых программ.

3) Электронные таблицы Microsoft Excel достаточно просты в работе и. при достаточно профессионально написанной программе, не должны вызывать трудностей в работе для пользователей с невысоким уровнем подготовки. Особенно это касается процедуры запуска программ.

4) Электронные таблицы Microsoft Excel достаточно легко и просто подключать для совместной работы, что значительно расширяет диапазон их применения.

Данная компьютерная программа позволяет обработать результаты около 500 экспериментов (точек), включающих до 8 независимых переменных (причём это количество может быть значительно увеличено) и состоит из двух разделов – определение парной корреляции и построение математической на основе многофакторного эксперимента.

Парная корреляция основана на известных уравнениях определения коэффициентов в уравнениях элементарных функций методом наименьших квадратов. Все расчёты выполнены в виде таблиц (рис. 1). Таблицы, в свою очередь, состоят из ряда столбцов, из которых исследователю следует заполнять только два – значения переменных X и Y (Yфакт), а остальные (в том числе и номер опыта) заполняются автоматически.

2	№п/п	X	Yфакт	Yрасч							
3				Y = A×X+B	Y=X/(A×X+B)	Y=B+A/X	Y=A+B×Ln (X)	Y=A*X^B	Y=A*Exp(B*X)	Y=1/(A×X+B)	Y=A×X²+B×X+C
4				r = 1	r = 0,9972	r = 0,8656	r = 0,9675	r = 0,9842	r = 0,8428	r = 0,8835	r = 1
6	1	1	3	3	2,73121712	-4,0588235	2,72487127	2,89071778	0,754643181	3,73516572	3
7	2	2	5	5	5,146624683	5,51470588	6,593702712	10,7697213	2,236820813	4,488499818	5
8	3	3	7	7	7,298004346	8,70588235	8,856824027	23,2452556	6,63011006	5,622480671	7
9	4	4	9	9	9,226403927	10,3014706	10,46253415	40,123909	19,65215951	7,523136323	9
10	5	5	11	11	10,96478165	11,2588235	11,70801969	61,2756841	58,25052224	11,36503955	11
11	6	6	13	13	12,53990543	11,8970588	12,72565547	86,6030321	172,6590576	23,22611465	13
12	7	2	5	5	5,146624683	5,51470588	6,593702712	10,7697213	2,236820813	4,488499818	5
13	0										
14	0										
15	0										
526	0										
529											

Рис. 1. Таблица для расчёта парной корреляции

Компьютер просчитывает методом наименьших квадратов коэффициенты A, B и C для каждой из элементарных функций и на основе этого вычисляет расчётное значение функции Yрасч для каждого из значений аргумента X для каждой элементарной функции. После этого, опять-таки, для каждой элементарной функции определяется коэффициент корреляции r и на основе максимального значения этого коэффициента выбирается наиболее соответствующая функция Yк (рис. 2). Кроме того, когда несколько функций имеют одинаковый коэффициент корреляции, компьютер указывает остальные функции.

2	№п/п	X	Yфакт	Yрасч								Yк
3				Y = A×X+B	Y=X/(A×X+B)	Y=B+A/X	Y=A+B×Ln (X)	Y=A*X^B	Y=A*Exp(B*X)	Y=1/(A×X+B)	Y=A×X²+B×X+C	**Y=A*X^B**
4				r = 0,9774	r = 0,758	r = 0,7418	r = 0,8867	r = 1	r = 0,725	r = -0,5119	r = 1	**1**
6	1	1	1	-4	1,366806137	-53	4,254541503	1	15,52048937	1,706257799	1	**1**
7	2	2	4	4	3,2721202	1	20,12504385	4	4566,419945	2,161185817	4	**4**
8	3	3	9	12	6,112266112	19	29,4086926	9	1343526,652	2,946897143	9	**9**
9	4	4	16	20	10,79889807	28	35,9955462	16	395290815,6	4,630251124	16	**16**
10	5	5	25	28	20	33,4	41,10470679	25	1,16302E+11	10,79889807	25	**25**
11	6	6	36	36	46,2992126	37	45,27919495	36	3,42182E+13	-32,50243479	36	**36**
12	7	7	49	44	762,2222222	39,5714286	48,80867439	49	1,00676E+16	-6,487781555	49	**49**
13	0											
14	0											
15	0											
526	0											
529												
530	**Наиболее подходит функция Y=A*X^B с коэффициентами: A = 1; B = 2.**											
531	**Кроме того, возможно также использовать функцию Y=A×X²+B×X+C с коэффициентами: A = 1; B = 0 и C = 0**											
532	**Экстремальные координаты функции (минимум) будут следующие: X=2; Y = 4**											

Рис. 2. Вывод на экран монитора результаты расчётов

Если же в выбранной функции имеются точки экстремума, то компьютер указывает их, а также указывает на то, что представляют собой точки экстремума – минимум или максимум.

Кроме того, на экран выводится графическое изображение выбранной функции (рис. 3) и сопоставление фактических и расчётных данных (рис. 4). При этом под графиком появляется надпись, указывающая, сколько реальных точек (количество опытов) входит в этот график.

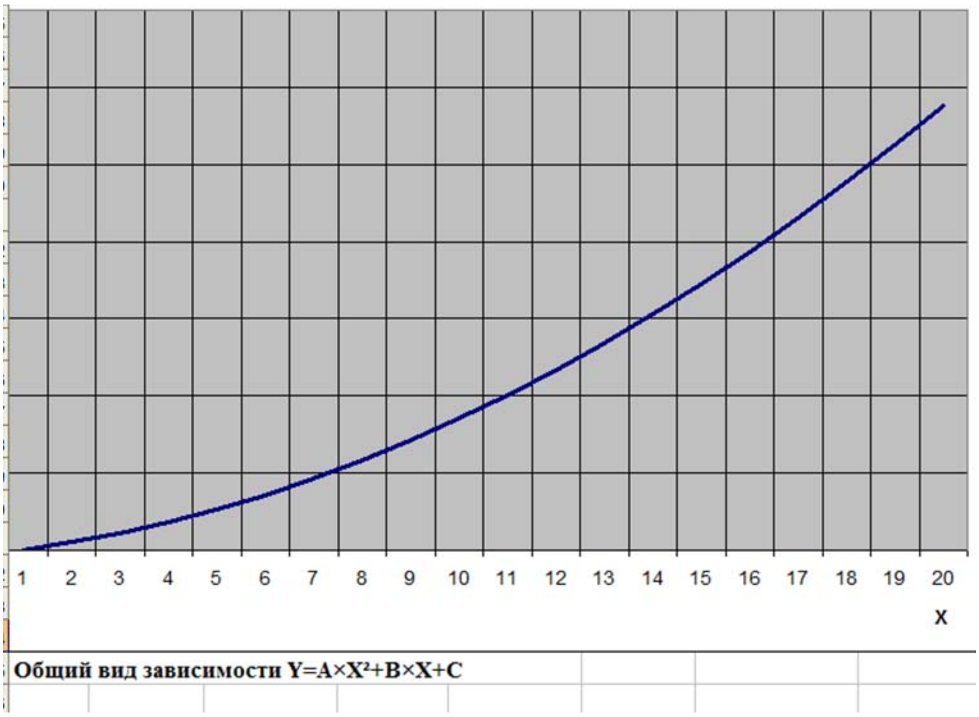


Рис. 3 Вид выбранной функции

Рис. 4 Сопоставление фактических и расчётных значений

На основании изложенной методики была разработана компьютерная программа для расчёта многофакторного эксперимента, которая состоит из блока ввода исходных данных (рис. 5) и единичных блоков парной корреляции, аналогичных вышерассмотренному.

Блок ввода исходных данных состоит из таблицы ввода аргументов, фактических данных, автоматического счётчика проведенных экспериментов и расчётных значений полученной зависимости. Исследователь вводит исходные данные по аргументам (X1- X8) и фактические значения функции (Yфакт). Далее компьютерная программа автоматически рассчитывает зависимости по каждому аргументу и строит обобщенную математическую модель.

	№п/п	X1	X2	X3	X4	X5	X6	X7	X8	Yфакт	Yк
5	1	1	2	3	0	-1	1	1	1	1	1,00
6	2	2	3	4	1	2	2	2	2	4	4,00
7	3	3	4	5	2	5	3	3	3	9	9,00
8	4	4	5	6	3	8	4	4	4	16	16,00
9	5	5	6	7	4	11	5	5	5	25	25,00
10	6	6	7	8	5	14	6	6	6	36	36,00
11	7	7	8	9	6	17	7	7	7	49	49,00
12	8	8	9	10	7	20	8	8	8	64	64,00
13	9	9	10	11	8	23	9	9	9	81	81,00
14	10	10	11	12	9	26	10	10	10	100	100,00
15	0										1
526	0										1
529											

530 **Внимание! Если Вы хотите просмотрть все результаты сопоставлния фактических и расчётных значений функции, откройте строки с 5 по 14!**

531 Математическая модель данного процесса описывается уравнением:

532 **Yк=[((1×X1^2+0×X1+0)×(1×X2^2+-2×X2+1)×(1×X3^2+-4×X3+4)×(1×X4^2+2×X4+1)×(0,1111×X5^2+0,8889×X5+1,7778)×(1×1^2)×(1×1^2)×(1×1^2)]^(1/8)**

Рис. 5 Таблица ввода исходных данных многофакторного эксперимента

Для наглядности насколько расчётные данные соответствуют фактическим, каждой точке определяется квадратическое отклонение и определяется суммарное среднеквадратическое отклонение по всему массиву.

Если в ходе проведения исследований окажется, что переменных (аргументов) будет не восемь, а меньше, то в таблице исходных данных лишние аргументы следует обнулить (рис. 6).

Поскольку математическая модель изучаемого процесса представляет собой среднегеометрическую величину от функций, получаемых от каждого аргумента (уравнение 4), то вполне может сложиться ситуация, когда при чётном количестве аргументов какое-либо значение из вычисленных функций будет иметь отрицательный результат. Тогда общее произведение (конечная функция) будет иметь отрицательное значение и, согласно правилам математики, чётная степень не может браться из отрицательного значения.

№п/п	X1	X2	X3	X4	X5	X6	X7	X8	Yфакт	Yк	δ²
1	1	2							1	1,67	0,45
2	2	3							4	3,47	0,28
3	6	4							9	16,64	58,37
4	4	5							16	13,58	5,86
5	5	6							25	22,25	7,56
6	6	7							36	33,28	7,40
7	7	8							49	46,65	5,53
8	8	9							64	62,35	2,73
9	9	10							81	80,37	0,40
10	10	11							100	100,72	0,51
0										1	0,00
											Σδ²/N = 8,91

Внимание! Если Вы хотите просмотрть все результаты сопоставлния фактических и расчётных значений функции, откройте строки с 5 по 14!

Математическая модель данного процесса описывается уравнением:

Yк=[((1,3414×X1^2+-3,7944×X1+5,2416)×(1×X2^2+-2×X2+1)×]^(1/2)

при общем коэффициенте корреляции 0,9959

Рис. 6 Таблица ввода исходных данных для расчётов с уменьшенным количеством аргументов

В этом случае следует сделать так, чтобы число аргументов стало нечетным, о чём появляется соответствующее сообщение в строке над таблицей (рисунок 7). Для этого следует либо добавить ещё один аргумент (если количество аргументов меньше 8), добавив пустое множество (например, повторить ещё раз уже имеющийся аргумент), либо убрать наименее значимый из аргументов, выбрав аргумент с наименьшим коэффициентом корреляции (рис. 7).

Кроме того, исследователь имеет возможность определить значение конечной функции Yк для различных значений аргументов (рис. 8), т.е получить прогноз для рассчитанной математической модели.

Внимание! В даном распределении желательно иметь нечётное количество аргументов! Рекомендуется убрать аргумент X2

№п/п	X1	X2	X3	X4	X5	X6	X7	X8	Yфакт	Yк	δ²
1	1	1	1	0	-1	1	5	1	1	1,13	0,02
2	2	2	2	1	2	2	6	5	4	3,86	0,02
3	3	16	3	2	5	3	7	9	9	#ЧИСЛО!	#ЧИСЛО!
4	4	4	4	3	8	4	8	13	16	13,41	6,71
5	5	5	5	4	11	5	9	17	25	20,20	23,07
6	6	6	6	5	14	6	10	21	36	28,42	57,44
7	7	7	7	6	17	7	11	25	49	38,26	115,34
8	8	8	8	7	20	8	12	29	64	50,07	194,08
9	9	9	9	8	23	9	13	33	81	64,65	267,47
10	10	10	10	9	26	10	14	37	100	84,45	241,81
0										1	0,00
0										1	0
	r(X1)=1	r(X2)=0,8894	r(X3)=1	r(X4)=1	r(X5)=1	r(X6)=1	r(X7)=1	r(X8)=1			#ЧИСЛО!

Внимание! Если Вы хотите просмотрть все результаты сопоставлния фактических и расчётных значений функции, откройте строки с 5 по 14!

Yк=[((1×X1^2)×(1/(-0,0364×X2+0,4022))×(1×X3^2)×(1×X4^2+2×X4+1)×(0,1111×X5^2+0,8889×X5+1,7778)×(1×X6^2)×(1×X7^2+-8×X7+16)×(0,0625×X8^2+0,375×

Рис. 7. Сообщение о необходимости иметь нечётное количество аргументов и массив, который можно удалить

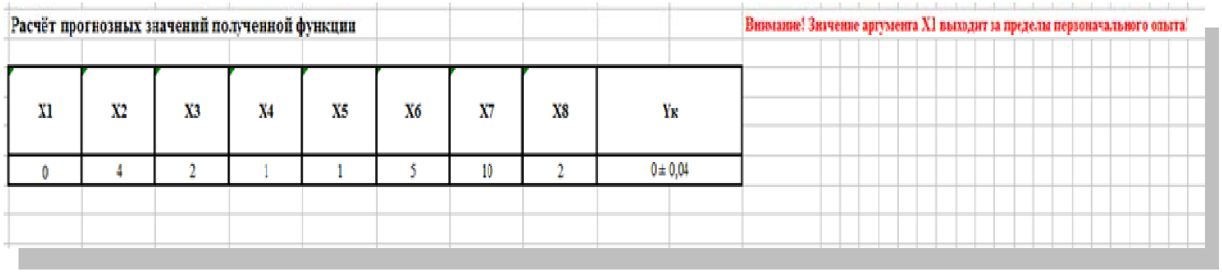
Расчёт прогнозных значений полученной функции

X1	X2	X3	X4	X5	X6	X7	X8	Yк
3	4	2	1	1	5	10	2	7,3223 ± 0,04

Рис. 8. Таблица расчёта прогнозных значений расчётной функции

Для этого в таблицу вводят прогнозные значения аргументов и в ячейке Yк получают прогнозное значение функции. Если введенное значение аргумента выходит за пределы значений в исходной таблице (предварительного эксперимента), на полях таблицы появляется соответствующее сообщение (рис. 9).

Для того, чтобы случайно не повредить программу, она защищена от несанкционированного доступа в запрограммированные ячейки. При попытке ввести в них какие-либо данные, работа программы автоматически останавливается, а на экран выводится соответствующее сообщение (рис. 10).

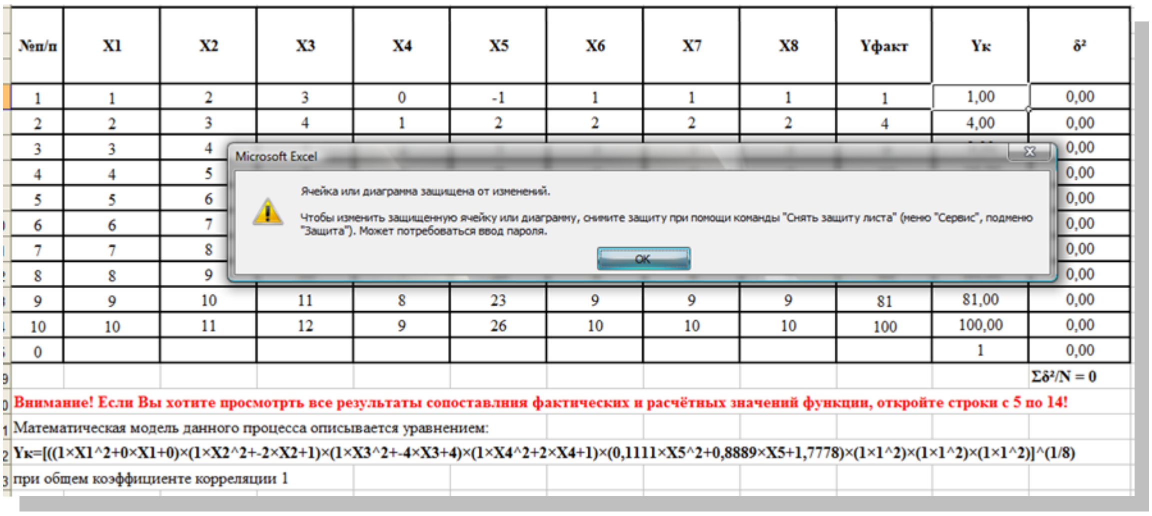


Расчёт прогнозных значений полученной функции

Внимание! Значение аргумента X1 выходит за пределы первоначального опыта!

X1	X2	X3	X4	X5	X6	X7	X8	Yк
0	4	2	1	1	5	10	2	0± 0,04

Рис. 9. Вывод предупреждения о том, что введённый аргумент выходит за пределы предварительного эксперимента



№п/п	X1	X2	X3	X4	X5	X6	X7	X8	Yфакт	Yк	δ²
1	1	2	3	0	-1	1	1	1	1	1,00	0,00
2	2	3	4	1	2	2	2	2	4	4,00	0,00
3	3	4									0,00
4	4	5									0,00
5	5	6									0,00
6	6	7									0,00
7	7	8									0,00
8	8	9									0,00
9	9	10	11	8	23	9	9	9	81	81,00	0,00
10	10	11	12	9	26	10	10	10	100	100,00	0,00
0										1	0,00

Σδ²/N = 0

Внимание! Если Вы хотите просмотрть все результаты сопоставлния фактических и расчётных значений функции, откройте строки с 5 по 14!

Математическая модель данного процесса описывается уравнением:

Yк=[((1×X1^2+0×X1+0)×(1×X2^2+-2×X2+1)×(1×X3^2+-4×X3+4)×(1×X4^2+2×X4+1)×(0,1111×X5^2+0,8889×X5+1,7778)×(1×1^2)×(1×1^2)×(1×1^2)]^(1/8)

при общем коэффициенте корреляции 1

Рис. 10. Вывод сообщения о попытке несанкционированного доступа в запрограммированную ячейку

Для сопоставления данного метода и классического метода планируемого эксперимента были проведены расчёты по обеим методикам. Оказалось, что, хотя и функции, описывающие процесс, были разными, среднеквадратичные отклонения этих функций были примерно одинаковы. Конечно, не следует идеализировать разработанный метод. В условиях, когда имеется возможность провести эксперименты строго на двух (или более) фиксированных уровнях, конечно, приоритет будет принадлежать классическому методу планирования эксперимента. Однако в тех случаях, когда наблюдаемые значения не сводятся в определенные разряды и приходится в виде общего правила иметь дело с неравноотстоящими значениями исследуемых величин (а это часто имеет место при исследовании действующих предприятий) новая методика вполне может найти своё место. Кроме того, использование предлагаемой методики позволяет значительно сократить количество опытов, необходимых для построения математической модели.

Выводы

В данной представлена новая методика построения математической модели функционирования произвольного объекта, представляющая собой среднегеометрическую величину элементарных функций, определяемых методом наименьших квадратов.

Сопоставление вновь разработанной методики и традиционного метода показало вполне удовлетворительные результаты.

Разработанный метод позволяет получать математическую модель функционирования объекта в тех случаях, когда её невозможно получить классическим методом планирования эксперимента.

При достаточно большом количестве опытов разница между результатами, полученные методом планирования эксперимента и вновь разработанным методом неощутимо мала.

Список литературы

1. Журавский А. А., Торяник Э. И., Крышень И. Г. и др. Автоматическое построение математической модели функционирования объекта. – Кокс и химия. – 200. – № 3. – С. 22–28.

2. Кучма Н. В., Зоря Е. С., Торяник Э. И, и др. Применение статистических методов при анализе влияния изменений угольной шихты на качество кокса. Углехимический журнал, 2003 г., №№ 5–6. – С. 15–24.

3. Л. Н. Тищенко, Д. И. Мазоренко, М. В. Пивень, С. А. Харченко, В. В. Бредихин, А. В Мандрыка//Моделирование процессов зерновых сепараторов. – Харьков: ХНТУСХ, «Міськдрук», 2010. – 360с., ил. – На рус.яз.

METHOD OF GETTING MATHMATICAL FREE OBJEKT

O.V. SINYAEVA, graduate student, M. M. ABDUJEV, Cand. Tech. Scie.
A. A. GURAVSKY, Pf. Filosf.

It is often challenging to fulfill all the requirements of scientific experiments while research. Thus, a new method has been worked out to design a stimulator of an object under study. It allows to avoid limits which may be caused by the traditional methods of experiment planning.

The comparison of the new and traditional methods shows numerous advantages of the former. Furthermore, a computer version of the new method has been developed to facilitate data processing of scientific experiments.

Поступила в редакцию 25.11 2011 г.