

В.Е. БОНДАРЕНКО, докт. техн. наук; **О.В. ШУТЕНКО**;
Н.В. АУЛОВА; НТУ «ХПИ»

ФОРМИРОВАНИЕ ПОДМНОЖЕСТВ ОДНОРОДНЫХ ВРЕМЕННЫХ РЯДОВ ПОКАЗАТЕЛЕЙ КАЧЕСТВА ТРАНСФОРМАТОРНОГО МАСЛА

Предложен метод для формирования однородных временных рядов показателей качества трансформаторного масла на основе однофакторного дисперсионного анализа. Показана высокая эффективность предложенного метода.

The method for formation of homogeneous time of sequence of parameters of quality of transformer oil is offered on the basis of the one-factor dispersive analysis. High efficiency of the offered method is shown.

Постановка задачи. Для оценки степени старения (износа) жидкой изоляции высоковольтных трансформаторов необходимо располагать априорной (предварительной) информацией о поведении показателей качества масла на длительных интервалах эксплуатации. Единственный способ получить такого рода информацию является использование результатов периодического контроля состояния масла, в качестве обучающей выборки. Однако в процессе эксплуатации трансформаторы находятся в различных условиях (как по режимам работы, так и по качеству заливаемого масла), что приводит к наличию как мультипликативного так и аддитивного смещения между рядами показателей масла. Другими словами временные ряды показателей неоднородны. Отсутствие информации о режимах работы трансформаторов значительно усложняет проблему формирования статистически однородных обучающих выборок.

Цель статьи. Данная статья посвящена описанию метода формирования однородных массивов временных рядов показателей качества трансформаторного масла на основе однофакторного дисперсионного анализа.

Анализ публикаций. В [1] для формирования массивов однородных данных был использован дисперсионный анализ ковариационных моделей. Недостатком данного метода является детерминированный характер модели, используемой для анализа, что в конечном итоге не позволяет адекватно учесть физические особенности процесса старения масла. Данного недостатка лишен предложенный в [2] критерий максимума корреляционного отношения, который позволяет обеспечить минимальное значение дисперсии результатов измерений в каждый момент времени. К недостаткам данного подхода можно отнести отсутствие граничных значений корреляционного

отношения и как следствие отсутствие реальной возможности количественно оценить степень неоднородности. Поэтому вопросы, связанные с формированием однородных подмножеств показателей качества масла требуют дальнейшего рассмотрения.

Метод решения. Для выделения однородных временных рядов используем математический аппарат однофакторного дисперсионного анализа. В качестве фактора влияющего на изменение средних значений показателя качества масла примем время эксплуатации. Пусть \bar{x}_i – среднее значение показателя X для i -го уровня варьирования фактора. Модель исследования имеет вид:

$$x_{ij} = \bar{x}_i + \xi_{ij} \quad (1)$$

В модели (1) имеется три переменных: x_{ij} – наблюдаемые значения показателя x ; \bar{x}_i – постоянный коэффициент, представляющий эффект соответствующего уровня фактора (дрейф значения показателя обусловленный старением масла); ξ_{ij} – случайный остаток, (обусловленный как погрешностями результатов измерений, так и неоднородностью временных рядов показателей).

Для остатка выполняются следующие допущения:

1) $M[\xi_{ij}] = 0$ для всех i и j .

2) $\text{cov}(\xi_{ij}, \xi_{st}) = 0, i \neq s, j \neq t$, т.е. все случайные величины ξ_{ij} взаимно независимы.

3) $D \xi_{ij} = M[\xi_{ij}^2] = \sigma_2$ для всех i и j .

4) Случайные величины ξ_{ij} распределены по нормальному закону, т.е. $\xi_{ij} \sim NN(0, \sigma^2)$.

Обозначим через \bar{x}_i среднее арифметическое из n измерений выполненных на i -том интервале времени:

$$\bar{x}_i = \frac{1}{n} \sum_{j=1}^n x_{ij},$$

где n – объем выборочных значений на i -том интервале времени.

Обозначим через \bar{x} общую среднюю арифметическую всех результатов измерений:

$$\bar{x} = \frac{1}{m \cdot n} \sum_{i=1}^m \sum_{j=1}^n x_{ij},$$

где m – число анализируемых временных интервалов.

Известно [3], что сумма квадратов отклонений Q всех значений \bar{x}_{ij} от \bar{x}

имеет вид:

$$Q = n \cdot \sum_{i=1}^m (\bar{x}_i - \bar{x})^2 + \sum_{i=1}^m \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2, \quad (2)$$

или $Q = Q_1 + Q_2$

где Q – общая или полная сумма квадратов отклонений отдельных наблюдений от общей средней \bar{x} ; Q_1 – сумма квадратов отклонений между сериями (рассеяние за счет исследуемого фактора) – рассеивание по факторам; Q_2 – сумма квадратов отклонений внутри серий (остаточное рассеивание)

$$Q_1 = n \cdot \sum_{i=1}^m (\bar{x}_i - \bar{x})^2;$$

$$Q_2 = \sum_{i=1}^m \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2.$$

Очевидно, что однородность результатов измерений будет тем выше, чем меньше значение суммы квадратов отклонений внутри серий (Q_2). Количественную оценку степени однородности можно выполнить в рамках однофакторного дисперсионного анализа. Выдвигаем нулевую гипотезу H_0 : средние значения показателя на каждом интервале времени одинаковы $V_1 = V_2 = \dots = V_m$, что эквивалентно равенству всех x_i , $i = 1, m$. Это возможно в двух случаях: либо если рассматриваемые ряды стационарны по математическому ожиданию, либо если случайная составляющая значительно превышает систематическую. Поскольку предварительный этап обработки результатов периодического контроля [4] включает в себя отсев стационарных по математическому ожиданию временных последовательностей, то справедливость основной гипотезы H_0 будет эквивалентно превышению остаточного рассеяния над рассеянием за счет исследуемого фактора, т.е. о неоднородности результатов измерений.

Для проверки основной гипотезы H_0 оценим дисперсию результатов измерения показателя X , используя Q , Q_1 , и Q_2 , предполагая, что разброс средних \bar{x}_i относительно \bar{x} связан с однородностью выборки, т.е.

$$\sigma_{x_i}^2 = \frac{\sigma^2}{n} (H_0)$$

а) на основании уравнения $Q = \sum_{i=1}^m \sum_{j=1}^n (x_{ij} - \bar{x})^2$ имеем:

$$\sigma^2 = \frac{Q}{m \cdot n - 1} \quad (3)$$

б) с учетом уравнения для Q_1 :

$$\frac{1}{m-1} Q_1 = n \cdot \frac{1}{m-1} \sum_{i=1}^m (\bar{x}_i - \bar{x})^2 = n \cdot \sigma_x^2 = n \cdot \frac{\sigma^2}{n} = \sigma^2$$

$$\sigma_x^2 = \frac{1}{m-1} \cdot \sum_{i=1}^m (\bar{x}_i - \bar{x})^2 = \frac{Q_1}{n(m-1)}, \text{ но } \sigma_x^2 = \frac{\sigma^2}{n}, \text{ откуда}$$

$$\sigma^2 = n \cdot \sigma_x^2 = \frac{Q_1}{m-1} \quad (4)$$

в) для Q_2 имеем:

$$Q_2 = \sum_{i=1}^m \sum_{j=1}^n (\bar{x}_{ij} - \bar{x}_i)^2 = \sum_{i=1}^m \frac{(n-1)}{(n-1)} \sum_{j=1}^n (\bar{x}_{ij} - \bar{x}_i)^2 = \sum_{i=1}^m (n-1) \cdot \frac{1}{(n-1)} \sum_{j=1}^n (\bar{x}_{ij} - \bar{x}_i)^2 = m(n-1) \cdot \sigma^2,$$

тогда получим

$$\sigma^2 = \frac{Q_2}{m(n-1)} \quad (5)$$

Схема однофакторного дисперсионного анализа приведена в таблице 1.

Таблица 1 – Схема однофакторного дисперсионного анализа (параметрическая модель) (m -число строк, n -число столбцов)

Обозначение суммы квадратов	Источник изменчивости (дисперсии)	Число степеней свободы	Сумма квадратов	Средний квадрат
Q_1	Отклонения между сериями (за счет исследуемого фактора)	$m - 1$	$n \cdot \sum_{i=1}^m (\bar{x}_i - \bar{x})^2$	$\frac{Q_1}{m-1} = \sigma^2 + n \cdot \sigma_x^2$
Q_2	Отклонения внутри серии (остаточное рассеивание)	$m(n - 1)$	$\sum_{i=1}^m \sum_{j=1}^n (\bar{x}_{ij} - \bar{x}_i)^2$	$\frac{Q_2}{m(n-1)} = \sigma^2$
Q	Общая сумма квадратов отношений	$mn - 1$	$\sum_{i=1}^m \sum_{j=1}^n (\bar{x}_{ij} - \bar{x})^2$	$\frac{Q_2}{m(n-1)}$

Проверка основной гипотезы H_0 сводится к сравнению дисперсии σ^2 , получена за счет Q_1 , с той же дисперсией, но полученной за счет Q_2 , что легко проверить по отношению

$$F = \frac{\frac{Q_1}{m-1}}{\frac{Q_2}{m(n-1)}} \sim F_{(m-1); m(n-1)} \quad \text{или} \quad F = \frac{Q_1}{Q_2} \cdot \frac{m(n-1)}{(m-1)} \quad (6)$$

Статистика F является случайной величиной имеющей F -распределение с $(m - 1)$ и $m(n - 1)$ степенями свободы. Если при заданном α (обычно $\alpha = 0,05$) $F > F_{\alpha}$ (преобладает числитель), то фактор значим. Таким образом, временные ряды могут считаться статистически однородными, если они обеспечивают максимальное значение F -статистики в рамках однофакторного дисперсионного анализа.

Практические результаты. Проиллюстрируем процедуру формирования однородных подмножеств данных на примере такого показателя как цвет трансформаторного масла. На рис. 1, *а* приведен массив исходных данных, на рис 1, *б* три подмножества однородных данных полученных из данного массива на основе однофакторного дисперсионного анализа. Данные подмножества были получены на основе критерия максимума значения F -статистики. Число временных рядов показателя образующих данное подмножество, объем выборочных значений и результаты дисперсионного разложения приведены в табл. 2.

Таблица 2 – Результаты дисперсионного разложения для однородных временных рядов цвета трансформаторного масла

Результаты анализа	Массивы данных		
	M1	M2	M3
Объем выборочных значений			
N	80	40	54
nt	6	3	3
Суммы дисперсионного разложения			
Q_0	65,050	86,375	46,967592592
Q_1	55,34915475	75,501923077	44,385449735
Q_2	9,7008452504	10,873076923	2,5821428571
Число степеней свободы			
v_0	79	39	53
v_1	6	4	8
v_2	73	35	45
Значения F -статистики			
$F_{\text{расч.}}$	69,42	60,759	96,69
$F_{\text{крит.}}$	2,22	2,64	2,15
Значение коэффициента парной корреляции ρ_{t-x} и корреляционного отношения η_{t-x}			
ρ_{t-x}	0,916	0,923	0,929
η_{t-x}	0,921	0,945	0,963

Кроме результатов дисперсионного анализа в табл. 2 приведены также значения коэффициента парной корреляции показателя на время эксплуатации ρ_{r-x} и корреляционного отношения η_{r-x} . Как видно из табл. 2 для всех трех подмножеств данных значение систематической составляющей намного превосходит значение остаточного рассеивания, о чем свидетельствуют рассчитанные значения F -статистик по каждому из подмножеств. Это говорит о том, что полученные подмножества данных являются однородными.

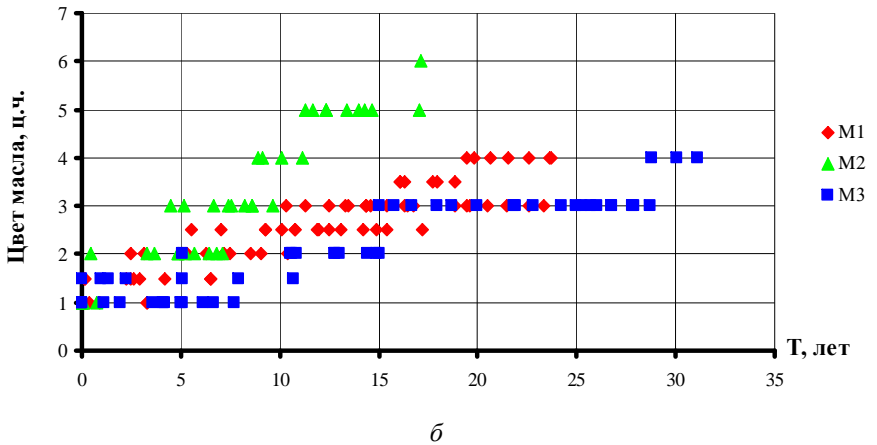
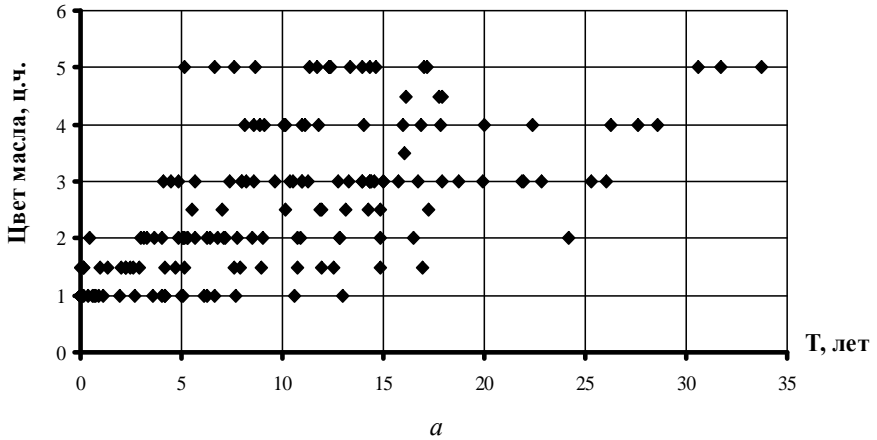


Рис.1. Однородные подмножества цвета трансформаторного масла:
 а – массив исходных данных после отсева искаженных и стационарных временных последовательностей $N = 175, nt = 14$;

b – однородные подмножества полученные в результате однофакторного дисперсионного анализа

Важным, является то обстоятельство что значения корреляционного отношения и коэффициента парной корреляции указывают на наличие практически функциональной связи между временем эксплуатации и цветом масла, что подтверждает высокую эффективность процедуры формирования однородных подмножеств по критерию максимума корреляционного отношения. Другими словами результаты формирования однородных подмножеств методом дисперсионного анализа и по критерию максимума корреляционного отношения совпадают.

Выводы.

1. Предложено использовать однофакторный дисперсионный анализ для формирования однородных массивов показателей качества масла. Практическая реализация метода на примере цвета трансформаторного масла, показала его высокую эффективность.

2. Значимое влияние времени эксплуатации на дрейф значений цвета трансформаторного масла, свидетельствует об интенсивном развитии процессов старения.

3. Полученные значения коэффициента парной корреляции и корреляционного отношения свидетельствуют о наличии практически функциональной связи между цветом масла и временем эксплуатации.

4. Сравнительный анализ полученных значений F -статистик и значений корреляционного отношения показал, что результаты формирования однородных подмножеств методом дисперсионного анализа и по критерию максимума корреляционного отношения совпадают.

Направления дальнейших исследований. Дальнейшим этапом работы является использование полученных однородных подмножеств показателей масла в качестве обучающих выборок для формирования решающего правила принятия решений при оценке степени старения масла.

Список литературы: 1. *Бондаренко В. Е., Шутенко О. В.* Метод выделения оптимального числа наиболее информативных показателей качества при синтезе регрессионных моделей для оценки степени старения жидкой изоляции трансформаторов // Технічна електродинаміка, Тематичний випуск Силова електроніка та енергоефективність, Частина 5, Київ, 2006., с. 88-93. 2. *Шутенко О. В.* Формирование однородных массивов показателей качества трансформаторного масла в условиях априорной неопределенности результатов испытаний // Інтегровані технології та енергозбереження. Щоквартальний науково-практичний журнал. – Харків: НТУ «ХПИ», 2006. – № 4. – С. 42–50. 3. *Джонсон Н., Лион Ф.* Статистика и планирование эксперимента в технике и науке. – М.: Мир, 1981. – 520 с. 4. *Бондаренко В.Е., Щапов П.Ф., Шутенко О. В.* Повышение эффективности эксплуатационного измерительного контроля трансформаторных масел. Монография. – Харьков.: НТУ «ХПИ», 2007. – 452 с.

Поступила в редколлегию 16.07.08