

УДК 004.942

В.Ф. МИРГОРОД<sup>1</sup>, Е.В. ДЕРЕНГ<sup>2</sup><sup>1</sup> АО «Элемент», Одесса, Украина<sup>2</sup> Институт проблем моделирования в энергетике НАН Украины, Киев, Украина

## ВЫДЕЛЕНИЕ КЛАСТЕРОВ ПРИ МНОГОМЕРНОМ ТРЕНДОВОМ АНАЛИЗЕ

*Рассмотрен подход к повышению надежности диагностических выводов о техническом состоянии силовых и энергетических установок на базе газотурбинных двигателей, основанный на построении статических моделей изменения состояния в виде отклонений от исходных (либо эталонных) статических характеристик и анализе двумерных комплексных трендов временных рядов таких отклонений. Комплексный трендовый анализ выполняется на основе метода главных компонент. Оптимальная аппроксимация выделенной трендовой компоненты выполняется методом наименьших квадратов. Предложен критерий разделения трендов на различные группы (кластеры) по уровню статистической связи между ними.*

**Ключевые слова:** техническое состояние, диагностическая модель, временные ряды, трендовый анализ.

### Введение

Управление жизненным циклом силовых и энергетических установок на основе газотурбинных двигателей в настоящее время осуществляется в соответствии со стратегией эксплуатации по их техническому состоянию. Последовательная реализация этапов перехода к указанной стратегии представляет собой важную научно-техническую проблему. Разрешение указанной проблемы обуславливает необходимость разработки новых, более совершенных методов и технических средств контроля состояния силовых и энергетических установок, как объектов диагностирования. Актуальная научно-прикладная задача состоит в создании методов и программных средств такого контроля, позволяющих дать статистически обоснованную оценку и прогноз изменения технического состояния на основе информации, фиксируемой аппаратными средствами систем технической диагностики (СТД).

### 1. Формулирование проблемы

Исходные данные в системах технической диагностики формируются в виде временных рядов: временных срезов многомерной совокупности параметров на заданном промежутке времени эксплуатации.

Объектом исследования являются процессы изменения технического состояния силовых и энергетических установок в длительной эксплуатации.

Предметом данного исследования являются статистические модели (СМ) порождения данных, методы трендового контроля и анализа, позволяющие установить факт наличия тренда и закономерности его развития [1, 2, 4, 6].

В системах технической диагностики находят широкое применение различные трендовые статистики (Хальда - Аббе, кумулятивных сумм, F-критерий и др.) [4]. Известные методы трендового контроля позволяют установить лишь факт отсутствия тренда на заданном уровне значимости [8]. В известных методах не учитывается наличие естественного тренда, вызванного выработкой ресурса. Поэтому методы трендового контроля могут быть использованы лишь для краткосрочного анализа, а опыт их применения свидетельствует о недопустимо высоком уровне ложных тревог (ошибок второго рода), что снижает доверие к результатам диагностических решений. Более совершенными являются методы трендового анализа, позволяющие восстановить трендовую компоненту и выполнить анализ тенденций ее развития. Однако методы трендового анализа предлагаются в настоящее время лишь в скалярном варианте [5, 6]. Разработка методов многомерного трендового анализа является важной для теории и необходимой для прикладных приложений, но еще нерешенной в настоящее время задачей.

Целью настоящего исследования является выделение связанных трендов в многомерных временных рядах на основе оптимальной аппроксимации трендовой компоненты временного ряда, образованного двумерной совокупностью параметров регистрации технического состояния силовых и энергетических установок в длительной эксплуатации.

### 2. Решение проблемы

Особенность регистрации технического состояния силовых и энергетических установок состоит в том, что в процессе эксплуатации добавляются новые данные (временные срезы), которые образуют

расширяющуюся совокупность данных (термогазо-динамических и вибропараметров):

$$\begin{aligned}\bar{x}_1 &= [x_1 x_2 \dots x_n], \\ \bar{x}_2 &= [x_1 x_2 \dots x_n x_{n+1}], \\ &\dots \\ \bar{x}_k &= [x_1 x_2 \dots x_n x_{n+k}].\end{aligned}\quad (1)$$

Известно, что информация об изменении статистических свойств временного ряда содержится в прямоугольной траекторной матрице [3, 6] размером  $n \times k$

$$\begin{aligned}\bar{x}_1 &= [x_1 x_2 \dots x_n], \\ \bar{x}_2 &= [x_2 x_3 \dots x_{n+1}], \\ &\dots \\ \bar{x}_k &= [x_k x_{k+1} \dots x_{n+k-1}].\end{aligned}\quad (2)$$

В качестве статистической модели порождения данных при трендовом анализе принимается следующая модель аддитивной совокупности трендовой, циклической и помеховой (шумовой) компонент [9]:

$$\begin{aligned}\bar{x}_k &= [x_k x_{k+1} x_{k+2} \dots x_{k+n-1}] = \\ &= \bar{x}_{tr} + \bar{x}_{cicle} + \bar{x}_{noise}.\end{aligned}\quad (3)$$

Задача оптимальной аппроксимации трендовой компоненты для совокупности исходных данных (1) и статистической модели порождения данных вида (3) решена в [9].

В отличие от известных подходов, с целью выделения совместных трендов и трендов различия, предлагается следующее попарное комплексное объединение временных рядов всех регистрируемых параметров  $\bar{x}_k = \bar{y}_k + j\bar{z}_k$ . Для каждой из полученной таким образом совокупности прямоугольных комплексных траекторных матриц, в соответствии с методом главных компонент, далее решается стандартная задача на собственные значения и собственные векторы:

$$X_t X_t^H \bar{u}_i = \sigma_i \bar{u}_i = \sigma_i \bar{u}_i, \quad (4)$$

где индекс  $H$  означает операцию комплексного сопряжения и транспонирования, а столбцы  $\bar{u}_i$  образуют матрицу из ортогональных векторов матрицы  $X_t X_t^H$ .

Матрица главных компонент объединенного комплексного временного ряда имеет следующий вид [1, 5]

$$F = U^H X_t, \quad (5)$$

строки которой  $\bar{f}_j$  (главные компоненты) упорядочены по величине собственных чисел (выборочным дисперсиям) матрицы корреляций  $X_t X_t^H$ .

Каждая из строк траекторной матрицы (2) может быть представлена в виде следующего разложения по главным компонентам (5):

$$\bar{x}_s = \sum_{i=1}^k b_{si} \bar{f}_i, \quad (6)$$

где  $s = \overline{1, k}$ ,  $b_{si}$  – коэффициенты влияния [1], определяемые решением переопределенной ( $n > k$ ) системы линейных алгебраических уравнений

$$F^H \bar{b}_s = \bar{x}_s. \quad (7)$$

Решение (7) может быть найдено с использованием псевдообратной матрицы [5]:

$$\bar{b}_s^H = (FF^H)^{-1} F \bar{x}_s^H. \quad (8)$$

Такое решение является оптимальным по критерию наименьших квадратов [5], поскольку одновременно минимизирует норму ошибки  $\|F^H \bar{b}_s - \bar{x}_s\|$  для любой заданной строки (1) и длину

искомого вектора  $\bar{b}_s$ . Так как  $(FF^H) = \text{diag}\{\lambda_i\}, i = \overline{1, k}$ , то из (7) следует соотношение для определения оптимальных, в указанном смысле, коэффициентов влияния

$$b_{sj} = \bar{x}_s \bar{f}_j^H / (\bar{f}_j \bar{f}_j^H) = \lambda_j^{-1} \bar{x}_s \bar{f}_j^H \quad (9)$$

где  $\bar{f}_j$  – строки матрицы главных компонент.

Если собственное число  $\lambda_1 = \lambda_{\max}$  соответствует компоненте, которая в исходной выборке является трендовой, согласно (7) получаем ее представление в виде

$$\bar{x}_{tr,s} = b_{s1} \bar{f}_1, \quad (10)$$

где  $b_{s1} = \lambda_{\max}^{-1} \bar{x}_s \bar{f}_1^H$ .

Трендовая компонента (10) содержит вещественную и мнимую часть. Если объединяемые временные ряды  $\bar{y}_k$  и  $\bar{z}_k$  не имеют статистически значимой связи, то на комплексной плоскости трендовой компоненты все значения аргумента комплексного вектора распределены приблизительно равномерно с центром в нуле.

Заметим, что если временные ряды  $\bar{y}_k$  и  $\bar{z}_k$  отличаются лишь масштабным коэффициентом, то на комплексной плоскости трендовой компоненты все ее значения принадлежат биссектрисе I – III квадрантов, то есть аргумент комплексного вектора равен 45 градусов. Потому мерой статистической связи трендов в объединяемых временных рядах может быть выбрана близость линии регрессии на комплексной плоскости  $\text{Im}\{\bar{x}\} - \text{Re}\{\bar{x}\}$  к указанной биссектрисе.

На практике более удобным является использовать для анализа непосредственно комплексные компоненты, поскольку они являются нормированными.

Для проверки обоснованности предложенного подхода выполнено решение задачи, рассмотренной в [9], по оценке технического состояния авиационного газотурбинного двигателя в длительной эксплуатации. На рис. 1, 2 и 3 представлены плоскости комплексных значений первой главной компоненты, соответствующей максимальному собственному числу матрицы  $X_t X_t^H$  для различных пар объединяемых параметров двигателя.

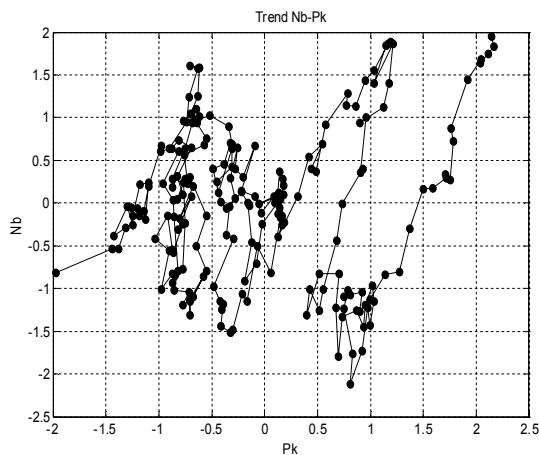


Рис. 1. Плоскость первой главной компоненты для комплексного объединения отклонений оборотов турбины высокого давления и степени повышения давления

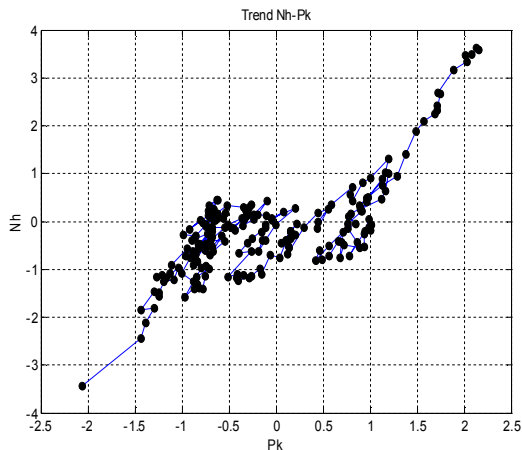


Рис. 2. Плоскость первой главной компоненты для комплексного объединения для отклонений оборотов турбины низкого давления и степени повышения давления

Предлагаемый подход позволяет кластеризовать полученные в [9] тренды на две статистически связанные группы. Такими группами являются тренды: обороты турбины низкого давления – степень повышения давления за компрессором, температура газов – расход топлива.

Действительно, в автоматическую регулирование указанного двигателя заложен закон поддержания

приведенных оборотов турбины высокого давления, поэтому при деградации характеристик компрессора высокого давления для поддержания требуемых оборотов необходимо увеличивать расход топлива. В свою очередь такое увеличение расхода топлива влечет увеличение температуры газов за турбиной, что обуславливает связанность указанных трендов. С другой стороны, падение степени повышения давления приводит к отклонению оборотов турбины низкого давления от номинальных значений. Поэтому формируются две группы параметров двигателя. Тренды параметров между группами не имеют значимой статистической связи.

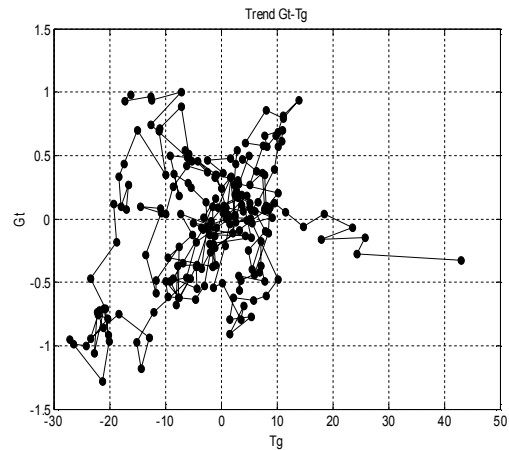


Рис. 3. Плоскость первой главной компоненты для комплексного объединения для отклонений расхода топлива и отклонений температуры газов за турбиной

## Заключение

Основные результаты работы заключаются в том, что предлагается и обосновывается новый метод двумерного трендового анализа временных рядов, образованных объединением параметров регистрации сложных технических объектов на примере оценки технического состояния силовых и энергетических установок в их длительной эксплуатации.

Предлагаемый метод позволяет разделить многомерные временные ряды, образованные параметрами регистрации состояния ГТД, на группы (кластеры), имеющие существенную статистическую связь, что позволяет установить соответствующие параметры трендовых статистик для обнаружения указанных трендов.

Разработаны и реализованы программные средства для компьютерной реализации предлагаемого подхода на основе интерактивной среды MATLAB и ее приложений.

На основе предложенного подхода решена прикладная задача кластеризации многомерных

трендов для діагностики авіаційного двигателя в процесі довготривалої експлуатації.

Перспективи дальніших досліджень заключаються в розповсюдженні пропонуваного підходу на тренди, розмірності яких більше двох, а також в розв'язанні задач прогнозування.

### Література

1. Прикладна статистика. Класифікація і зниження розмірності [Текст] / С.А. Айвазян, І.С. Енюков, В.М. Бухштабер, Л.Д. Мешалкин. – М.: Финансы и статистика, 1989. – 607 с.
2. Бендат, Дж. Прикладний аналіз випадкових даних [Текст] / Дж. Бендат, А. Пирсон. – М.: Мир, 1989. – 540 с.
3. Главные компоненты временных рядов: метод "Гусеница" [Текст] / под ред. Д.Л. Данилова, А.А. Жиглявского. – СПб.: СПТУ, 1997. – 270 с.
4. Синтез систем управления и диагностирования газотурбинных двигателей [Текст] / С.В. Епифанов, В.И. Кузнецов, И.И. Богаенко и др. – К.: Техника, 1998. – 312 с.

5. Марпл, мл. С.Л. Цифровой спектральный анализ и его приложения [Текст] / С.Л. Марпл мл. – М.: Мир, 1990. – 584 с.

6. Миргород, В.Ф. Применение диагностических моделей и методов трендового анализа для оценки технического состояния газотурбинных двигателей [Текст] / В.Ф. Миргород, Г.С. Ранченко, В.М. Кравченко // *Авиационно-космическая техника и технология*. – 2008. – № 9(56). – С. 192–197.

7. Elsner I.B. Singular Spectrum Analysis: A New Tool in Time Series Analysis [Text] / I.B. Elsner, A.A. Tsonis. – New York, London: Plenum Press, 1996. – 164 p.

8. Perron, P. Trend and Random Walks in Macroeconomic Time Series: Further Evidence from a New Approach [Text] / P. Perron // *Journal of Economic Dynamic and Control*. – 2000. – No. 12. – P. 297–332.

9. Миргород, В.Ф. Оптимальная аппроксимация трендовой компоненты временного ряда [Текст] / В.Ф. Миргород, И.М. Гвоздева // *Електро-технічні та комп'ютерні системи*. – 2011. – № 04(80). – С. 121–125.

Поступила в редакцію 6.06.2013, розглянута на редколегії 13.06.2013

**Рецензент:** д-р техн. наук, проф. С.Г. Антошук, Одеський національний політехнічний університет, Одеса.

### ВИДІЛЕННЯ КЛАСТЕРІВ ПРИ БАГАТОВИМІРНОМУ ТРЕНДОВОМУ АНАЛІЗІ

*В.Ф. Миргород, Є.В. Деренг*

Розглянутий підхід до підвищення надійності діагностичних висновків про технічний стан силових і енергетичних установок на базі газотурбінних двигунів, заснований на побудові діагностичних моделей зміни стану у вигляді відхилень від початкових (або еталонних) статичних характеристик і аналізі двовимірних комплексних трендів тимчасових рядів таких відхилень. Комплексний трендовий аналіз виконується на основі методу головних компонент. Оптимальна апроксимація виділеної трендової компоненти виконується методом найменших квадратів. Запропонований критерій розподілу трендів на різні групи (кластери) по рівню статистичного зв'язку між ними.

**Ключові слова:** технічний стан, діагностична модель, тимчасові ряди, трендовий аналіз.

### CLUSTERS EXTRACTION AT MULTIDIMENSIONAL TREND ANALYSIS

*V.F. Mirgorod, E.V. Dereng*

An approach to improve the reliability of diagnostic conclusions on the technical state of power and power plants based on gas turbine engines is considered. The approach is based on diagnostic models constructing for the state changes in the form of deviations from the baseline (or reference) static characteristics and the analysis of two-dimensional complex trends of time-series trends of such deviations. Comprehensive trend analysis is performed on base of principal component method. The optimal approximation of a selected trend components is performed by least squares methods. A criterion for the separation of trends for different groups (clusters) in terms of the statistical interaction among them is proposed.

**Key words:** technical state, the diagnostic model, time series, trend analysis.

**Миргород Владимир Федорович** – ведучий научний співробітник АО «Елемент», Одеса, Україна, e-mail: odessa@element.od.ua

**Деренг Евгения Владимировна** – аспірант Інститута проблем моделювання в енергетиці ім. Г.Е. Пухова НАН України, Київ, Україна, e-mail: odessa@element.od.ua