УДК 168+519.233.5

Радченко С. Г.

Д-р техн. наук, профессор кафедры технологии машиностроения Национального технического университета Украины «Киевский политехнический институт» им. И. Сикорского, Киев, Украина

КОРРЕКТНОЕ СТАТИСТИЧЕСКОЕ МОДЕЛИРОВАНИЕ В УСЛОВИЯХ НЕПОЛНОЙ ИСХОДНОЙ ИНФОРМАЦИИ

Актуальность. Рассмотрена задача статистического моделирования сложных систем и процессов в условиях неполной исходной информации.

Целью данной работы является использование метода формализованного получения структуры многофакторной статистической модели и устойчивого оценивания ее коэффициентов для получения высокоточных статистических моделей упругих деформаций технологической системы токарного станка.

Методы. При решении прикладных задач анализ исходных данных получения статистических моделей показал, что часто они строятся в условиях неполной исходной информации и решаемая задача является некорректно поставленной. В таких условиях проблемами построения моделей является получение структуры модели и ее устойчивость. Предложена расширенная концепция ортогональности получаемой модели: план эксперимента, структура модели и структурные элементы модели ортогональны друг к другу. Ортогональная структура многофакторной статистической модели позволяет получить статистически независимые оценки коэффициентов моделируемую функцию. Такая структура может быть определена однозначно со статистически значимыми коэффициентами. Нормирование ортогональных эффектов позволяет получить максимально устойчивую структуру модели и, следовательно, ее коэффициентов. Решаемая задача будет корректно поставленной.

Результаты. Применение рассмотренного метода формализованного получения структуры многофакторной статистической модели и устойчивого оценивания ее коэффициентов использовано для получения высокоточных статистических моделей упругих деформаций обрабатываемой на токарном станке стальной заготовки. Выполнен полный факторный эксперимент, где факторами служили сила резания, длина заготовки, диаметр заготовки, а откликом (функцией) – величина упругих деформаций системы. По результатам эксперимента построены статистические регрессионные модели деформаций $\hat{y}_1$ и $\hat{y}_2$. В структуре моделей факторы представлены ортогональными контрастами. При формировании структуры модели в нее вводятся статистически значимые эффекты. Проведенные проверки полученных моделей по критериям качества показали их высокую информативность, устойчивость, адекватность, статистическую эффективность. Использование моделей на станках с числовым программным управлением позволяет сократить число проходов режущего инструмента и, следовательно, время обработки детали.

Выводы. Результаты использования расширенной концепции ортогональности и структуры модели полного факторного эксперимента при получении моделей упругих деформаций технологической системы токарного станка подтвердили перспективность применения рассматриваемого подхода, его эффективность и целесообразность при построении регрессионных статистических моделей сложных систем и процессов.

Ключевые слова: статистическое моделирование, некорректные задачи, устойчивая структура статистической модели, расширенная концепция ортогональности.

НОМЕНКЛАТУРА

$\hat{y}$ – статистическая модель;

$x_i^{(1)}, x_i^{(2)}, ..., x_i^{(s_i-1)}$ – ортогональные контрасты фактора X_i;

s_i – число различных уровней фактора X_i;

k – общее число факторов, $1 \le i \le k$;

(1), (2), …, $(s_i - 1)$ – порядок контрастов фактора X_i;

N_Π – число структурных элементов полного факторного эксперимента, равное числу опытов эксперимента;

$x_{iu}^{(p)}$ – значение p-го ортогонального контраста i-го фактора для u-той строки матрицы планирования, $1 \le u \le N$, $1 \le p \le s_i - 1$;

$x_{ju}^{(q)}$ – значение q-го ортогонального контраста j-го фактора для u-той строки матрицы планирования, $1 \le q \le s_i - 1$, $1 \le i < j \le k$;

$\mathbf{X}$ – матрица эффектов полного факторного эксперимента;

$\sigma^2(\varepsilon)$ – теоретическое значение дисперсии воспроизводимости результатов опытов;

N – число опытов в плане эксперимента;

$\mathbf{E}$ – единичная матрица;

$\left| r_{ij}(x_i^{(p)} \times x_j^{(q)}) \right|$ – абсолютное значение коэффициента парной корреляции факторов X_i и X_j;

$F_{N-1;N(n-1)}^{\text{эксп}}$ – экспериментальное значение F-критерия (Фишера);

$F_{\alpha;N-1;N(n-1)}^{\text{крит}}$ – критическое значение F-критерия (Фишера);

$\bar{y}_u$ – среднее значение результатов опытов в u-той строке результатов;

$\bar{y}$ – среднее значение всех результатов опытов;

$\hat{y}_u$ – рассчитанное по модели значение отклика в u-той строке результатов;

y_{ul} – значение результата опыта в u-той строке результатов для l-го повторного опыта;

n – число повторных опытов в каждой строке результатов;

u – текущее значение номера строки в плане эксперимента; $1 \le u \le N$;

DOI 10.15588/1607-3274-2017-4-6

l – текущее значение номера повторного опыта;

α – уровень значимости;

$SS_{\text{ад}}$ – сумма квадратов адекватности для модели;

$SS_{\text{восп}}$ – сумма квадратов воспроизводимости результатов экспериментов;

$f_{\text{ад}}$ – число степеней свободы для проверки адекватности;

$f_{\text{восп}}$ – число степеней свободы для проверки воспроизводимости результатов экспериментов;

R – коэффициент множественной корреляции;

k' – общее число значимых коэффициентов уравнения регрессии;

$SS_{k'}$ – сумма квадратов отклонений, связанная с k' коэффициентом модели;

$SS_{\text{ост}}$ – сумма квадратов отклонений остаточная;

$f_{k'}$ – число степеней свободы для $SS_{k'}$;

$f_{\text{ост}}$ – число степеней свободы для $SS_{\text{ост}}$;

$\lambda_{\max}$ – максимальное собственное число для информационной матрицы Фишера $\mathbf{X}^{\mathrm{T}}\mathbf{X}$;

$\lambda_{\min}$ – минимальное собственное число для информационной матрицы Фишера $\mathbf{X}^{\mathrm{T}}\mathbf{X}$;

e_{ul} – остаток в u-той строке для l-го повторного опыта.

ВВЕДЕНИЕ

Многофакторные математические модели широко применяется при создании и совершенствовании сложных систем и процессов. В качестве исходной информации используются результаты экспериментальных исследований, статистических испытаний, экспертное оценивание, результаты сложных вычислений. Для определения коэффициентов моделей используется регрессионный анализ и метод наименьших квадратов.

Модель получают путем аппроксимации результатов экспериментов, которые являются результатом суммарного влияния управляемых, неуправляемых и неконтролируемых факторов. Определяемые коэффициенты смешаны с некоторыми другими коэффициентами, которые коррелируют с определяемыми. Необходимо по полученным результатам эксперимента $\mathbf{Y}$ восстановить математическую модель $\mathbf{Y} = \mathbf{XB} + \boldsymbol{\varepsilon}$ в условиях присутствия случайных и систематических ошибок $\boldsymbol{\varepsilon}$. Задача классифицируется как обратная. Типичным условием при получении моделей является отсутствие необходимой информации о статистически значимом влиянии факторов, структуре определяемой модели, законе распределения результатов наблюдений и других данных.

Большинство обратных задач являются некорректно поставленными. Корректными задачами называются классы математических задач, отвечающих некоторым условиям определенности их решений. Задача называется корректной задачей (или корректно поставленной), если выполнены следующие условия (условия корректности): задача имеет решение, каждым исходным данным соответствует только одно решение – однозначность задачи, решение устойчиво. Задача, для которой не выполняется хотя бы одно из условий, характеризующих корректно поставленную задачу, называется некорректно поставленной, или некорректной. Исследователь должен системно проанализировать условия получения модели и найти необходимое решение в условиях исходной неопределенности необходимой информации.

Цель работы – использование метода формализованного получения структуры многофакторной статистической модели и устойчивого оценивания ее коэффициентов для получения высокоточной статистической модели упругих деформаций обрабатываемой заготовки в условиях неполной исходной информации.

1 ПОСТАНОВКА ЗАДАЧИ

В соответствии с расширенной концепцией ортогональности выбираются план эксперимента, структура модели и ортогональные элементы модели. Рассматривается упругая система токарного металлорежущего станка. При обработке на станке цилиндрической стальной заготовки длиной l, диаметром d на нее влияет радиальная сила резания P, которая деформирует заготовку. Величина упругой деформации фиксируется при различных значениях влияющих факторов. По полученным результатам эксперимента строятся математические модели деформаций $\hat{y}_1$ и $\hat{y}_2$. Построенные модели проверяются по критериям качества: адекватность, информативность, устойчивость, статистическая эффективность.

2 ОБЗОР ЛИТЕРАТУРЫ

На указанные свойства сложных систем обращал внимание проф. В. А. Вознесенский. «К сожалению, в технологических, технико-экономических и других реальных сложных системах неизвестны ни вид функции φ (уравнение состояния системы – Р.С.), ни полный набор факторов $\mathbf{X}_\varphi$, ни числовые значения констант (коэффициентов) $\mathbf{B}_\varphi$, ни законы распределения случайных величин ε_φ, ни граничные условия» [1].

В публикациях по регрессионному анализу при выборе структуры многофакторной модели типичные решения сводятся к постулированию структуры в виде полинома определенного порядка без обоснования предложенной структуры. Полученная модель может быть неадекватной и неустойчивой. Отмечается «...принципиальное отличие неформализуемой задачи выбора вида искомой математической модели исследуемого явления и чисто формальной задачи нахождения по экспериментальным данным значений параметров уже выбранной аппроксимирующей функции» [2].

Д.ф.-м.н. С.А. Айвазян, д.ф.-м.н. И.С. Енюков и д.ф.-м.н. Л.Д. Мешалкин отмечают, что «на первый план выходит задача правильного определения структуры модели (т. е. выбора общего вида функции $f(X)$), решение которой обеспечивает возможность количественного измерения эффекта воздействия на $Y(X)$ каждой из объясняющих переменных $x^{(1)}, x^{(2)}, ..., x^{(p)}$ в отдельности. Однако как раз это место (правильный выбор общего вида функции $f(X)$) и является самым слабым во всей технике статистического исследования зависимостей: к сожалению, не существует стандартных приемов и методов, которые образовывали бы строгую теоретическую базу для решения этой важнейшей задачи» [3].

В [4] отмечается, что «...этап исследования, посвященный выбору общего вида функции регрессии (параметризация модели), бесспорно является ключевым: от того, насколько удачно он будет реализован, решающим образом зависит точность восстановления неизвестной функции регрессии $f(X)$. В то же время приходится признать, что этот этап находится, пожалуй, в самом невыгодном положении: к сожалению, не существует системы стандартных рекомендаций и методов, которые образовали бы строгую теоретическую базу для его наиболее эффективной реализации».

Авторы предлагают использовать априорную информацию о содержательной сущности анализируемой зависимости; предварительный анализ геометрической структуры исходных данных; различные статистические приемы обработки исходных данных, позволяющие сделать наилучший выбор из нескольких сравниваемых вариантов.

Не приведены [4] рекомендации по выбору плана эксперимента для получения модели.

В рассмотренных публикациях известных специалистов по регрессионному анализу не рассмотрена проблема некорректно поставленных задач, устойчивости получаемых статистических моделей, проверки их на устойчивость и по другим критериям качества. Такой подход типичен и для других публикаций [5, 6]. В книге [5] изложены классическая теория и современные результаты, полученные в регрессионном анализе. Приведены основы регрессионного и дисперсионного анализа в предположении, что результаты наблюдений распределены нормально, методики выбора статистических гипотез и подбора оптимальной модели с использованием информационных критериев. В [6] изложена теория и практика планирования научных экспериментов: латинские квадраты; планы 2^k, 2^{k-p}, 3^k, $2^m\times3^n$; планы второго порядка, планы Бокса-Бенкина. Получение структуры статистической модели, устойчивость оценивания коэффициентов не рассматривается.

В монографии [7] исследованы проблемы использования математических и эвристических методов получения многофакторных статистических моделей реальных сложных систем и процессов. Рассмотрено взаимодействие математических и эвристических методов в получении наилучших возможных моделей для стандартных и нестандартных областей факторного пространства.

Оптимизация светодиодной системы освещения витаминной космической оранжереи в [8] осуществляется с использованием экспериментально-статистического подхода и полученной статистической модели, которая характеризуется хорошими параметрами качества, так как было использовано оптимальное планирование эксперимента.

3 МАТЕРИАЛЫ И МЕТОДЫ

Для аппроксимации исходных данных будем использовать класс полиномиальных математических моделей. Их применение обосновано теоремами Вейерштрасса, Стоуна, Джексона [9].

Элементы структуры определяемой статистической модели $\hat{y}$ выбираются из элементов структуры модели полного факторного эксперимента.

По теореме Бродского В. З. [9] в модели полного факторного эксперимента все эффекты ортогональны друг другу и число эффектов равно числу опытов. Если эффекты выразить в виде ортогональных нормированных контрастов, т. е.

$$\sum_{u=1}^{N} x_{iu}^{(p)} = 0,\ \sum_{u=1}^{N} x_{iu}^{(p)} \times x_{ju}^{(q)} = 0,$$

$$\sum_{u=1}^{N} [x_{iu}^{(p)}]^2 = N,\ \sum_{u=1}^{N} [x_{iu}^{(p)} \times x_{ju}^{(q)}]^2 = N,$$

что эквивалентно системе ортогональных полиномов Чебышева, то матрица дисперсий-ковариаций будет диагональной

$$(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\, \sigma^2(\varepsilon) = (1/N)\mathbf{E}\ \sigma^2(\varepsilon).$$

Модель полного факторного эксперимента соответствует наилучшим возможным критериям *D*-, *A*-, *E*-, *G*-оптимальности, ортогональности и адекватна результатам экспериментов.

В общем случае использование полного факторного эксперимента требует много опытов и при решении некоторых реальных прикладных задач невыполнимо. В качестве квазиполного факторного эксперимента необходимо использовать дробный факторный эксперимент, применять многофакторные регулярные планы и планы на основе ЛП_τ равномерно распределенных последовательностей [9, 7].

Выбор структуры модели проводится с использованием алгоритма RASTA3 и программного средства «Планирование, регрессия и анализ моделей» (ПС ПРИАМ) [9].

Использование многофакторных регулярных планов позволяет получить все главные эффекты ортогональными. Корреляция взаимодействий должна быть ограничена коэффициентом парной корреляции $\left| r_{ij}(x_i^{(p)} \times x_j^{(q)}) \right| \le 0{,}3)$. Бóльшая коррелированность эффектов создает смешивание определяемых коэффициентов, увеличивает их среднеквадратические ошибки, и решаемая задача становится некорректно поставленной.

Под устойчивой структурой многофакторной статистической модели полиномиального вида понимается структура, характеризующаяся неизменностью множества главных эффектов и взаимодействий при изменениях значений результатов экспериментов (откликов), порождаемых случайными ошибками (погрешностями) результатов наблюдений, измерений, вычислений и неопределенностью искомой структуры модели.

При формировании структуры модели необходимо вводить в модель статистически значимые факторы. Так как информация о них часто отсутствует, то необходимо оптимизировать значения тех факторов, которые не изменяются в процессе функционирования системы и тем самым уменьшить число факторов, вводимых в структуру определяемой системы. Общая методика проведения многофакторного статистического моделирования приведена в [9, 10].

Полученную математическую модель необходимо проверить на статистическую значимость. Предполагается равномерное дублирование опытов: $n_u = n = \text{const}$.

$$F^{\text{эксп}}_{N-1;N(n-1)} = (\sum_{u=1}^{N}(\bar{y}_u - \bar{y})^2/(N-1))/(\sum_{u=1}^{N}\sum_{l=1}^{n}(y_{ul} - \bar{y}_u)^2/N(n-1)).$$

Если $F^{\text{эксп}}_{N-1;N(n-1)} > F^{\text{крит}}_{\alpha;N-1;N(n-1)}$, то получение математической модели целесообразно; в противном случае нецелесообразно, так как суммарное влияние группы факторов X_1, ..., X_k будет статистически незначимо на фоне случайных ошибок повторных опытов, порождаемых изменчивостью неуправляемых и неконтролируемых факторов.

Математическая модель должна быть проверена на адекватность [9]. Если

$$F^{\text{эксп}}_{f_{\text{ад}};f_{\text{восп}}} = (SS_{\text{ад}}/f_{\text{ад}})/(SS_{\text{восп}}/f_{\text{восп}}) = s^2_{\text{ад}}/s^2_{\text{восп}} \le F^{\text{крит}}_{\alpha;f_{\text{ад}};f_{\text{восп}}},$$

$$SS_{\text{ад}} = n\sum_{u=1}^{N}(\bar{y}_u - \hat{y}_u)^2, \; SS_{\text{восп}} = \sum_{u=1}^{N}\sum_{l=1}^{n_u}(y_{ul} - \bar{y}_u)^2,$$

$$f_{\text{ад}} = N - k', \; f_{\text{восп}} = N(n-1),$$

то модель адекватна и прогноз результатов по модели не противоречит результатам опытов. В противном случае модель не адекватна; необходимо проанализировать конкретную ситуацию и наметить действия к разрешению противоречия.

Модель необходимо проверить на информативность. Вычисляют значение коэффициента множественной корреляции R

$$R = \left\{1 - \left[\sum_{u=1}^{N}(\bar{y}_u - \hat{y}_u)^2\right] \Big/ \left[\sum_{u=1}^{N}(\bar{y}_u - \bar{y})^2\right]\right\}^{1/2}.$$

Гипотезу о значимости множественного коэффициента корреляции проверяют по F-критерию [9]:

$$F^{\text{эксп}}_{f_{k'};f_{\text{ост}}} = (SS_{k'}/f_{k'})/(SS_{\text{ост}}/f_{\text{ост}}) = s^2_{k'}/s^2_{\text{ост}},$$

$$SS_{k'} = \sum_{u=1}^{N}(\bar{y}_u - \bar{y})^2 - \sum_{u=1}^{N}(\bar{y}_u - \hat{y}_u)^2,$$

$$SS_{\text{ост}} = \sum_{u=1}^{N}\sum_{l=1}^{n}(y_{ul} - \hat{y}_u)^2,$$

$$f_{k'} = k' - 1, \; f_{\text{ост}} = Nn - k'.$$

Если $F^{\text{эксп}}_{f_{k'};f_{\text{ост}}} \le F^{\text{крит}}_{\alpha;f_{k'};f_{\text{ост}}}$, то гипотеза о статистической незначимости R принимается: разброс значений отклика по уравнению регрессии $\hat{y}$ относительно общего среднего по всем результатам опытов $\bar{y}$ отличается статистически незначимо от разброса результатов опытов y_{ul} относительно уравнения регрессии $\hat{y}$, т. е. прогноз по уравнению регрессии не дает больше полезной информации, чем знание $\bar{y}$.

Если $F^{\text{эксп}}_{f_{k'};f_{\text{ост}}} > F^{\text{крит}}_{\alpha;f_{k'};f_{\text{ост}}}$, то гипотеза о статистической значимости R принимается; прогноз по уравнению регрессии дает больше полезной информации, чем знание $\bar{y}$.

Коэффициенты математической модели должны быть устойчивы к малым случайным изменениям в исходных данных, полученных в процессе экспериментирования. Предложено для количественного показателя устойчивости коэффициентов математической модели использовать 2 меры обусловленности: по Нейману-Голдстейну и число обусловленности матрицы $\mathbf{X}^{\mathrm{T}}\mathbf{X}$.

Мера обусловленности матрицы по Нейману-Голдстейну

$$\mathrm{P}(\mathbf{X}^{\mathrm{T}}\mathbf{X}) = \lambda_{\max}/\lambda_{\min},$$

Другая мера обусловленности матрицы $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ обозначается cond.

$$\mathrm{cond}(\mathbf{X}^{\mathrm{T}}\mathbf{X}) = \left\|\mathbf{X}^{\mathrm{T}}\mathbf{X}\right\| \times \left\|(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\right\|,$$

где $\|\cdot\|$ – обозначение нормы матрицы.

Предполагается, что матрица $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ невырождена.

Модель максимально устойчива, если приведенные критерии равны 1; хорошо устойчива, если критерии не превышают 10.

Полученную математическую модель необходимо проверить по всей области моделирования. При использовании ПС ПРИАМ строится трехмерное изображение поверхности отклика, т. е. $\hat{y}_w = f_w(X_i, X_j)$. Каждый из двух факторов изменяется на 20 уровнях. Все остальные факторы принимают средние значения. Всего вычисляется 400 значений отклика.

Полученное изображение поверхности отклика анализируется и определяются минимальное и максимальное расчетные значения $\hat{y}_w$. Они сравниваются с физически возможными значениями отклика. В случае имеющегося противоречия модель должна быть проанализирована и работа по ее получению продолжена.

Более жесткой проверкой является поиск минимума и максимума по модели $\hat{y}_w$ с использованием ЛП_τ равномерно распределенных последовательностей с последующим их сравнением с возможными физически не противоречивыми значениями моделируемого критерия качества.

Проводится также проверка свойств остатков

$$e_{ul} = y_{ul} - \hat{y}_u,$$

что позволяет подтвердить или опровергнуть предположения относительно распределения ошибок и систематических, т. е. не случайных, изменений от последовательности проведения опытов и от значений отклика, рассчитанных по модели в точках факторного пространства, в которых проводились опыты.

Получаемая статистическая модель должна быть семантичной. Семантичность (информационная) – получение информации о механизмах происходящих явлений в системах и процессах путем изучения структур моделей.

p-ISSN 1607-3274. Радіоелектроніка, інформатика, управління. 2017. № 4
e-ISSN 2313-688X. Radio Electronics, Computer Science, Control. 2017. № 4

Получение семантических многофакторных моделей возможно для моделей, линейных относительно параметров, и формализованного формирования структурной связи в виде оператора структурной связи, задаваемого множеством эффектов модели полного факторного эксперимента. В указанном типе моделей коэффициенты интерпретируются как оценки частных производных моделируемой функции. Анализ полученных статистических моделей позволяет вскрыть механизмы явлений, происходящих в системе и процессе.

Семантичность позволяет использовать коэффициенты полученной статистической модели в целях интерпретации в терминах предметной (моделируемой) области знаний. Для выполнения данного требования необходимо, чтобы тип моделей был линейный по параметрам, а коэффициенты модели были статистически независимы или близки к независимым (коэффициент парной корреляции между эффектами должен быть не более 0,3 по абсолютной величине) и нормированы. Тогда абсолютная величина коэффициента модели показывает силу влияния соответствующего эффекта, а знак – направление влияния. Коэффициенты главных эффектов будут приближенно описывать производные первого, второго и других порядков, а взаимодействий – частные производные соответствующих эффектов.

4 ЭКСПЕРИМЕНТЫ

Экспериментальное исследование эффективности предложенной методики проведено при математическом моделировании упругих деформаций технологической системы токарного станка для консольно закрепленной в трехкулачковом патроне цилиндрической стальной заготовки в состоянии поставки. Заготовка была свободно зажата (результаты y_1, мкм) или поджата задним центром (y_2, мкм). Нагрузка, имитирующая действие составляющей силы резания P_y, устанавливалась с помощью образцового динамометра сжатия в пределах 200–2000 Н. Величина упругих перемещений в нагружаемом сечении фиксировалась с помощью индикатора часового типа. Замеры деформаций проводились с точностью до 1 мкм. Сила изменялась в диапазоне $P(X_1)$ = 200–1000 Н, длина консольной части заготовки $l(X_2)$ = 20–140 мм, диаметр заготовки $d(X_3)$ = 18–50 мм.

Для обеспечения высокоточного прогнозирования деформаций технологической системы был проведен полный факторный эксперимент $3^1\times4^1\times5^1//60$, в котором сила изменялась на трех уровнях (200; 600; 1000 Н), длина консольной части заготовки в точке замера деформации – на четырех уровнях (20; 60; 100; 140 мм), диаметр заготовки – на пяти уровнях (18; 26; 34; 42; 50 мм). Было проведено 60 опытов; каждый опыт повторен 2 раза (табл. 1).

5 РЕЗУЛЬТАТЫ

Структура модели полного факторного эксперимента $3^1Ч4^1Ч5^1//60$ следующая:

$$(1+x_1^{(1)}+x_1^{(2)})(1+x_2^{(1)}+x_2^{(2)}+x_2^{(3)})(1+x_3^{(1)}+x_3^{(2)}+x_3^{(3)}+x_3^{(4)}) \to N_{60},$$

где $x_1^{(1)}, x_2^{(1)}, x_3^{(1)}$, $x_1^{(2)}, x_2^{(2)}, x_3^{(2)}, x_2^{(3)}, x_3^{(3)}, x_3^{(4)}$ – соответственно линейные, квадратичные, кубичные, четвертой степени контрасты факторов X_1, ..., X_3; N_{60} – число структурных элементов для схемы полного факторного эксперимента.

Все эффекты (главные и взаимодействия) были нормированы.

Вычисление математических моделей и всех их критериев качества было проведено с использованием ПС ПРИАМ. Полученные математические модели имеют вид

$$\begin{gathered}\hat{y}_1 = 121,43-129,31x_3+119,83x_2-158,30x_2x_3+79,95x_1+\\ +59,38z_3-85,08x_1x_3+79,10x_1x_2+75,66x_2z_3-103,28x_1x_2x_3+\\ +38,93x_1z_3-37,19z_2x_3+26,14z_2+49,26x_1x_2z_3-31,71x_2g_3-\\ -23,13g_3+19,13z_2z_3-23,68x_1z_2x_3+16,65x_1z_2-21,71x_1x_2g_3-\\ -15,79x_1g_3++12,00x_1z_2z_3-8,70z_2g_3+7,74x_2v_3-5,84x_1z_2g_3+\\ +3,70v_3+5,98x_1x_2v_3-4,12g_2x_3+2,53g_2+2,90z_2v_3+3,54x_1v_3+\\ +2,12x_1g_2+1,88g_2z_3-2,63x_1g_2x_3+1,91x_1z_2v_3+1,14x_1g_2z_3+\\ +1,02z_1x_3+1,29z_1x_2x_3-1,09z_1x_2z_3-0,87g_2g_3+0,78z_1z_2x_3-\\ -0,64z_1z_2z_3-0,43z_1z_3-0,40z_1-0,73z_1g_2v_3+\\ +0,48z_1z_2g_3+0,62z_1g_2g_3,\ \text{мкм},\end{gathered}$$

$$\begin{gathered}\hat{y}_2 = 30,55+18,95x_1+10,33x_2-10,51x_3-6,95x_1x_3+6,18x_1x_2-\\ -5,83x_2x_3+3,43z_3+3,09v_3-2,32g_3+3,46x_2v_3+2,25x_1z_3+\\ +2,22x_2z_3-2,60x_2g_3+1,28z_2-2,27x_1x_2x_3-1,20z_1+1,58x_1x_2z_3-\\ -1,35x_1g_3+1,43x_1v_3-1,55x_1x_2g_3-0,72z_2z_3-1,30z_1v_3-\\ -1,64z_1x_2v_3+1,07z_1x_3+1,17x_1x_2v_3+0,73x_1z_2x_3-0,48x_1z_2z_3,\text{мкм}.\end{gathered}$$

Формулы для вычисления ортогональных контрастов:

$$x_1^{(1)} = x_1 = 0,0025\ (X_1-600);$$

$$x_1^{(2)} = z_1 = 1,5\ (x_1^2-0,666667);$$

$$x_2^{(1)} = x_2 = 0,00166667\ (X_2-80);$$

$$x_2^{(2)} = z_2 = 2,25\ (x_2^2-0,555556);$$

$$x_2^{(3)} = g_2 = 3,75\ (x_2^3-0,911111x_2);$$

$$x_3^{(1)} = x_3 = 0,0625\ (X_3-34);\quad x_3^{(2)} = z_3 = 2\ (x_3^2-0,5);$$

$$x_3^{(3)} = g_3 = 3,33333\ (x_3^3-0,85x_3);$$

$$x_3^{(4)} = v_3 = 7,77778\ (x_3^4-1,10714\,x_3^2+0,128571).$$

Адекватность математических моделей

Для $\hat{y}_1$ $F^{\text{эксп}}_{13;60} = 1,6 < F^{\text{крит}}_{0,05;13;60} = 2,297$,

Для $\hat{y}_2$ $F^{\text{эксп}}_{32;60} = 1,88 > F^{\text{крит}}_{0,05;32;60} = 1,71$.

Обе модели приемлемы для использования.

Доли рассеивания, объясняемые моделями $\hat{y}_1$ и $\hat{y}_2$ соответственно: 0,999996; 0,9979 – очень высокие. Коэффициенты множественной корреляции R = 0,999998; 0,998933 весьма близки к 1 и статистически высокозначимы: F-отношения для R равны 63232; 554 при критических значениях 1,83; 1,95. Информативность моделей очень высокая. Критерий Бокса и Веца для информативности больше 49 для первой модели и равен 14 для второй.

Число обусловленности для обеих моделей cond = 1: коэффициенты максимально устойчивы относительно их случайных возмущений.

Таблица 1 – Рабочая матрица плана эксперимента, результаты повторных опытов и расчетные значения откликов

п/п	Кодированные значения уровней варьирования факторов			Факторы			Функции					
				Натуральное обозначение			Натуральное обозначение					
				P	l	d	y_1			y_2		
				Кодир. обозначение			Кодированное обозначение					
1	F_1	F_2	F_3	X_1	X_2	X_3	y_1			y_2		
2	0	0	0	200	20	18	Результаты повторных опытов		Расчет. значение отклика	Результаты повторных опытов		Расчет. значение отклика
3	1	1	1	600	60	26						
4	2	2	2	1000	100	34						
5	–	3	3	–	140	42						
6	–	–	4	–	–	50	y_{11}	y_{12}	$\hat{y}_1$	y_{21}	y_{22}	$\hat{y}_2$
7	О	1		200	20	18	12	13	12,6	8	8	8,5
8	п	2		200	20	26	11	10	11,0	8	7	8,5
9	ы	3		200	20	34	8	8	8,5	7	7	7,4
10	т	4		200	20	42	8	9	8,6	7	8	8,1
11		5		200	20	50	8	9	8,9	8	9	8,2
12		6		200	60	18	38	39	38,8	13	13	11,1
13		7		200	60	26	22	22	22,1	9	10	7,3
14		8		200	60	34	12	12	11,3	8	9	7,7
15		9		200	60	42	11	12	12,0	8	8	7,4
16		10		200	60	50	9	9	8,9	8	7	8,2
17		11		200	100	18	118	118	117,3	17	17	17,1
18		12		200	100	26	50	50	50,3	9	10	9,7
21		15		200	100	50	15	16	15,3	9	10	8,8
22		16		200	140	18	268	266	267,0	25	26	26,8
23		17		200	140	26	100	98	98,9	16	14	15,6
31		25		600	20	50	22	23	21,9	18	19	17,5
32		26		600	60	18	120	121	120,1	40	39	41,2
33		27		600	60	26	68	68	67,6	27	28	26,9
36		30		600	60	50	29	29	28,7	19	19	20,5
37		31		600	100	18	348	350	350,2	58	54	54,9
38		32		600	100	26	148	147	147,1	30	31	31,0
41		35		600	100	50	43	44	44,3	24	25	24,7
42		36		600	140	18	789	790	789,5	70	72	69,7
43		37		600	140	26	281	280	280,6	36	41	38,3
44		38		600	140	34	127	125	126,3	46	54	47,9
45		39		600	140	42	80	80	80,3	31	32	30,8
46		40		600	140	50	66	67	66,5	29	30	30,0
54		48		1000	60	34	66	66	65,4	40	38	39,1
57		51		1000	100	18	580	584	581,6	89	83	85,0
58		52		1000	100	26	238	237	237,2	49	49	50,7
60		54		1000	100	42	87	87	87,3	40	39	40,5
61		55		1000	100	50	73	73	72,4	39	40	39,3
62		56		1000	140	18	1296	1292	1293,9	101	105	104,4
63		57		1000	140	26	463	461	462,1	60	65	61,8
64		58		1000	140	36	218	214	215,8	62	69	66,1
65		59		1000	140	42	137	140	138,3	52	54	53,1
66		60		1000	140	50	112	112	112,0	48	50	48,1
Примечание. Приведены отдельные фрагменты рабочей матрицы												

Эффективность извлечения полезной информации – 100%. Коэффициенты моделей отвечают критериям *D*-, *A*-, *E*-, *G*-оптимальности. Они также ортогональны, информационно семантичны.

6 ОБСУЖДЕНИЕ

Математическое моделирование по схеме полного факторного эксперимента позволило получить прецизионные математические модели, удовлетворяющие всему комплексу необходимых требований.

Общий вывод: математические модели соответствуют всем необходимым критериям качества и могут быть использованы для прогноза упругих деформаций технологической системы токарного станка.

Недостаточность исходной информации при статистическом моделировании сложных систем создает условия некорректно решаемой задачи. Необходим системный подход к формированию условий, при которых решаемая задача будет корректно поставленной. Таким условием является использование расширенной концепции ортогональности и структуры модели полного факторного эксперимента. Используемый подход позволяет получить многофакторные статистические модели с наилучшими возможными критериями качества. Модели $\hat{y}_1$ и $\hat{y}_2$ адекватны, информативны, устойчивы, эффективны. Они характеризуются как семантические, что позволяет использовать их в предметной области для интерпретации влияния воздействующих факторов.

Используемый метод построения моделей позволяет получить статистически независимые коэффициенты с минимально возможными их среднеквадратическими ошибками.

Использование аналогичных многофакторных математических моделей совместно с обработкой на станках с числовым программным управлением позволяет для нежестких деталей программно предсказать траекторию движения инструмента и добиться более точной обработки без увеличения числа проходов инструмента.

Примеры успешно решенных реальных задач приведены в [9].

С разработанными методами моделирования и полученными результатами можно ознакомиться в [11, 12].

ВЫВОДЫ

1. При отсутствии необходимой информации для получения статистических моделей необходимо использовать расширенную концепцию ортогональности: план эксперимента, структура моделей и эффекты моделей должны быть ортогональными или близкими к ортогональным. Структуру определяемой модели следует выбирать из элементов структуры модели полного факторного эксперимента.

2. Использование ортогональной структуры модели, когда структурные элементы модели ортогональны друг другу, позволяет получить несмещенные коэффициенты модели, минимальные среднеквадратические ошибки коэффициентов, максимально устойчивую структуру модели. Критерии качества такой модели являются наилучшими из возможных.

3. Если выполнить необходимые условия получения статистических моделей не представляется возможным, необходимо применять квазиусловия путем использования эвристических решений. Использование методологии теории планирования эксперимента позволяет получить наилучшие условия решения задачи моделирования и решать ее как корректно поставленную.

4. Получение характеризующихся наилучшими возможными критериями качества статистических моделей упругих деформаций консольно закрепленной детали при обточке ее на токарном станке подтвердило правильность выбранного подхода и метода получения высокоточных моделей. Использование моделей на станках с числовым программным управлением позволяет сократить время обработки заготовки за счет сокращения числа проходов режущего инструмента.

БЛАГОДАРНОСТИ

Работа выполнена на кафедре «Технология машиностроения» Национального технического университета Украины «Киевский политехнический институт» им. И. Сикорского. Автор благодарен ст. преподавателю Лапачу С. Н. за полезные замечания при обсуждении статьи и доц. Беланенко В. Г. за участие в проведении эксперимента.

СПИСОК ЛИТЕРАТУРЫ

1. Современные методы оптимизации композиционных материалов / [Вознесенский В. А., Выровой В. Н., Керш В. Я. и др.] ; под ред. В. А. Вознесенского. – К. : Будівельник, 1983. – 144 с.
2. Новицкий П. В. Оценка погрешностей результатов измерений / П. В. Новицкий, И. А. Зограф. – 2-е изд., перераб. и доп. – Л. : Энергоатомиздат. Ленингр. отд-ние, 1991. – 304 с.
3. Айвазян С. А. Прикладная статистика : Исследования зависимостей : справ. изд. / С. А. Айвазян, И. С. Енюков, Л. Д. Мешалкин ; под. ред. С. А. Айвазяна. – М. : Финансы и статистика, 1985. – 487 с.
4. Прикладная статистика. Основы эконометрики: Учебник для вузов: В 2 т. 2-е изд. испр. – Т. 1 : Айвазян С. А., Мхитарян В. С. Теория вероятностей и прикладная статистика. – М. : ЮНИТИ-ДАНА, 2001. – 656 с.
5. Малов С. В. Регрессионный анализ. Теоретические основы и практические рекомендации / С. В. Малов. – Санкт-Петербург : Изд-во СПбГУ, 2013. – 280 с.
6. Hinkelmann K. Design and Analysis of Experiments, Introduction to Experimental Design / Klaus Hinkelmann, Oscar Kempthorne. – 2 edition. – Vol. 1. – Wiley-Interscience, 2007. – 631 p. – (Wiley Series in Probability and Statistics).
7. Радченко С. Г. Формализованные и эвристические решения в регрессионном анализе : монография / Радченко С. Г. – К. : «Корнійчук», 2015. – 236 с.
8. Оптимизация светодиодной системы освещения витаминной космической оранжереи / [Коновалова И. О., Беркович Ю. А., Ерохин А. Н. и др.] // Авиакосмическая и экологическая медицина. – 2016. – Т. 50, № 3. – С. 17–22.
9. Радченко С. Г. Методология регрессионного анализа : монография / С. Г. Радченко. – К. : «Корнійчук», 2011. – 376 с.
10. Радченко С. Г. Анализ методов моделирования сложных систем / С. Г. Радченко // Математичні машини і системи. – 2015. – № 4. – С. 123–127.
11. Лаборатория экспериментально-статистических методов исследований (ЛЭСМИ). – [Электронный ресурс]. – Режим доступа: http://www.n-t.org/sp/lesmi
12. Сайт кафедры «Технология машиностроения» Механико-машиностроительного института Национального технического университета Украины «Киевский политехнический институт». – [Электронный ресурс]. – Режим доступа: http://tm-mmi.kpi.ua/index.php/ru/1/publications/

Статья поступила в редакцию 07.02.2017.
После доработки 04.04.2017.

Радченко С. Г.
Д-р техн. наук, професор кафедри технології машинобудування Національного технічного університету України «Київський політехнічний інститут» ім. І. Сікорського, Київ, Україна

КОРРЕКТНЕ СТАТИСТИЧНЕ МОДЕЛЮВАННЯ В УМОВАХ НЕПОВНОЇ ПЕРВИННОЇ ІНФОРМАЦІЇ

Актуальність. Розглянуто завдання статистичного моделювання складних систем і процесів в умовах неповної первинної інформації.

Метою роботи є використання методу формалізованого отримання структури багатофакторної статистичної моделі і стійкого оцінювання її коефіцієнтів для одержання високоточних статистичних моделей пружних деформацій технологічної системи токарного верстата.

Методи. При вирішенні прикладних задач аналіз первинних даних отримання статистичних моделей показав, що часто вони будуються в умовах неповної первинної інформації і розв’язувана задача є некоректно поставленою. В таких умовах проблемами побудови моделей є отримання структури моделі і її стійкість. Запропоновано розширену концепція ортогональності одержуваної моделі: план експерименту, структура моделі та структурні елементи моделі ортогональні. Ортогональна структура багатофакторної статистичної моделі дозволяє отримати статистично незалежні оцінки коефіцієнтів модельованої функції Така структура може бути визначена однозначно зі статистично значущими коефіцієнтами. Нормування ортогональних ефектів дозволяє отримати максимально стійку структуру моделі і, отже, її коефіцієнтів. Розв’язувана задача буде коректно поставленою.

Результати. Застосування розглянутого методу формалізованого отримання структури багатофакторної статистичної моделі і стійкого оцінювання її коефіцієнтів використано для одержання високоточних статистичних моделей пружних деформацій оброблюва-

ної на токарному верстаті сталевої заготовки. Виконано повний факторний експеримент, де факторами служили сила різання, довжина заготовки, діаметр заготовки, а відгуком (функцією) – величина пружних деформацій системи. За результатами експерименту побудовано статистичні регресійні моделі деформацій $\hat{y}_1$ і $\hat{y}_2$. У структурі моделей фактори представлені ортогональними контрастами. При формуванні структури моделі в неї вводяться статистично значущі ефекти. Проведені перевірки одержаних моделей за критеріями якості показали їх високу інформативність, стійкість, адекватність, статистичну ефективність. Використання моделей на верстатах з числовим програмним управлінням дозволяє скоротити кількість проходів ріжучого інструмента і, отже, час обробки деталі.

Висновки. Результати використання розширеної концепції ортогональності і структури моделі повного факторного експерименту при отриманні моделей пружних деформацій технологічної системи токарного верстата підтвердили перспективність застосування розглянутого підходу, його ефективність і доцільність при побудові регресійних статистичних моделей складних систем і процесів.

Ключові слова: статистичне моделювання, некоректні задачі, стійка структура статистичної моделі, розширена концепція ортогональності.

Radchenko S. G.

Dr. Sc., Professor Department of Mechanical Engineering, I. Sikorsky National Technical University of Ukraine «Kyiv Polytechnic Institute», Kyiv, Ukraine

CORRECT STATISTICAL MODELING IN CONDITIONS OF INCOMPLETE INITIAL INFORMATION

Context. Urgent problem of statistical modeling of the complex systems and processes in conditions of incomplete initial information has been considered.

Objective. The work objective is the use of the method of formalized obtaining of the structure of multifactor statistical model and stable estimation of its coefficient for obtaining highly precise statistical models of elastic deformations of technological system of a lathe.

Methods. When solving applied problems, the analysis of initial data on obtaining statistical models has shown that they are often constructed in conditions of incomplete initial information and the problem which is solved is incorrectly formulated one. In such conditions the model structure obtaining and its stability prove to be the problems in models construction. The author proposes an extended conception of orthogonality of the obtained model: the experiment design, model structure and structure elements of the model are orthogonal. The orthogonal structure of the multifactor statistical model allows obtaining statistically independent estimates of coefficients of the modeled function. Such a structure may be defined unambiguously with statistically significant coefficients. Normalization of orthogonal effects permits obtaining a maximally stable model structure, and, consequently, its coefficients. The problem will be well-posed.

Results. Application of the considered method of formalized obtaining of the structure of multifactor statistical model and stable estimation of its coefficients is used for obtaining accurate statistical models of elastic deformations of a steel work-piece processed by a lathe. A complete factor experiment has been fulfilled; the factors were as follows: cutting force, work-piece length and diameter, and a response – the value of elastic deformations of the system. Statistical regression moments $\hat{y}_1$ and $\hat{y}_2$ were constructed as the experiment result. In the structure of models the factors are presented by orthogonal contrasts. Statistically significant effects are introduced in the model structure under its formation. The checks of the obtained models by quality criteria have shown their high informativeness, stability, adequacy, statistical efficiency. Using the models on lathes with numerical programmed control allows decreasing the number runs of the cutting tool and, consequently, the time of work-piece processing.

Conclusion. The results of the use of the extended conception of orthogonality and structure of the model of a complete factor experiment, when obtaining the models of elastic deformations of technological system of a lathe, have confirmed the great prospects of application of the considered approach, its effectiveness and expediency in constructing regression statistical models of complex systems and processes.

Keywords: statistical modeling, ill-posed problems, stable structure of statistical model, extended conception of orthogonality.

REFERENCES

1. Voznesenskij V. A., Vy'rovoj V. N., Kersh V. Ya. i dr. pod red. V. A. Voznesenskogo Sovremenny'e metody' optimizacii compozicionny'x materialov. Kyiv, Budivelnyk, 1983, 144 p.
2. Noviczkij P. V., Zograf I. A. Izdanie 2, pererabotannoe i dopolnennoe Ocenka pogreshnostej rezultatov izmerenij. Leningrad, Energoatomizdat, 1991, 304 p.
3. Ajvazyan S. A., Enyukov I. S., Meshalkin L. D., pod red. S. A. Aivazyan. Prikladnaya statistika: Issledovanie zavisimostej: Spravochnoe izdanie. Moscow, Finansy i statistika, 1985, 487 p.
4. Ajvazyan S. A., Mxitaryan V. S. Pikladnaya statistika. Osnovy ekonometriki: Uchebnik dlya vuzov, v 2 tomax, izdanie 2, ispravlennoe. Vol. 1, Teoriya veroyatnostej i prikladnaya statistika. Moscow, YUNITI-DANA, 2001, 656 p.
5. Malov S .V. Regressionny'j analiz. Teoreticheskie osnovy i prakticheskie rekomendacii. St.Petersbourg, Izdatelstvo SPbGU, 2013, 280 p.
6. Hinkelmann K. Kempthorne O. Design and Analysis of Experiments, Introduction to Experimental Design, 2 edition, Vol. 1,Wiley-Interscience, 2007, 631 p. (Wiley Series in Probability and Statistics).
7. Radchenko S. G. Formalizovanny'e i e'vristicheskie resheniya v regressionnom analize: monografiya. Kyiv, «Kornijchuk», 2015, 236 p.
8. Konovalova I. O., Berkovich Yu. A., Eroxin A. N. i dr. Optimizaciya svetodiodnoj sistemy' osveshheniya vitaminnoj kosmicheskoj oranzherei, *Aviakosmicheskaya i e'kologicheskaya medicina*, 2016, Vol. 50, No. 3, pp. 17–22.
9. Radchenko S. G. Metodologiya regressionnogo analiza : monografiya. Kyiv, «Kornijchuk», 2011, 376 p.
10. Radchenko S. G. Analiz metodov modelirovaniya slozhnih sistem, *Matematichni mashini i sistemi*, 2015, No. 4, pp. 123–127.
11. Laboratory of experimental statistical method of investigation (LESMI) [Electronic resource]. Access mode: http://www.n-t.org/sp/lesmi
12. Site of the Department Mechanical Engineering, Mechanical Engineering Institute, National Technical University of Ukraine «Kyiv Polytechnic Institute» [Electronic resource]. Access mode: http://tm-mmi.kpi.ua/index.php/ru/1/publications/