

Клименко Павел Петрович

E-mail: rts@tsure.ru.

Maksimov Aleksandr Viktorovich

Taganrog Institute of Technological – Federal State-Owned Educational Establishment of Higher Vocational Education “Southern Federal University”.

E-mail: kafmps@tppark.ru.

81, Petrovskay street, Taganrog, 347900, Russia.

Phone: +78634328052; +78634325825.

Ghelozhe Yury Andreevich

E-mail: rts@tsure.ru.

Klimenko Pavel Petrovich

E-mail: rts@tsure.ru.

УДК 681.3.07: 628.162.8

М.В. Игнатьев, В.М. Игнатьев

РЕГРЕССИОННЫЕ МОДЕЛИ ПРОЦЕССОВ ОБЕЗЗАРАЖИВАНИЯ ПИТЬЕВОЙ ВОДЫ

Строятся нелинейные регрессионные модели обеззараживания питьевой воды. Остатки этих моделей проверялись по статистикам Уайта и Дарбина-Уотсона.

Регрессионные модели.

M.V. Ignat'ev, V.M. Ignat'ev

REGRESSION MODELS DEACTIVATE OF DRINKING WATER

Create non-linear regression models deactivating of drinking water. Residuals of this models tested on statistics White and Darbin-Watson.

Regression models.

Для оценки эффективности применения дезактивантов при обеззараживании питьевой воды проводились эксперименты [1]. При проведении экспериментов в качестве дезактивантов использовались ионы серебра, меди и цинка с фиксацией четырёх факторов: c – удельная массовая концентрация препарата, мг/л; t – температура обеззараживаемой воды, °С; τ – продолжительность процесса обеззараживания, сутки или минуты; K – показатель порядка уменьшения содержания клеток бактерий *E-coli* в 1 литре воды. Показатель степени обеззараживания питьевой воды (K) оценивался как логарифм числа выживших бактерий *E-coli*. Результатирующими факторами процесса являются степень обеззараживания K и продолжительность процесса τ . Экспериментальные данные однородны по критерию Колмогорова-Смирнова, следовательно, для данных можно подобрать закон распределения или построить его с помощью кривых Пирсона [2]. Эксперимент включал 78 опытов (n) по обеззараживанию питьевой воды ионами меди, 126 опытов с применением ионов серебра и 52 опыта с применением ионов цинка [1]. На основании полученных статистических данных были построены многофакторные нелинейные модели, пригодные для расчётов результирующих факторов, которые приведены в табл. 1.

Таблица 1

Вид модели и значения критериев Фишера F и $F_{\text{теор}}$

Номер модели	n	Вид уравнения	Корреляция	F	$F_{\text{теор}}$
Медь	1	$\tau = \exp(0,783) t^{-0,264} c^{-0,334} K^{1,167}$ (0,132) (0,037) (0,38) (0,07)	0,888	92,5	8,563
	2	$K = \exp(-0,37) \tau^{0,676} t^{0,22} c^{0,286}$ (0,114) (0,04) (0,026) (0,025)	0,93	156,5	8,563
Серебро	3	$\tau = \exp(1,082) c^{-0,322} t^{-0,863} K^{0,788}$ (0,052) (0,006) (0,016) (0,016)	0,993	2783,2	8,55
	4	$K = \exp(-2) c^{0,325} t^{0,346} \tau^{1,006}$ (0,22) (0,027) (0,03) (0,06)	0,846	102,11	8,55
Цинк	5	$\tau = \exp(1,5) c^{0,168} t^{-0,203} K^{0,662}$ (0,22) (0,436) (0,088) (0,11)	0,665	12,7	8,583
	6	$K = \exp(-1,05) c^{-0,22} t^{0,307} \tau^{0,656}$ (0,266) (0,04) (0,081) (0,108)	0,78	25,4	8,583

Расчёты по оценке значимости моделей проводились с помощью методов дисперсионного анализа по критерию Фишера при 5-процентном уровне значимости [2]. Для всех уравнений, приведённых в табл. 1, выполняется условие $F > F_{\text{теор}}$, следовательно, все полученные модели значимы. Робастные оценки всех коэффициентов уравнений значимы (см. значения оценок, приведённые в табл. 1 в 4 столбце ниже коэффициентов регрессионных уравнений).

В табл. 1 коэффициент парной корреляции процесса обеззараживания и модели (1) в первой строке таблицы равен $r = 0,888$ и значим по критерию Стьюдента на 5 %

уровне значимости, так как $T_{\text{экс}} = r \cdot \sqrt{\frac{n-3}{1-r^2}} = 16,724$, $T_{\text{ст}}(0,975; 77) = 1,991$ и

$T_{\text{экс}} > T_{\text{ст}}$. Значимость r указывает на наличие линейной связи данных опыта времени экспозиции τ и значений, получаемых с помощью построенной регрессионной модели (1) из первой строки табл. 1. Для всех приведённых в данном исследовании моделей коэффициенты корреляции значимы. При отсутствии значимости коэффициента парной корреляции нельзя рассматривать уравнение в качестве модели [3].

Долю общего разброса данных относительно среднего значения выборки, которую объясняет регрессионная зависимость, определяет скорректированный коэффициент детерминации

$$R_{\text{скор}}^2 = 1 - (1-r^2) \cdot \frac{n-1}{n-k-1} = 1 - (1-0,888^2) \cdot \frac{78}{78-4} = 0,777,$$

где $R_{\text{скор}}^2$ – скорректированный коэффициент детерминации; R^2 – коэффициент детерминации (0,888²); n – объем выборки (78); k – количество параметров в регрессионной зависимости ($k=3$).

Согласно условию Гаусса-Маркова, при построении регрессионных моделей разности между значениями экспериментов и значениями, полученными с помощью моделей, называемых остатками регрессии e , и рассматриваются как данные с равномерным законом распределения, математическим ожиданием, равным нулю, и постоянными дисперсией, среднеквадратичным отклонением, вариацией [2]. Условием гомоскедастичности для модели является постоянство вариации для всех точек выборки. Если вариация остатков различна и превышает заданное значение, то модель обладает условием гетероскедастичности.

Для проверки остатков регрессионной модели на гетероскедастичность были использованы методы Уайта [4] и Дарбина-Уотсона [2].

Первый метод требует построения параболической модели для остатков следующего вида:

$$e^2 = a_0 + a_1 c + a_2 c^2 + a_3 t + a_4 t^2 + a_5 K + a_6 K^2, \quad (1)$$

где $a_0 \div a_6$ – искомые коэффициенты.

Полученный при сравнении параболического уравнения (1) и остатков e коэффициент детерминации R^2 строит статистику Уайта

$$W = n R^2,$$

которая подчиняется закону распределения хи-квадрат.

Статистика Уайта сравнивается с критическим значением – квантилем хи-квадрат

$$W_{\text{теор}}(\alpha = 0,05; \gamma = 6),$$

где α – уровень значимости модели; γ – число степеней свободы уравнения (1).

Результаты проверки моделей на гетероскедастичность методом Уайта сведены в табл. 2.

Таблица 2

Модели и статистика Уайта при проверке на гетероскедастичность

Параметр		n	Вид уравнения	$R^2_{\text{скор}}$	W
Медь	τ	78	$\tau=1,524 c^{-0,364} t^{-0,313} K^{1,713}$	0,79	10,418
	K	78	$K=0,868 c^{0,184} t^{0,235} \tau^{0,417}$	0,877	6,457
Серебро	τ	126	$\tau=3,059 c^{-0,371} t^{-0,268} K^{1,196}$	0,908	14,139
	K	126	$K=0,431 c^{0,235} t^{0,237} \tau^{0,7}$	0,733	0,366
Цинк	τ	52	$\tau=10,586 c^{0,176} t^{-0,147} K^{1,013}$	0,878	0,101
	K	52	$K=0,127 c^{-0,156} t^{0,217} \tau^{0,724}$	0,896	1,555

Как видно из табл. 2, статистика Уайта не приемлема только для третьей модели $W_3 > W_{\text{теор}}$ ($W_{\text{теор}} = 12,59$), следовательно, указывает, что модель (3) – гетероскедастична. Остальные модели гомоскедастичны.

Второй метод Дарбина-Уотсона использует сериальную корреляцию ρ_s (автокорреляцию между остатками, отстоящими друг от друга на s шагов). Статистика Дарбина-Уотсона [2]:

$$d = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2}. \quad (2)$$

сравнивается с табличными граничными значениями d_L и d_U , которые зависят от величины выборки n , числа степеней свободы регрессионного уравнения k .

Для всех построенных моделей в табл. 1 и 2 число степеней свободы $k = 3$.

При $\rho_1 > 0$ существует три варианта выбора.

1. Если $d < d_L$, то статистика d значима и гипотеза, что все $\rho_s = 0$, отвергается.
2. Если $d > d_U$, то статистика d незначима и гипотеза, что все $\rho_s = 0$, принимается.
3. Если $d_L \leq d \leq d_U$, то определённый вывод сделать нельзя.

При отрицательном значении коэффициента автокорреляции $\rho_1 < 0$ пересчитывается значения $d' = 4 - d$ и применяются затем приведённые выше варианты выбора.

Данные проверки построенных моделей на гетероскедастичность с помощью метода Дарбина-Уотсона сведены в табл. 3.

Таблица 3

Статистика Дарбина-Уотсона при проверке на гетероскедастичность

Номер модели	n	Статистика D	d_L	d_U	ρ_1	Гомо-/гетероскедастичность	
Медь	1	78	1,488	1,41	1,565	0,182	гетеро
	2	78	2,072	1,41	1,565	-0,081	гомо
Серебро	3	126	0,693	1,53	1,63	0,643	гетеро
	4	126	0,6536	1,53	1,63	0,674	гетеро
Цинк	5	52	0,94	1,31	1,515	0,49	гетеро
	6	52	0,8225	1,31	1,515	0,58	гетеро

Только у модели (2) вариация остатков незначительна на однопроцентном уровне значимости. Остальные модели необходимо улучшать, вводя новые переменные более высоких порядков или переходя к другим нелинейным моделям. Как видно из сравнения данных табл. 2 и 3, метод Дарбина-Уотсона, используемый для проверки регрессионных моделей на гетероскедастичность, более требователен к остаткам, чем метод Уайта.

Для степени обеззараживания питьевой воды ионами меди наибольшую значимость по критерию Фишера дает уравнение вида

$$K = 1,102 - 0,331 \ln t + 0,226 (\ln t)^2 + 0,553 \ln c + 0,0084 (\ln c)^2 + 0,89 \ln \tau + 0,326 (\ln \tau)^2; \quad (3)$$

для ионов серебра вид уравнения следующий:

$$K = -0,2535 B^2 + 3,0115 B - 1,5827, \quad (4)$$

где $B = 47,041 c^{-0,31} T^{-0,811} / \tau$; $T = 0,984 e^{0,035 t}$;

для ионов цинка уравнения имеет следующий вид:

$$K = 0,364 - 0,8 c + 0,134 c^2 + 0,0121 t + 0,1127 \tau. \quad (5)$$

Расчёты по оценке значимости моделей (3) и (4) сведены в табл. 4. Модели значимы на 5-процентном уровне значимости, так как значение критерия модели превышает значение квантиля Фишера. Скорректированный коэффициент детерминации данных экспериментов и значений, рассчитанных по уравнению регрессии $R^2_{\text{скор}}$, приведён в скобках в первом столбце табл. 4. Модель (5) с использованием ионов цинка имеет меньшую значимость, чем имеет уравнение, приведённое в последней строке табл. 2.

Таблица 4

Значимость моделей для степени обеззараживания

Модель	Источник	Сумма квадратов	Число степеней	Дисперсия
Ионы меди ($R^2_{\text{скор}} = 0,968$)	Регрессия	76,019	6	12,67
	Остаток	2,56	71	0,036
	Вариация	78,58	77	-
	Критерий Фишера	Модели 351,374	Теоретический 3,73	
Ионы серебра ($R^2_{\text{скор}} = 0,949$)	Регрессия	179,63	3	59,878
	Остаток	4	122	0,033
	Вариация	183,63	125	-
	Критерий Фишера	Модели 1823	Теоретический 19,5	

Рекомендуется выбор лучшей модели осуществлять в следующей последовательности: коэффициент парной корреляции между данными эксперимента и модели, значимость уравнения регрессии по критерию Фишера, наибольшее значение скорректированного коэффициента детерминации, проверка модели на условие гетероскедастичности по критериям Уайта и Дарбина-Уотсона.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Дровозова Т.И., Игнатъев М.В., Гутенёв В.В., Денисов В.В. Основы энергосберегающей технологии фотохимического обеззараживания воды и напитков на её основе. Монография. – Новочеркасск: НГМА, 2006. – 196с. – Деп. в ВИНТИ. №1414-В2006.
2. Дрейпер Н.Р., Смит Г. Прикладной регрессионный анализ. – М.: Диалектика, 2007. – 912 с.
3. Фёрстер Э., Рёнц Б. Методы корреляционного и регрессионного анализа. – М.: Финансы и статистика, 1983. – 302 с.
4. Ниворожкина Л.И., Кокина Е.П., Кравцов В.Б. Эконометрическое моделирование с использованием пакета программ «Econometric Views». – Ростов-на-Дону: РГЭУ, 2005. – 108 с.

Игнатъев Михаил Викторович

Новочеркасская государственная мелиоративная академия (НГМА).
E-mail: mig-77@inbox.ru.
346428, г. Новочеркасск, ул. Пушкинская, 111.
Тел.: 88635222714; 88635242960.

Игнатъев Виктор Михайлович

Южно-Российский государственный технический университет (НПИ).
E-mail: mig-77@inbox.ru.
346428, г. Новочеркасск, ул. Просвещения, 132.
Тел.: 88635255628.

Ignatjev Mihail Viktorovich

Novocherkassk State Land Reclamation Academy.
E-mail: mig-77@inbox.ru.
111, Pushkinskay street, Novocherkassk, 346428, Russia.
Phone: +78635222714; +78635242960.

Ignatjev Viktor Mihailovich

Higher Professional Education «South Russia State Technical University».
E-mail: mig-77@inbox.ru
132, Prosvechniy street, Novocherkassk, 346428, Russia.
Phone: +78635255628.

УДК 621.315

В.Л. Земляков

МЕТОДЫ ОПРЕДЕЛЕНИЯ И КОНТРОЛЯ ПАРАМЕТРОВ ПЬЕЗОЭЛЕКТРИЧЕСКИХ РЕЗОНАТОРОВ

Получены соотношения и предложены новые методы определения и контроля параметров пьезоэлектрических резонаторов.

Пьезокерамический элемент; преобразователь; пьезоэлектрический резонатор; определение параметров.