УДК 519:24
Н. Мулик
Тернопільський державний технічний університет імені Івана Пулюя

ГІСТОГРАМНИЙ АНАЛІЗ ГАЗОНАВАНТАЖЕНЬ ТА ВИБІР ІНФОРМАТИВНИХ ПАРАМЕТРІВ

Проведено гістограмний аналіз газонавантажень з використанням відомих методів статистичної обробки випадкових періодичних процесів, які полягають у побудові гістограм для кожної окремо виділеної φ *-серії стохастично періодичного процесу. В якості функції щільності розподілу вибрана крива Пірсона І типу. Для характеристики режиму газоспоживання запропоновано нові інформативні параметри.*

Вступ. Відомо, що газова промисловість є провідною галуззю паливно-енергетичного комплексу України, від роботи якої значною мірою залежать стабільність та розвиток національної економіки. Ефективність функціонування газових систем, крім технічного оснащення і програмного забезпечення автоматизованих систем управління, значною мірою залежить від оперативності та вірності прийняття рішень диспетчерами цих систем. Існуючі показники (середньодобовий рівень витрат газу, максимальні та мінімальні витрати газу за годину) та критерії оцінки газоспоживання не дають змоги розв'язати багато задач практики (наприклад, задачі прогнозу та виявлення несанкціонованого відбору газу тощо). Однією з причин є те, що в методах обробки навантажень не враховується їх важлива властивість – ритмічність.

Статистичний аналіз відносять до основних задач при вивченні будь-якого сигналу, зокрема і газонавантаження. Широке застосування серед статистичних методів дослідження отримав гістограмний аналіз, суть якого полягає в оцінюванні щільності розподілу.

Постановка задачі. Математична модель газонавантажень була обґрунтована в [1] у вигляді лінійного випадкового процесу:

$$\xi(t) = \int_{-\infty}^{\infty} \varphi(t-\tau)\, d\pi(\tau), \quad t \in (-\infty, \infty), \tag{1}$$

де $\varphi(s)$ – ядро лінійного випадкового процесу, невипадкова функція, для якої справджуються умови: $\int_{-\infty}^{\infty} \varphi^2(s)ds < \infty,\ \ \forall t,\ \varphi(s) = 0,\ \ s < 0;$

$\pi(\tau)$– узагальнений неоднорідний пуассонівський процес з незалежними приростами, причому $\mathbf{P}\{\pi(0)=0\}=1$, $\mathbf{D}[d\pi(\tau)] < \infty$.

У роботах [2] та [3] постулюється стохастична періодичність газонавантажень (період $T = 24$ год.), яка, очевидно, пояснюється фізіологічними особливостями функціонування людського організму, а також режимом роботи більшості людей. Виходячи з цього за математичну модель газоспоживання приймаємо лінійний випадковий періодичний процес.

Метою даної роботи є обґрунтувати методи статистичної обробки газонавантажень на основі побудованої математичної моделі, висунути статистичну гіпотезу відносно розподілу та запропонувати нові інформативні параметри газоспоживання.

Розв'язання задачі. Однією з важливих задач математичної статистики є оцінка щільності розподілу. Існують різні методи її знаходження, але історично першим і найбільш ґрунтовно розробленим є метод гістограм.

Оскільки газонавантаження в силу своєї стохастичної періодичності належить до класу нестаціонарних процесів, то для його обробки не можна використовувати методи статистичного аналізу, розроблені для стаціонарних процесів. Автори джерела [4] для гістограмного аналізу випадкових періодичних процесів розробили і запропонували використовувати метод φ -серій.

Зупинимося коротко на означенні стохастично періодичного процесу, на понятті φ -серій та їх основних властивостях.

Сепарабельний випадковий процес $\{\xi(t),\ t \in \mathbf{R}\}$ називається T *-періодичним за Слуцьким*, якщо існує таке число $T > 0$, що скінченновимірні вектори ($\xi(t_1)$, $\xi(t_2)$, ... , $\xi(t_k)$) і ($\xi(t_1+T)$, $\xi(t_2+T)$, ... , $\xi(t_k+T)$), де $\{t_1,\ t_2,\ ..., t_k\}$ - множина сепарабельності процесу $\xi(t)$, при всіх цілих $k > 1$ є стохастично еквівалентними у широкому розумінні. Для стохастично T -періодичного випадкового процесу його функція розподілу є періодичною з періодом T за сукупністю часових аргументів:

$$F(x_1,..., x_n;t_1...,t_n) = P\{\xi(t_1) < x_1,...,\xi(t_n) < x_n\} = F(x_1,..., x_n;t_1+T,...,t_n+T)\,. \tag{2}$$

Означимо тепер поняття φ -серії. Як відомо, методи статистичного аналізу зазвичай орієнтовані на використання ЕОМ. Тому, коли вихідною є реалізація неперервного аргументу (що відповідає розробленій математичній моделі газонавантажень), її необхідно попередньо дискретизувати з деяким кроком дискретизації h:

$$h = \frac{T}{N},\ N \geq 2, \tag{3}$$

де T - період процесу, N - число відліків на одному періоді.

Нехай φ_i - число, що належить проміжку $[0,T)$ і визначається наступним чином:

$$\varphi_i = i \cdot h,\ i = \overline{0, N-1}. \tag{4}$$

Для заданої фази φ_i на дійсній осі вводиться множина точок:

$$\{\varphi_i + lT,\ l \in \mathbf{Z}\}, \tag{5}$$

яку називають φ_i-сіткою.

Згідно з [5], φ_i-серією називається вкладена відносно стохастично періодичного процесу $\xi(t)$ послідовність

$$\xi(\varphi_i + lT), \quad l \in \mathbf{Z}, \tag{6}$$

областю визначення якого є φ_i-сітка.

Враховуючи (4), формулу (6) можна записати у вигляді:

$$\xi(\varphi_i + lT) = \xi(ih + lT) = \xi_i(l), i = \overline{0, N-1} \quad l \in \mathbf{Z}. \tag{7}$$

Для стохастично періодичних процесів характерна така властивість: відліки процесу, взяті через період T (φ_i-серії), утворюють стаціонарні та стаціонарно зв'язані у вузькому розумінні випадкові послідовності. Тобто задана на решітці з періодом T і будь-якою початковою фазою $\varphi_i \in [0,T]$ вкладена відносно процесу $\xi(t)$ послідовність (6) є стаціонарною, а будь-які дві такі послідовності з різними початковими фазами є стаціонарно зв'язаними.

Беручи до уваги стаціонарність φ_i-серій, а також те, що спостереження ведуться з кроком дискретизації $h = 1$год., на періоді Т=24год. можемо зафіксувати 24 початкових фази $\varphi_i = ih, \quad i = \overline{0,23}$ і відповідно виділити для цих початкових фаз 24 стаціонарні послідовності. Для кожної з цих послідовностей будуємо гістограму, приклади окремих гісторам наведені нижче.

Після побудови оцінки відповідної щільності розподілу її необхідно згладити. Задача згладжування не має однозначного розв'язання. Існують різноманітні методи згладжування, наприклад: метод найменших квадратів, метод згладжування з використанням системи перетворень Джексона, підхід Пірсона, який використовує метод моментів. К.Пірсон запропонував описувати одновимірні розподіли системою кривих, яка складається з тринадцяти типів. Даний метод зручний в тих випадках, коли необхідно просто підібрати криву щільності і не стоїть задача обґрунтування фізичної природи функції розподілу.

Слідуючи Пірсону [6], можна записати наступне рівняння:

$$\frac{dp(x)}{dx} = -p(x) \cdot \frac{a + x}{c_0 + c_1 x + c_2 x^2} \tag{8}$$

де $p(x)$- щільність розподілу, a, c_0, c_1, c_2 - дійсні параметри.

В залежності від значень параметрів a, c_0, c_1, c_2, які виражаються через перші чотири моменти, Пірсон отримував всі тринадцять типів кривих як розв'язок рівняння (8). З них I, IV, VI типи – основні, а решта –допоміжні (їх можна розглядати як підтипи основних типів).

Розглянемо основні математичні співвідношення, які використовуються для підбору кривої Пірсона.

Коефіцієнт асиметрії (характеризує асиметричність графіка функції щільності розподілу) задається виразом:

$$\beta_1 = \frac{\mu_3}{\mu_2^3}, \tag{9}$$

де μ_k - центральний момент k-того порядку.

Наступне співвідношення дозволяє знайти коефіцієнт ексцесу за відомими значеннями другого та четвертого центральних моментів:

$$\beta_2 = \frac{\mu_4}{\mu_2^2}. \tag{10}$$

Якщо початок координат вибраний в точці, що відповідає математичному сподіванню m_1 (перший початковий момент) і якщо величина d визначається формулою

$$d = 2(5\beta_2 - 6\beta_1 - 9), \tag{11}$$

то значення констант диференціального рівняння (8) задаються виразами:

$$a = \left(\sqrt{\mu_2}\,(\beta_2 + 3)\sqrt{\beta_1}\right)/d\,; \tag{12}$$

$$c_0 = \mu_2(4\beta_2 - 3\beta_1)/d\,; \tag{13}$$

$$c_1 = a\,; \tag{14}$$

$$c_2 = (2\beta_2 - 3\beta_1 - 6)/d\,. \tag{15}$$

Тип отриманої кривої залежить від значень коренів квадратного рівняння, що відповідає квадратному тричлену в знаменнику правої частини (8). Варто звернути увагу на критерій Пірсона

$$\kappa = \frac{c_1}{4c_0c_2}, \tag{16}$$

який виділяє різні типи кривих і базується на розрізненні типів коренів вище згаданого квадратного рівняння. Якщо $\kappa < 0$ (корені квадратного рівняння дійсні та мають різний знак), то отримуємо криву I типу, якщо $0 < \kappa < 1$ (комплексні корені рівняння) – криву IV типу, якщо $\kappa > 1$ (дійсні корені одного знаку) – криву VI типу. Для підбору неосновних типів II, III, V, VII-XIII необхідно застосовувати додаткові критерії.

На практиці для підбору кривої та при обчисленні параметрів рівняння (8) у формулах (9)-(15) замість теоретичних моментів m_1 та μ_k використовують їх оцінки $\overline{m}_1$ та $\overline{\mu}_k$ (метод моментів). Оцінки параметрів, отримані методом моментів, є функціями від емпіричних моментів.

Оцінку математичного сподівання для фіксованої φ_i-серії знаходимо із співвідношення:

$$\overline{m}_1 = \frac{1}{L}\sum_{l=1}^{L}\xi_i(l), \tag{17}$$

де L – обсяг вибірки φ_i-серії (кількість добових циклів).

Оцінки центральних моментів знаходимо за формулами:

$$\overline{\mu}_k = \frac{1}{L}\sum_{l=1}^{L}\left(\xi_i(l) - \overline{m}_1\right)^k. \tag{18}$$

При проведенні аналізу реалізацій газоспоживання в якості функції щільності розподілу була вибрана крива Пірсона I типу в літній період та крива Пірсона II типу в опалювальний сезон.

Для того, щоб більш достовірно судити про розподіл досліджуваної вибірки, необхідно побудувати метод перевірки статистичних гіпотез відносно обраної функції щільності розподілу $p(x)$. Для цього застосуємо критерій згоди Пірсона (χ^2-тест) [7]. Для отримання вірогідного результату такої перевірки було опрацьовано реалізацію газоспоживання, яка містила 92 добові цикли.

При проведенні перевірки вибрані наступні параметри тесту. Число ступенів свободи було вибрано рівним $r=4$, рівень значимості $\alpha=0.05$ і, відповідно, квантиль χ^2-розподілу Пірсона з r ступенями свободи $\chi^2_{0.95,4}=9.48$. На рисунку 1 наведено графік газоспоживання с.Підгайці на 12 годину дня.

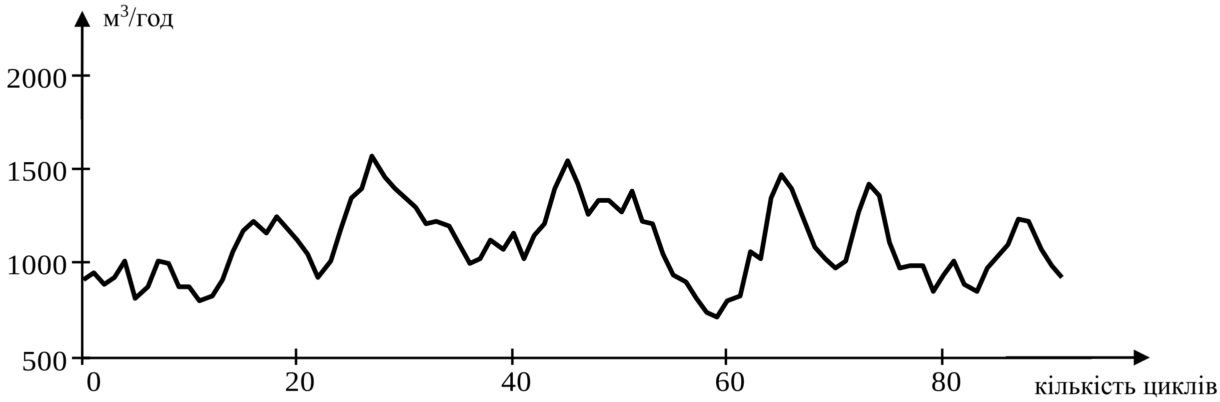

Рисунок 1 - Графік газоспоживання на 12-ту годину дня, зареєстрованого в 92 добових циклах в зимовий період

Гістограму та графік згладжуючої кривої Пірсона II типу, побудованих за результатами даних про газонавантаження на 12-ту годину дня в 92 добових циклах с.Підгайці , подано на рисунку 2.

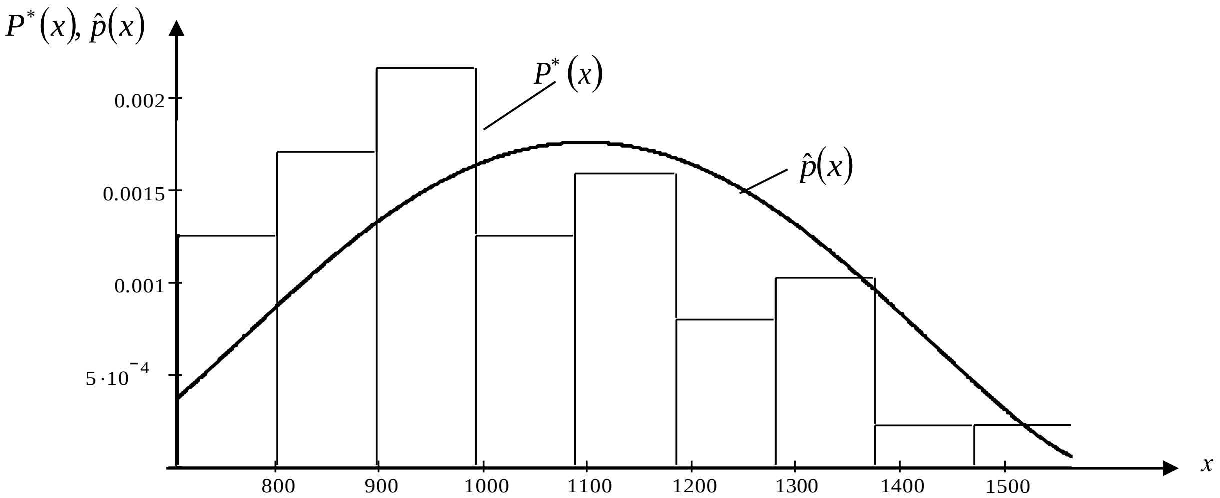

Рисунок 2 - Гістограма $P^*(x)$ та згладжуюча її крива Пірсона II типу $\hat{p}(x)$для вибірки з 92 значень газоспоживання на 12-ту годину дня в зимовий період

На рисунку 3 наведено графік газоспоживання на 2 годину ночі зареєстрованої реалізації газонавантажень у 92-х добових циклах в зимовий період, а на рисунку 4 подано відповідну гістограму.

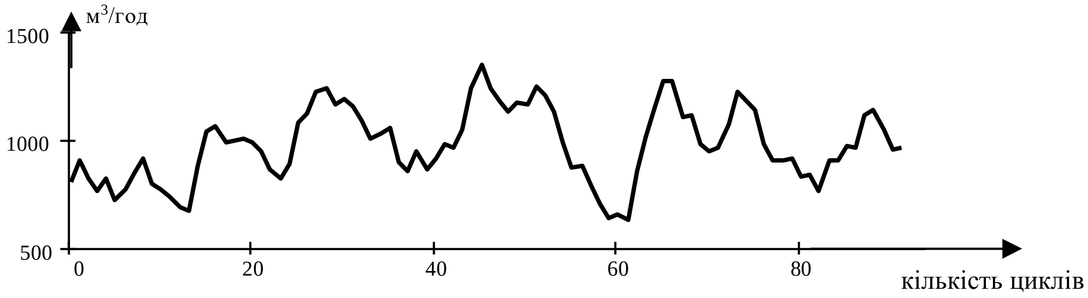

Рисунок 3 - Графік газоспоживання на 2-гу годину ночі, зареєстрованого в 92 добових циклах в зимовий період

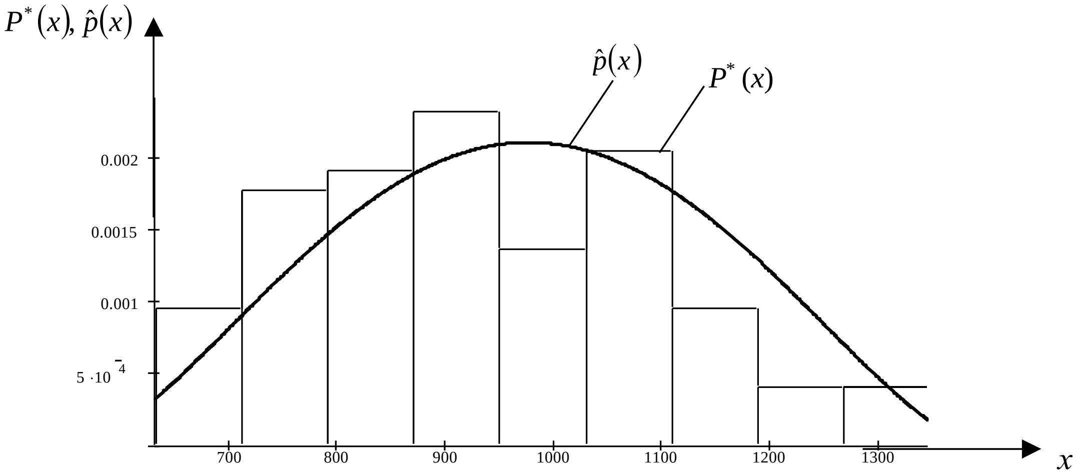

Рисунок 4 - Гістограма $P^*(x)$ та згладжуюча її крива Пірсона ІІ типу $\hat{p}(x)$ для вибірки з 92 значень газоспоживання на 2-гу годину ночі в зимовий період

Графік газоспоживання на 2 годину ночі зареєстрованої реалізації газонавантажень у 68-х добових циклах в літній період побудований на риунку 5, а на рисунку 6 подано відповідну гістограму.

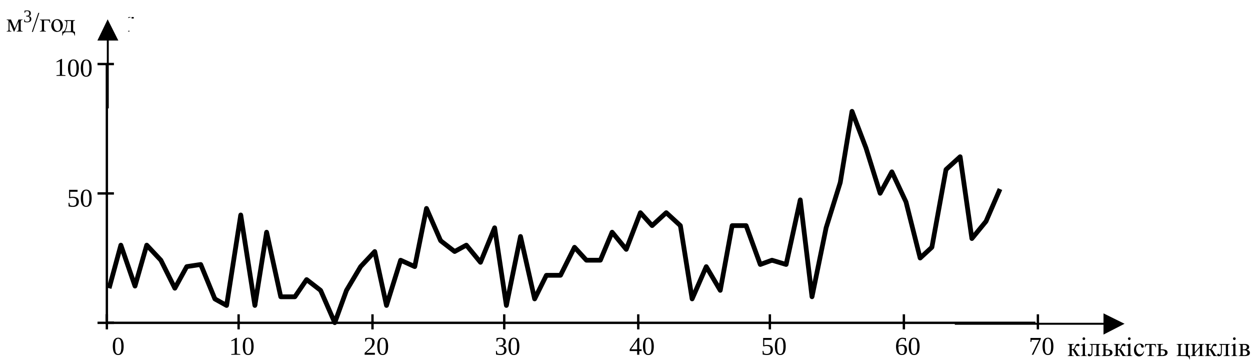

Рисунок 5 - Графік газоспоживання на 2-гу годину ночі, зареєстрованого в 68 добових циклах в літній період

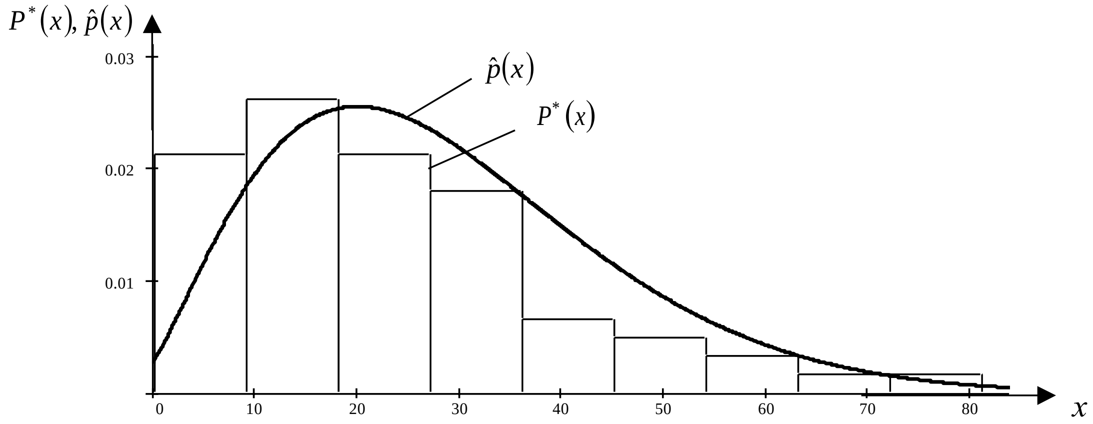

Рисунок 6 - Гістограма $P^*(x)$ та згладжуюча її крива Пірсона I типу $\hat{p}(x)$для вибірки з 68 значень газоспоживання на 2-гу годину ночі в літній період

На основі результатів перевірки гіпотези про розподіл φ_i -серій за критерієм згоди Пірсона встановлено, що ці результати не суперечать гіпотезі про те, що функція щільності розподілу описується кривої Пірсона I чи II типів. Результати перевірки наведені в таблиці 1.

Таблиця 1 - Результати перевірки гіпотези про розподіл газоспоживання в зимовий період за критерієм згоди χ^2-Пірсона

Година (номер φ_i -серії)	0	1	2	3	4	5	6	7	8	9	10	11
Тип кривої Пірсона	1	2	2	2	2	2	2	1	2	2	1	1
Критерій χ^2	3,92	3,90	3,63	5,19	6,69	6,42	3,77	5,08	9,01	5,28	4,08	8,86
Година(номер φ_i -серії)	12	13	14	15	16	17	18	19	20	21	22	23
Тип кривої Пірсона	2	2	1	1	1	2	4	2	2	2	2	2
Критерій χ^2	4,61	9,2	6,08	9,21	6,28	13,9	–	32	7,21	5,22	3,57	4,57

Вибір розподілу φ_i-серій є підставою для вибору параметрів для оцінювання газонавантажень. Використовувати функцію щільності розподілу як критерій оцінки при проведенні статистичного аналізу є незручно, тому необхідно вибрати нові інформативні параметри газоспоживання, які однозначно задають його розподіл та відповідають за різні стани газонавантажень. Такими параметрами у нашому випадку будуть коефіцієнти рівняння (8), які обчислюються згідно з формулами (12)-(15) за відомими оцінками перших чотирьох центральних моментів.

Відзначимо, що більшість систем для аналізу газонавантажень використовують моментні функції лише першого (математичне сподівання) та другого порядків (дисперсія та кореляційна функція). Вищенаведені статистичні оцінки дозволяють досліджувати газоспоживання в рамках вищих моментних функцій, що дає можливість підвищити інформативність комп’ютерної діагностики функціонального стану газової системи.

Висновки. В роботі проведено гістограмний аналіз газонавантажень з використанням відомих методів статистичної обробки випадкових періодичних процесів, що ґрунтуються на властивості стаціонарності φ -серій. В якості функції щільності розподілу вибрана крива Пірсона I(II) типу, для характеристики стану газоспоживання запропоновано використовувати нові інформативні параметри – коефіцієнти диференціального рівняння, що визначає різні типи кривих Пірсона.

Histogrammic gasloading analysis based on the known methods of statistical treatment of stochastic periodic processes is carried out in the article. These methods deal with thebuilding of the bar chart for every φ *-series of stochastic periodic process. The Pirson's curve of the first type is chosen as probability density function. New informative parameters for characteristic of gasusing conditions are proposed.*

Література

1. Б.Г. Марченко, Н.В. Мулик, М.Є. Фриз. Обґрунтування математичної моделі газонавантаження // Вісник Тернопільського державного технічного університету. – 2005. – №2. – С.138-143.
2. О.В. Мацюк, М.В.Приймак. Моделі газонавантажень з врахуванням стохастичної періодичності та можливості їх статистичного аналізу // Розвідка та розробка нафтових і газових родовищ. – 2003. – №2(7).– С.64-68.
3. Б.Г.Марченко, Н.В.Мулик. Математичні моделі газонавантажень // Збірник тез 8-ої науково-технічної конференції ТДТУ ім.І.Пулюя. -2004.- С.58.
4. О.В.Маєвський, М.В. Приймак, Л.М.Щербак. Гісторамний аналіз випадкових процесів та його використання в прикладних дослідженнях// Вимірювальна та обчислювальна техніка в технологічних процесах.–2003.–№2.–С.26-31.
5. Б.Г. Марченко, М.В.Приймак. Побудова моделі та аналіз стохастично періодичних навантажень енергосистем // Праці інс-ту електродинаміки.- Київ: ІЕД НАН України, 1999. – Вип.1. – С.129-153.
6. Дж. Боллард. Справочник по вычислительным методам статистики. – М: Финансы и статистика, 1982. – 369 с.
7. Р.А.Фишер. Статистические методы для исследователей. – М: Госстатиздат, 1958. – 267 с.

Одержано 13.12.2005 р.