чувствительность ИКРП, межэлектродное расстояние $d_0$, коэффициент модуляции L, РВЭ рабочей поверхности зонда $Ф_1$ и рабочая площадь зонда S.

Для обеспечения достоверности при идентификации и интерпретации результатов измерений и диагностики необходимо нормировать параметры ИКРП, отрабатывать основные положения и методы поверки.

Необходимым условием совершенствования метода КРП является развитие метрологического обеспечения, что и будет предметом наших дальнейших исследований.

**Список литературы:** 1. *Вудраф Д.* Современные методы исследования поверхности. Пер. с англ. / *Д. Вудраф, Т. Делчар.* – М.: Мир, 1989. – 564 с. 2. *Царев Б. М.* Контактная разность потенциалов / *Царев Б. М.* – М.: ГИТТЛ, 1955. – 280 с. 3. *Компанеец И. В.* Физические основы конденсаторных методов измерения контактной разности потенциалов / *И. В. Компанеец* // Вестник НТУ «ХПИ». Тематический выпуск «Автоматика и приборостроение», 2009. – № 23. - С. 89-95. 4. *Компанеец И. В.* Оценка чувствительности измерителя контактной разности потенциалов / *И. В. Компанеец, В. М. Комолов, А. М. Шкилько* // Вестник НТУ «ХПИ». Тематический выпуск «Новые решения в современных техноогиях», 2010. – № 46. - С. 89-94. 5. *Компанієць І. В.* Вимірювач контактної різниці потенціалів / *І. В. Компанієць, А. М. Шкілько* // Метрологія та прилади. – 2010. – №4. – С. 33-36.

*Поступила в редколлегию 25.11.2010*

**УДК 539.3: 519.876.5**

***А.М. МИЛЬЦЫН*,** канд. техн. наук, профессор, начальник отдела ТД Днепропетровского завода сварочных материалов, г. Днепропетровск
***Д.Г. ЗЕЛЕНЦОВ*,** доктор технических наук, старший научный сотрудник, профессор, Украинский государственный химико-технологический университет, г. Днепропетровск
***В.И. ОЛЕВСКИЙ*,** канд. техн. наук, доцент, Украинский государственный химико-технологический университет, г. Днепропетровск

## СЕЛЕКЦИЯ МНОГОФАКТОРНОЙ МОДЕЛИ ПО СОВОКУПНОСТИ СТАТИСТИЧЕСКИХ И ФИЗИЧЕСКИХ КРИТЕРИЕВ

Предложено использование физически обоснованных критериев селекции наилучшей многофакторной регрессионной модели, обеспечивающих ее работоспособность в области эксперимента и при экстраполяции. Продемонстрирована эффективность такого подхода.
Ключевые слова: модель, селекция, критерии

Запропоновано використання фізично обґрунтованих критеріїв селекції найкращої багатофакторної регресійної моделі, що забезпечують її працездатність в області експерименту і при екстраполяції. Продемонстровано ефективність такого підходу.
Ключові слова: модель, селекція, критерії

The use of physically based criteria of selection of best multifactor regression model is proposed to ensure its serviceability in the experiment field and for extrapolation. The effectiveness of this approach is demonstrated.
Keywords: model, selection, criteria

## 1. Введение

Известно, что работоспособность одной и той же конструкции, системы и технологического процесса можно представить различными математическими моделями, отличающимися числом переменных, видом уравнения и критериями работоспособности. Аппроксиманты, полученные на основе обработки выборки, зачастую имеют близкие, статистически неразличимые характеристики и выбор из них наилучшего приближения превращается в неформальную задачу, решение которой каждый исследователь производит по-своему. В связи с этим необходимо формализовать исследования по выбору наилучшей аппроксимирующей модели. Современный подход к этой проблеме базируется на использовании методов искусственного интеллекта и теории нечетких множеств. Наиболее часто применяется селективный отбор модели на основе группового учета аргумента [1]. В этом методе приняты несколько принципов построения и выбора модели. Основной из них состоит в том, что моделирование относится к некорректно поставленным задачам, и для нахождения однозначного решения необходимо задание внешнего дополнения - критерия оптимальности. Под этим понимается критерий, который вычисляется на базе информации, не использованной при оценке параметров модели. В качестве внешнего критерия селекции наиболее часто выбирается критерий минимума смещения, в соответствии с которым требуется, чтобы модели, построенные по одной части матрицы эксперимента как можно меньше отклонялись от моделей, построенных по другой ее части. Критерий минимума смещения является мощным инструментом отбора модели, но он не специфичен к рассматриваемой области моделирования и может свидетельствовать лишь об определенных качествах матрицы исходных данных. Распознание истинной зависимости в аналитической форме требует привлечения знаний о физической природе рассматриваемого объекта или процесса. Поэтому основой выбора наилучшей модели должны стать как статистические критерии, так и критерии работоспособности модели в особых точках факторного пространства, находящихся вне области эксперимента. Такой подход принят в методе структурно-экстраполяционного анализа [2], который показал свою эффективность при решении ряда практических задач. Адекватное поведение модели в таких точках позволяет говорить о ней, как о некой физически обоснованной зависимости, применимой к задаче в целом, а не как о локальной аппроксиманте эмпирических данных.

## 2. Методика выбора нелинейной полиномиальной модели

Пусть заданы $n$ факторов, математически интерпретируемых как независимые переменные $x_1, x_2, ..., x_n$ и образующие $n$-мерное факторное пространство. Будем полагать, что вследствие действия стохастических связей независимые переменные $x_i, i = \overline{1,n}$ определяются как случайные величины. Функция отклика $y$ при этом может быть определена как некоторый обобщенный или частный параметр, характеризующий работоспособность конструкции, объекта, системы и технологического процесса в заранее выбранном смысле по качественным и количественным критериям. Уравнение

$y = y(x_1, x_2, ..., x_n)$ в $(n+1)$ – мерном пространстве описывает некоторую гиперповерхность. Связь между параметрами $y$ и $\{x_i\}_{i=1}^{n}$ определяется математической моделью, исходя из концепции безусловного существования связи между ними. Математические модели в общем виде могут быть представлены различными функциональными зависимостями, например, в виде полиномов Колмогорова - Габора, степенных или логарифмических функций. При использовании зависимости вида

$$y = в_o + \sum_{i=1}^{n} в_i x_i + \sum_{i<j}^{n} в_{ij} x_i x_j + \sum_{i=1}^{n} в_{ii} x_i^2 + \dots,$$

где $в_0 = y|_{x=0}; в_i = \left.\frac{\partial y}{\partial x_i}\right|_{x=0}; \; в_{ij} = \left.\frac{\partial^2 y}{\partial x_i \partial x_j}\right|_{x=0}; \; в_{ii} = \left.\frac{1}{2}\frac{\partial^2 y}{\partial x_i^2}\right|_{x=0}$

задача сводится к отысканию выборочных оценок $\{в\}$ истинных коэффициентов $\{\beta\}$ основе экспериментальных данных. При этом может быть использован, в частности регрессионный анализ экспериментальных данных, включающих обработку результатов эксперимента по методу наименьших квадратов и статистический анализ полученного представления. Кроме традиционно используемых критериев величины коэффициента множественной детерминации и адекватности модели экспериментальным данным по формуле Фишера, существует еще ряд статистических критериев, дополняющих процедуру определения наилучшей модели. Так, качество математической модели может оцениваться на основе исследования невязок расчетных и экспериментальных данных. В случае выполнения предпосылок регрессионного анализа и полного соответствия структуры истинной и смоделированной зависимостей сумма невязок во всех опытах при достаточной точности вычислений должна быть близка к нулю. Таким образом, величина суммы невязок является интегральным показателем качества эмпирической зависимости. Значительное отличие её от нуля указывает либо на неудовлетворительный состав параметров, либо на плохое соответствие закона распределения параметров нормальному закону. В этом случае математическое описание не соответствует истинной качественной взаимосвязи параметров и на его основе опасно делать эвристические выводы как внутри, так и вне области эксперимента.

Другим важным показателем качества математической модели является некоррелируемость невязок. Независимость остатков указывает на отсутствие систематической ошибки модели и, следовательно, о соответствии экспериментальных данных условиям применимости регрессионного анализа. Этот факт проверяется критерием Дурбина - Ватсона [3].

Дальнейший анализ структуры модели должен базироваться на использовании физически обоснованных критериях [4]. Они отражают выполнение некоторых обязательных с физической точки зрения соотношений в особых точках факторного пространства, в которых предсказанные моделью значения функции отклика сравниваются с теоретически очевидным результатом. В этих точках требуется выполнение неравенств

$$\frac{\hat{y}-\hat{y}'}{\sqrt{S_0^2}}\sqrt{N} \le t\,,$$

где $\hat{y}, \hat{y}'$ - значения регрессии для физически эквивалентных состояний,

$S_0^2$ - средневзвешенная дисперсия воспроизводимости опытов,

$N$ - количество точек факторного пространства, используемых для оценки критерия,

$t$ - значение критерия Стьюдента при выбранном уровне значимости и $N-1$ степенях свободы.

Выбранная на основе совокупности статистических и физических критериев математическая модель может быть использована для практических расчетов и анализа нелинейного взаимодействия и влияния факторов на несущую способность, как в области эксперимента, так и при некоторой экстраполяции.

**3. Применение методики расчета**

Приведенная методика была применена к построению математической модели влияния семи вторичных факторов, на отклонение параметра устойчивости продольно сжатой цилиндрической оболочки от расчетного значения $\Delta T$. Факторы характеризовали основные виды отклонения геометрических параметров оболочки от идеальных обводов. Рассматривалась овально-коническая оболочка, близкая к круговой цилиндрической, с периодической в окружном направлении неплоскостностью торца, имеющая инициирующую потерю устойчивости лунку. Модель была получена при обработке результатов многофакторного эксперимента с тонкостенными оболочками диаметром 0,143 м, толщиной стенки 2,5х$10^{-4}$ м и длиной 0,2 м, изготовленными из стали марки Х18Н9-н. Статистические характеристики факторов приведены в таблице 1. Эксперимент проводился в соответствии с ядром плана $2^{7-2}$, являющегося четверть – репликой полного факторного эксперимента $2^7$. Для определения коэффициентов при нелинейных членах план эксперимента был дополнен до плана второго порядка при общем числе опытов $M$ = 47. Таким образом, был получен квазиортогональный центральный композиционный план при варьировании семи факторов на пяти уровнях. В дальнейшем ансамбль факторов был дополнен данными наблюдений за восьмым фактором, параметры которого также представлены в таблице 1.

Выбор наилучшей модели производился методом полного перебора всех регрессий [2]. Все генерируемые модели разбивались на классы по числу линеаризованных переменных. В каждом классе строились модели и из них выбирались наилучшие по критерию множественной детерминации $R^2$ в пределах погрешности его определения при заданном уровне значимости. Затем наилучшие модели по каждому из классов сравнивались по $F$-критерию Фишера.

Таблица 1

Факторы и их статистические данные

| Описание переменных | Условные обозначения | Единица измерения | Среднее значение | Параметр рассеивания |
|---|---|---|---|---|
| Глубина лунки | $w$ | м | 9,1x10$^{-4}$ | 3,830x10$^{-4}$ |
| Продольная длина лунки | $l_L$ | м | 3,3x10$^{-2}$ | 9,309x10$^{-3}$ |
| Поперечная ширина унки | $l_\varphi$ | м | 2,6x10$^{-2}$ | 6,903x10$^{-3}$ |
| Конусность | $\alpha$ | градусы | 94 | 4,5 |
| Овальность | $a/b$ | - | 0,9 | 0,039 |
| Амплитуда неплоскостности торца | $A_T$ | м | 1,5x10$^{-4}$ | 4,800x10$^{-5}$ |
| Число волн неплоскостности | $n_T$ | - | 12 | 5 |
| Разнотолщинность | $\delta_{\max} - \delta_{\min}$ | м | 3,2x10$^{-6}$ | 1,500x10$^{-6}$ |

Результаты анализа данных по моделям различных классов, имеющих наилучшее значение $R^2$, приведены на рис. 1. Коэффициент детерминации $R^2$, изменяющийся в пределах (0,1), асимптотически приближается к единице с увеличением числа переменных $n$, достигая $R^2$=0,80 при $n=11$, $R^2=0,90$ при $n=18$, а при $n=36$ статистически неотличим от единицы.

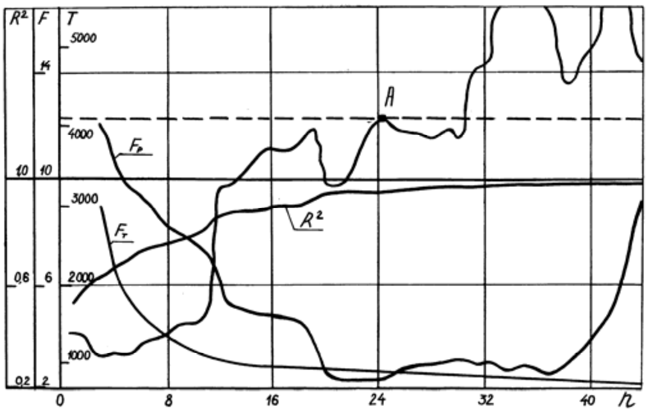


Рис. 1. Зависимость теоретических и экспериментальных критериев выбора наилучшей модели от числа переменных

При отборе моделей по $F$-критерию наилучшими считались модели, для которых выполнялось условие $F \le F_T$, где $F_T$ - табличное значение $F$-критерия при соответствующих уровне значимости $q$ и степенях свободы $f_1$ и $f_2$. Наилучшими моделями по $F$-критерию оказались модели с числом линеаризованных переменных $n$ от 20 до 25. На следующем шаге отбора наилучшей модели использовались физические критерии.

Так, для рассматриваемой задачи в качестве особой точки было принято, в частности, теоретическое значение верхней критической нагрузки $T_в / 2\pi E\delta^2 = 0,605$, соответствующее $\Delta T = 0$ и отсутствию отклонений формы оболочки от идеальных обводов. Для модели, окончательно выбранной в качестве наилучшей по совокупности критериев при уровне значимости $q = 5\%$, эти соотношения выполняются следующим образом (точка $A$ на рис. 1).

Для уравнения с натуральными переменными в соответствующей точке факторного пространства $l_\varphi = l_L = w = n_T = A_T = \delta_{\max} - \delta_{\min} = \alpha = 0$, $(a/b)=1$ значение функции имеет погрешность $\Delta T = 24,33$ кг при $T_в = 4095$ кг, что составляет 0,6%. Для стандартизованного уравнения в этой же точке $l_\varphi^0 = -3,76$, $A_T^0 = -3,06$, $n_T^0 = -2,84$, $\alpha^0 = -2,08$, $l_L^0 = -3,54$, $w^0 = -2,37$, $(a/b)^0 = -2,54$, $(\delta_{\max} - \delta_{\min})^0 = -2,13$ значение функции имеет погрешность $\Delta T^0 = -9,1004$ ($\Delta T = -29,485$) при $T_в = 4149$ кг, что составляет 0,7%.

Далее использовались критерии этой группы, отражающие качественную природу исходных переменных. Так, наличие плоского торца можно задать в виде $\Delta T(n_T = 0, \forall A_T) = \Delta T(A_T = 0, \forall n_T)$ для произвольного (фиксированного) уровня остальных переменных (табл. 2). Аналогично, сопоставление результатов для соотношений эквивалентных отсутствию лунки при $l_\varphi = 0$, $l_L = 0$ и $w = 0$ (табл. 3) показало, что разброс параметра несущей способности относительно среднего значения в центре плана не превышает выбранного уровня значимости.

Предложенные критерии позволили оставить единственную модель, удовлетворяющую наилучшим образом их совокупности. Полученная в результате процесса селекции наилучшая модель имеет вид

$$\begin{aligned}
&\Delta T^0 = -0,187(A_T^0)^2 - 0,239(n_T^0)^2 - 0,408(\delta_{\max} - \delta_{\min})^0 \alpha^0 + \\
&+0,174(a/b)^0 \alpha^0 - 0,974\alpha^0 l_L^0 - 0,375(\delta_{\max} - \delta_{\min})^0 A_T^0 + \\
&+0,135(a/b)^0 A_T^0 - 0,159 l_L^0 A_T^0 - 0,28 A_T^0 n_T^0 + 0,242 w^0 l_\varphi^0 - \\
&-0,086 n_T^0 l_\varphi^0 - 0,25(a/b)^0 (\delta_{\max} - \delta_{\min})^0 + 0,149 n_T^0 (\delta_{\max} - \delta_{\min})^0 - \\
&-1,066 w^0 (a/b)^0 + 0,124 n_T^0 l_T^0 + 0,194(\delta_{\max} - \delta_{\min})^0 - \\
&-0,186(a/b)^0 + 0,015 w^0 - 0,162 l_L^0 - 0,572 n_T^0 - 0,378.
\end{aligned}$$

Таблица 2

Проверка выполнения критерия наличия плоского торца

| Уровень сопутствующих факторов | $\frac{T_в - \Delta T}{2\pi E \delta^2}$ | Средняя погрешность, % |
|---|---|---|
| -1 | 0,381 | 1,05 |
| 0 | 0,268 | 0,18 |
| +1 | 0,406 | 2,22 |

Таблица 3

Проверка выполнения критерия отсутствия лунки

| Уровень сопутствующих факторов | $\frac{T_в - \Delta T}{2\pi E \delta^2}$ | Погрешность (%) | | | Средняя погрешность, % |
|---|---|---|---|---|---|
| | | $l_L = 0$ | $l_\varphi = 0$ | $w = 0$ | |
| -1 | 0,385 | 15,94 | 17,10 | 1,12 | 11,38 |
| 0 | 0,239 | 7,40 | 0,55 | 6,80 | 4,90 |
| +1 | 0,317 | 31,20 | 29,60 | 1,57 | 20,79 |

## 4. Выводы

Приведена методика, позволяющая получить наилучшее уравнение регрессии методом селекции модели по совокупности статистических и физически обоснованных критериев. Проанализированы преимущества использования таких критериев, выражающиеся в возможности получения модели, пригодной для экстраполяции за область эксперимента. Произведен расчет конкретной технической задачи и показана эффективность предлагаемой методики. Показано, что синтез наилучшей многофакторной модели второго порядка должен сопровождаться анализом ее физической адекватности по приведенной методике с учетом статистических характеристик экспериментальных данных.

**Список литературы:** 1. Зайченко, Ю. П. Синтез и адаптация нечетких прогнозирующих моделей на основе принципа самоорганизации [Текст] / Ю. П. Зайченко, И. О. Заец // Труды Одесского политехнического университета. – 2001. – Вып. 3(15). – С. 178 – 184. 2. Пилов П. И., Многофакторный структурно-экстраполяционный анализ в задачах управления эффективностью обогатительных процессов [Текст] / П. И. Пилов, А. М. Мильцын, В. И. Олевский // Збагачення корисних копалин: Наук.-техн. зб. – 2009. – Вип. 36(77)-37(78). – С. 204 – 217. 3. Смирнов, Н. В. Курс теории вероятностей и математической статистики [Текст] / Н. В. Смирнов, И. В. Дунин-Барковский. - М.: Наука, 1965. - 511 с. 4. Мильцын, А. М. Алгоритмизация построения, статистической обработки и анализа многофакторной регрессионной модели, содержащей управляемые, неуправляемые или смешанные переменные, в интерактивном режиме [Текст]: монография / А. М. Мильцын, В. И. Олевский // Днепропетровск, 1988. - 157 с. Деп. в ВИНИТИ 13.01.88, №188 - В88.

*Поступила в редколлегию 01.12.2010*

**УДК 621.114.32**

***В.С. ЄРЕМЕНКО,*** канд. техн. наук, доцент, НАУ, м. Київ
***А.В. ПЕРЕЇДЕНКО,*** студент, НАУ, м. Київ
***Є.О. ПІКОЛЕНКО,*** студент, НТУУ «КПІ», м. Київ

# РАНЖУВАННЯ ІНФОРМАТИВНИХ ОЗНАК ПРИ НЕРУЙНІВНОМУ КОНТРОЛІ КОМПОЗИЦІЙНИХ МАТЕРІАЛІВ

Описан способ ранжирования по информативности диагностических признаков при многопараметровом контроле. Приведены результаты применения даного способа для отбора признаков при контроле композиционных материалов методом низкоскоростного удара.
Ключевые слова: дисперсионный анализ, сотовые панели, классификация.

Описано спосіб ранжирування за інформативністю діагностичних ознак при багатопараметровому контролі. Наведено результати застосування даного способу для відбору ознак при контролі композиційних матеріалів методом низькошвидкісного удару.
Ключові слова: дисперсійний аналіз, стільникові панелі, класифікація.

This article is devoted to a method of ranking by informative diagnostic signs for multiparameter monitoring. The results of this method for feature selection in the control of composite materials by low-velocity impact.
Key words: analysis of variance, honeycombed, classification.