

2. Куц Ю.В., Щербак Л.М. Статистична фазометрія. Наукова монографія. – Тернопіль: ТДТУ ім. І. Пулюя, 2009. – 384 с.
3. Norden E. Huang. Hilbert-Huang transform and its Applications// World Scientific Publishing Co. Pte.Ltd, 2005 – 324p.
4. Patrick Flandrin. On Empirical mode decomposition and its algorithms// IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing NSIP-03, Grado (I), 2003 – p. 8-87.
5. Dergunov O. V., Empirical mode decomposition in signal analysis / Dergunov O. V., Trots V. M., Kuts Y. V., Shcherbak L. M. // Aviation in the XXI-st century 2010, (Kyiv, 21-23 September 2010) [etc.]. – К.: НАУ, 2010. – P. 12.21 – 12.26.
6. Al. Dergunov. Hilbert-Huang transform in signal analysis / O. Bliznuk., Al. Dergunov, Y. Kuts., A. Shekero., S. Shengur // «Defectoscopy'11»: proc. XXVI International conference: – Sofia: 2011. – P. 25-27.
7. Держунов О.В. Перспективи використання перетворення Гільберта-хуанга в задачах неруйнівного контролю / О.В. Держунов, Ю.В. Куц // Неруйнівний контроль та технічна діагностика: 7-а Національна НТК, 20-23 листопада 2012 р.: тези доп. – Київ, 2012.– С. 243 – 245

Поступила 29.09.2016 р.

УДК 519:2+512.643

В.А. Готович, О.Б. Назаревич, к.т.н., Г.В. Шимчук, ТДТУ, м. Тернопіль
Т.Л. Щербак, к.т.н., НАУ, м. Київ

СТАТИСТИЧНИЙ АНАЛІЗ ПРОЦЕСУ ГАЗОСПОЖИВАННЯ МІСТА НА РІЧНОМУ ІНТЕРВАЛІ СПОСТЕРЕЖЕННЯ

Abstract. Reasonable mathematical model of the city gas consumption at the annual observation range in the form of piece-periodic random process. On the basis of the method "Caterpillar-SSA» conducted a statistical analysis of a number of annual gas consumption of the city and the allocation of its main components: the annual trend, periodic components and the implementation of a random process.

Вступ. В Україні, як і в інших країнах світу науково-технічні проблеми використання енергоносіїв в умовах економічної кризи на сьогодні набули особливої актуальності. Це знайшло своє відображення у значній кількості публікацій, де наведені результати досліджень проблем газоспоживання, в тому числі [1-3].

Але в більшості робіт розглядається загальний підхід статистичного аналізу процесу газоспоживання, які в певній мірі не адоптовані до конкретних задач досліджень, а саме не вказується часовий інтервал спостережень (вимірювань) значень процесу газоспоживання, похибки вимірювань, об'єм спожитого газу та інші.

В даній роботі наведені результати конкретного випадку статистичного

аналізу часового рідущ процесу газоспоживання міста.

Постоновка завдання. Запропонувати модель газоспоживання, яка б враховувала дії значної кількості фактів, включаючи добову стохастичну періодичність процесу газоспоживання та його сезонність упродовж року, провести статистичний аналіз даних вимірювань газоспоживання міста на річному інтервалі часу з застосуванням запропонованої моделі, виділити відомі компоненти (складові) процесу газоспоживання, які мають чітку фізичну інтерпритацію.

Основні результати. На основі аналізу існуючих моделей процесу газоспоживання запропоновано математична модель, яка дозволяє враховувати не лише добову циклічність процесу газоспоживання, а й особливості роботи газотранспортної системи (ГТС) міста, що визначаються режимами її роботи в межах року. В загальному вигляді математична модель процесу газоспоживання міста представлена у вигляді:

$$\xi(\omega, t) = a(t) + \zeta(\omega, t), t \in T, \omega \in \Omega \quad (1)$$

де $a(t)$ – деяка детермінована функція (тренд, що враховує сезонність процесу газоспоживання), а кусково-однорідний періодичний випадковий процес $\zeta(\omega, t)$ описається виразом [4]

$$\zeta(\omega, t) = \sum_{k=1}^m \zeta_k(\omega, t) \cdot I(t, T_k), T_k \cap T_j = \emptyset, k \neq j, \bigcup_{k=1}^m T_k = T \quad (2)$$

де $\{\zeta_k(\omega, t), k = \overline{1, m}, t \in T_k\}$ – послідовність однорідних періодичних випадкових процесів, $I(\tau, T_k)$ – індикаторна функція задається у вигляді:

$$I(\tau, T_k) = \begin{cases} 1, & \tau \in T_k \\ 0, & \tau \notin T_k \end{cases} \quad (3)$$

Таким чином кусково-однорідний періодичний випадковий процес $\zeta(\omega, t)$ задається у вигляді суми однорідних випадкових періодичних процесів, кожний з яких задається на своєму часовому інтервалі.

Зупинимось на етапах оцінювання компонент такої математичної моделі. Для дослідження використано погодинні дані вимірювань процесу газоспоживання міста Тернопіль за 2009 рік. Точність вимірювання даних газоспоживання визначається відповідними характеристиками точності витратомірного комплексу «Флоутек» ГТС м. Тернополя.

На першому етапі проведено попередню обробку статистичних даних з метою вилучення наявних аномальних даних вимірювань. На рис. 1 наведено графік реалізації процесу газоспоживання міста Тернополя з кроком накопичення одна доба на річному інтервалі за 2009 рік.

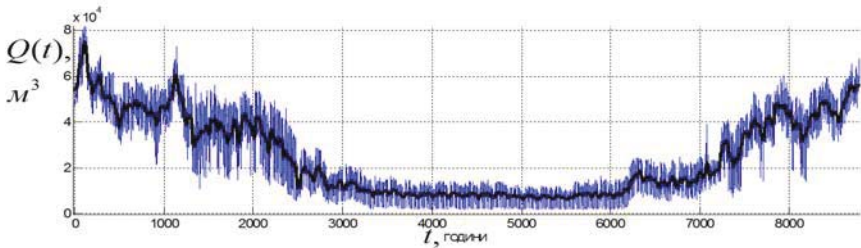


Рис. 1 – Графік реалізації процесу газоспоживання з кроком накопичення одна година на річному інтервалі за 2009 рік

На другому етапі обгрунтовано вибір методу «Гусениця-SSA» для оцінювання компонент математичної моделі (1). Найвні дані погодинних обсягів споживання газу вважаємо реалізацією нестационарного випадкового процесу, динаміка якого залежить від значної кількості випадкових факторів,

Для статистичного аналізу часових рядів нестационарних процесів використовується ряд методів: метод групового урахування аргументів (МГУА); методи, що ґрунтуються на авторегресії з інтегруванням ковзним середнім (ARIMA); емпірична модова декомпозиція (EMD); метод аналізу головних компонент; метод аналізу сингулярного спектру («Гусениця-SSA») та інші. В даній роботі обгрунтовано використання методу «Гусениця-SSA» [5] на основі наступних переваг у порівнянні з іншими методами опрацювання часових рядів а саме:

- Розклад досліджуваного сигналу за базисом власних сингулярних векторів траекторної матриці, який полягає у перетворенні одновимірного часового ряду в багатовимірний з подальшим застосуванням до отриманого багатовимірного часового ряду методу головних компонент.

- Можливість адаптації до задач дослідження, тому його можна адаптувати до моделі газоспоживання (1), що дозволяє оцінити основні компоненти моделі – тренд та послідовність однорідних періодичних випадкових процесів, які мають фізичну інтерпретацію.

- Інтерактивність методу, що дає можливість опрацювання часових рядів в режимі діалогу «машина-дослідник» з можливістю вибору довжини ковзного вікна даних.

- Інформаційне забезпечення методу, так як існує низка сертифікованих програмних продуктів (розроблених в середовищах Matlab, Octave, R-Project), що реалізують алгоритм методу «Гусениця-SSA».

На третьому етапі часовий ряд газоспоживання $Q(t)$, заданий на річному інтервалі розбивається на дві основні компоненти з використанням методу «Гусениця-SSA»:

$$Q(t) = A(t) + Z(t), \quad t \in [1, T], \quad T = 8760 \text{ год} \quad (4)$$

де $A(t)$ – оцінка річного тренду, що характеризує споживання газу протягом

року, частка якого становить близько 95% (2009 рік); $Z(t)$ – дискретний кусково-однорідний періодичний випадковий процес, який визначається вилученням із досліджуваного часового ряду $Q(t)$ тренду $A(t)$. Вважаємо, що

$$Z(t) = \sum_{k=1}^m Z_k(t) \cdot I(t, T_k), \quad T_k \subset T, t \in T, T_k \cap T_j = \emptyset, k \neq j, \bigcup_{k=1}^m T_k = T \quad (5)$$

де $Z_k(t), t \in T_k, k=\overline{1, m}$ – реалізація однорідного періодичного процесу $\zeta_k(\omega, \tau)$. Структура і характеристики компонент $Z(t)$ визначаються на наступних етапах.

На четвертому етапі на основі аналізу результатів застосування методу «Гусениця-SSA», запропоновано представити компоненту $Z(t)$ у вигляді:

$$Z(t) = B(t) + X(t), \quad (6)$$

де $B(t)$ – кусково періодична функції:

$$B(t) = \sum_{k=1}^m B_k(t) \cdot I(t, T_k), \quad k=\overline{1, m}, t \in T, T_k \cap T_j = \emptyset, k \neq j, \bigcup_{k=1}^m T_k = T,$$

$B_k(t)$ – детермінована періодична функція;

$X(t)$ – стохастичний залишок, який можна представити у вигляді кусково стаціонарного процесу:

$$X(t) = \sum_{k=1}^m X_k(t) \cdot I(t, T_k), \quad k=\overline{1, m}, t \in T, T_k \cap T_j = \emptyset, k \neq j, \bigcup_{k=1}^m T_k = T,$$

де $X_k(t)$ – стаціонарний випадковий процес.

Для використання моделі на практиці необхідно запропонувати метод визначення меж інтервалів $T_k, k=\overline{1, m}$.

На п'ятому етапі проведено розбиття річного інтервалу спостереження на підінтервали $T_k, k=\overline{1, m}$. Межі інтервалів можна визначити як моменти розладки кусково стаціонарного процесу.

Для визначення точок «розладки» часового ряду процесу газоспоживання використаємо метод Бродського-Дарховського [6-12], алгоритм якого детально описаний в [6].

Пошук моментів розладки проводився на основі стохастичного залишку $X(t)$ з використанням методу Бродського-Дарховського, що використовує модифіковану статистику Колмогорова-Смірнова:

$$Y(t) = \left[\left(1 - \frac{t}{N} \right) \cdot \left(\frac{t}{N} \right) \right]^V \left[\frac{1}{t} \sum_{l=1}^t X'(l) - \frac{1}{N-t} \sum_{l=t+1}^N X'(l) \right], \quad (7)$$

де $t = \overline{1, N}$ год, $t \in T$, $X'(l) = X^2(l)$ – квадрат стохастичного залишку $X^2(l)$ для l -ої точки, виділеного на етапі статистичного опрацювання погодинного часового ряду газоспоживання методом «Гусениця-SSA»; $N=8760$ год – довжина погодинного часового ряду газоспоживання на річному інтервалі спостереження; $0 < V < 1$ – параметр налаштування статистики на реальні дані, що визначає чутливість хибного спрацювання при пошуку моменту різкої зміни дисперсії стохастичного залишку $X(t)$. Для визначення значень параметра V для досліджуваного часового ряду процесу газоспоживання, було розроблене програмне забезпечення (в середовищі C), використання якого дозволило обчислювати значення параметра V в межах наступного інтервалу $V=0.3-0.4$. Модифікована статистика $Y(t)$ реагує на моменти зміни дисперсії стохастичного залишку $X(t)$.

Коротко опишемо результат застосування. Значення $X(t)$ погодинного залишку підносимо до квадрату $X(t)^2$ це дозволить статистиці реагувати на зміну дисперсії. Результат застосування статистики Бродського-Дарховського наглядно представлено на Рис. 2.

Далі продемонструємо розбиття на реальному річному часовому ряді газоспоживання міста Тернополя за 2009 рік.

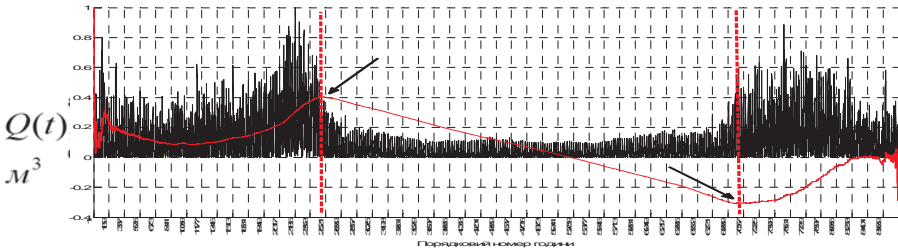


Рис. 2 – Результат застосування статистики Бродського-Дарховського для погодинного залишку $X^2(t)$

Отримані точки «розладки» в годинах округлюємо до найжчого цілого числа кратного 24 год – $\text{ceil}(x/24)*24$ (тобто на початок доби).

В результаті такого поділу отримуємо три ділянки: 1) [1:2520], 2) [2520:7008], 3) [7008:8760].

В результаті ми отримуємо п'ять окремих інтервали розбиття погодинного залишку газоспоживання міста на річному інтервалі: ділянка 1 [1:2064], ділянка 2 [2064:2520], ділянка 3 [2520:7008], ділянка 3 [7008:8232], ділянка 5 [7008:8760].

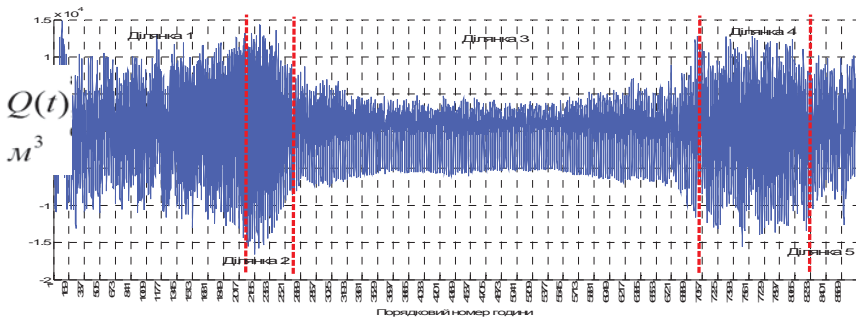


Рис. 3 – Результат поділу щоденного залишку $X(t)$ часового ряду газоспоживання на стаціонарні ділянки $X_k(t)$.

Як бачимо з Рис. 3, розбиття на 5 ділянок є природнім поділом на річні сезони: зима, весна, літо, осінь та знову зима і відповідно режими роботи газотранспортної мережі міста (опалюваний, неопалювальний, перехідні).

Як показують аналіз реальних даних таке розбиття на п'ять ділянок вказує на неспівпадіння календарної дати включення-виключення центрального опалення, а має дві додаткові перехідні ділянки: зима-весна та осінь-зима.

В результаті обробки даних газоспоживання міста Тернополя за 2009 рік, часовий ряд був розбитий на п'ять інтервалів, що відповідають різним режимам роботи газотранспортної системи міста:

- 1) перший – опалювальний режим 1, де $t \in \overline{1, 2064}$ год (86 діб);
- 2) другий – перехідний режим зима-весна, де $t \in \overline{2064, 2520}$ год (19 доба);
- 3) третій – неопалювальний режим, де $t \in \overline{2520, 7008}$ год (187 доби);
- 4) четвертий – перехідний режим осінь-зима, де $t \in \overline{7008, 8232}$ год (51 доба);
- 5) п'ятий – опалювальний режим 2, де $t \in \overline{8185, 8760}$ год (22 доби).

На шостому етапі було проведено апроксимацію компонент $A_k(t)$ детерміновано тренду $A(t)$ лінійними функціями на кожному з виділених інтервалів (рис. 4). Так для кожної ділянки $A_k(t)$ тренда $A(t)$ на основі використання методу найменших квадратів побудована лінійна залежність значень газоспоживання у вигляді:

$$A_k(t) = C_k + D_k t, \quad k = \overline{1, 5}, \quad t \in T_k \quad (8)$$

де C_k і D_k – відповідні коефіцієнти лінійної залежності були обчислені на основі реальних даних вимірювання газоспоживання.

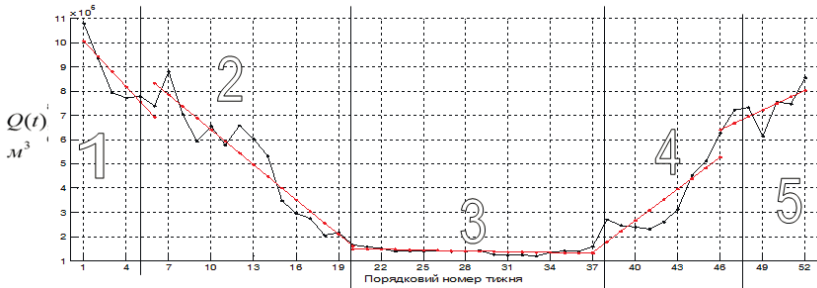


Рис. 4 – Графік апроксимації тренду $A(t)$ лінійними функціями

Розбиття часового інтервалу на підінтервали дозволило також проводити статистичний аналіз реалізації компонент $Z_k(t)$ кусково-однорідного періодичного процесу $\zeta(\omega, t)$.

Отримані результати попередніх досліджень дають можливість записати загальну модель газоспоживання міста на річному інтервалі.

На сьомому етапі з врахуванням вище наведеного загальна модель газоспоживання міста на річному інтервалі часу T має вигляд:

$$\xi(\omega, t) = \sum_{k=1}^m (a_k(t) + \zeta_k(\omega, t)) \cdot I(t, T_k), \quad (9)$$

$$k = \overline{1, m}, t \in T, T_k \cap T_j = \emptyset, k \neq j, \bigcup_{k=1}^m T_k = T$$

де $a_k(t) = c_k + d_k t$ – компоненти кусково лінійного детермінованого тренду $a(t)$;

$\zeta_k(\omega, t) = b_k(t) + \eta_k(\omega, t)$ – однорідний періодичний випадковий процес з періодом 24 год; $b_k(t)$ – детермінована періодична функція; $\eta_k(\omega, t)$ – стаціонарний випадковий процес;

m – кількість інтервалів розбиття; $I(t, T_k)$ – індикаторна функція задається у вигляді (3).

Для досліджуваного часового ряду газоспоживання міста Тернопіль за 2009 рік $m=5$ – кількість ділянок розбиття на річному інтервалі спостереження.

Висновки

Розроблено адитивну математичну модель процесу газоспоживання міста, що є сумою детермінованого річного тренду (частка потужності вкладу 90-96 %) та кусково-однорідного періодичного випадкового процесу (4-10 %), яка дала можливість врахувати, циклічний характер газоспоживання та режими роботи газотранспортної системи міста упродовж року.

Запропоновано метод аналізу газоспоживання з використанням методу «Гусениця-SSA», що дало можливість виділити компоненти запропонованої математичної моделі та провести їх оцінку.

Застосувавши модифіковану статистику Колмогорова-Смірнова адаптовано метод розбиття річного інтервалу газоспоживання на підінтервали, що уможливило виділення п'яти основних ділянок, які відображають особливості роботи газотранспортної системи міста в опалювальний, неопалювальний та перехідні періоди.

1. *Марченко Б.Г.* Обґрунтування математичної моделі газонавантажень / Б.Г. Марченко, Н.В. Мулик, М.С. Фриз // Вісник Тернопільського державного технічного університету ім. І.Пулкя. – Тернопіль, 2005. – №2. – С. 138 - 143
2. *Мулик Н.В.* Математична модель та метод прогнозу газоспоживання з урахуванням циклічності : автореф. дис. на здобуття наук. ступеня канд. техн. наук: 01.05.02 "Математичне моделювання" / Н.В. Мулик. — Тернопіль, 2006. — 19 с.
3. *Назаревич О.Б.* Інформаційна технологія моніторингу газоспоживання міста: автореф. дис. на здобуття наук. ступеня канд. техн. наук: 05.13.06 "Інформаційні технології" / О.Б. Назаревич. — Тернопіль, 2015. — 20 с.
4. *Щербак Т.Л.* Інформацій технологіагностики динаміки процесів електроспоживання організацій у штатному і нештатному режимах: автореф. дис. на здобуття наук. ступеня канд. техн. наук: 05.13.06 "Інформаційні технології" / Т.Л. Щербак. — Київ, 2010. — 20 с.
5. *Golyandina N.E.* Analysis of Time Series Structure: SSA and Related Techniques / Golyandina N.E., Nekrutkin V.V., and Zhigljavsky A.A. – Boca Raton: Chapman&Hall/CRC, 2000. – 305 p.
6. *Brodsky B.E.* Non-Parametric Statistical Diagnosis: Problems and Methods (Mathematics and Its Applications) / B.E. Brodsky, B.S. Darkhovsky. – Kluwer Academic Publishers, Dordrecht (the Netherlands), 2000. – 515p.
7. *Бродский Б.Е., Дарховский Б.С.* Асимптотический анализ некоторых оценок в апостериорной задаче о разладке // Теория вероятностей и ее применения. – 1990. – т.35, №3. – с. 551-557.
8. *Бродский Б.Е., Дарховский Б.С.* Алгоритм апостериорного обнаружения многократных разладок случайной последовательности // Автоматика и телемеханика. – 1993. – №1. – с. 62-67.
9. *Бродский Б.Е., Дарховский Б.С.* Проблемы и методы вероятностной диагностики // Автоматика и телемеханика. – 1999. – №8. – с. 3-50.
10. *Бродский Б.Е., Дарховский Б.С.* Непараметрический метод обнаружения моментов переключения двух случайных последовательностей // Автоматика и телемеханика. – 1989. – №10. – с. 66-74.
11. *Дарховский Б.С.* Непараметрический метод для апостериорного обнаружения момента “разладки” последовательности независимых случайных величин // Теория вероятностей и ее применения. – 1976. – т.21, №1. – с. 180-184.
12. *Дарховский Б.С.* Общий метод оценивания момента изменения вероятностных характеристик случайной последовательности // Статистические проблемы управления. – 1984. – вып.65. – с. 76-82.

Поступила 24.10.2016 р.