

УДК 620.311

ЗАСТОСУВАННЯ МЕТОДУ ГОЛОВНИХ КОМПОНЕНТ ДЛЯ ІДЕНТИФІКАЦІЇ ВПЛИВУ ПОКАЗНИКІВ НА РІВЕНЬ ЕЛЕКТРОСПОЖИВАННЯ

В. П. Розен, Я. М. Демчик

Національний технічний університет України «Київський політехнічний інститут ім. І.Сікорського»

Анотація. В статті описано застосування методу головних компонент для оцінювання вкладу годинного електричного навантаження в формування конфігурації добового графіка електричних навантажень об'єктів енергетичного ринку. Результати статті можна використати для визначення тарифних зон графіків електричних навантажень, які впливають на загальний результат аналізування об'єктів енергоринку.

Ключові слова: факторний аналіз, метод головних компонент, система електроспоживання, графік електричних навантажень.

Вступ.

Аналіз динаміки електроспоживання енергооб'єднання є складною задачею та є необхідним компонентом для ефективної роботи енергоринку. Суттєвий вплив на режим електроспоживання має велика кількість факторів, що формують конфігурацію графіка електричних навантажень (ГЕН). При цьому виникає необхідність використання сучасного математичного апарату, який в свою чергу забезпечить ефективний аналіз ГЕН. На теперішній час існує значна кількість методів аналізу даних: аналітичний, деталізація, порівняння та моделювання, балансовий, елімінування, статистичний, економіко-математичний, інтегральний та інші [1]. Але тільки інтегральний метод дозволяє оперувати та отримувати результат в чисельному вигляді. Необхідно враховувати також недоліки існуючих інтегральних методів оцінки: використання методу експертних оцінок, розрахунків на підставі малої кількості факторів, неможливість використання важкодоступних даних. Для цього запропоновано уникати оцінки експертів, враховувати як можна більшу кількість показників, що впливають на прийняття рішень та застосовувати дані отримані з систем автоматизованого обліку електроспоживання. В якості математичного апарату для аналізування ГЕН може виступити метод головних компонент як складовий факторного аналізу.

Метод головних компонент має деякі переваги перед простими методами факторного аналізу. На відміну від інших методів, він здатний виявити достатнє число характерних факторів при факторному аналізі даних [7].

Метод головних компонент виявляє k компонент — факторів, що пояснюють усю дисперсію й кореляції вихідних k випадкових величин; при цьому компоненти будуються в порядку спадання частки, що пояснюється ними, сумарної дисперсії вихідних величин, що дозволяє найчастіше обмежитися декількома першими компонентами.

1. Аналіз останніх досліджень і публікацій.

Основні ідеї методу головних компонентів знайшли своє відображення у працях С. Барта, К. Пірсона, Г. Томсона, Д. Гарнета, К. Хользінгера, Г. Голелінга. Метод головних компонент (МГК) в більшості застосовувався під час вирішення економічних задач підприємств. В статті [2] розглядалося питання застосування МГК для оцінки конкурентоспроможності підприємства. Доведено, що при застосуванні МГК можна побудувати новий інтегральний метод оцінки, який буде звільнений від суб'єктивної оцінки, спиратися на велику кількість факторів, а також буде математично обґрунтований [2]. В статті [3] викладено теоретичні аспекти оцінки фінансового стану підприємства та його динаміки з використання факторного аналізу та зокрема МГК. Використання методу головних компонент дозволяє пояснити велику кількість ознак невеликою кількістю ознак невеликою кількістю компонент.

В області енергетики метод застосовувався для вирішення задач енергетичної безпеки та оцінювання режимів електроспоживання [13]. Залишається відкритим питання виділення тарифних годин з ГЕН за допомогою МГК. Використання методу дозволило б ефективно аналізувати дані електроспоживання об'єктів.

© Розен В. П., Демчик Я. М., 2017

2. Мета та завдання.

Дослідження є підвищення адекватності прийняття рішення визначення головних компонент, що впливають на формування режиму електроспоживання об'єктів енергоринку. Використання МГК дозволить зменшити об'єм даних та виявити спільні узагальнюючі ознаки.

3. Виклад основного матеріалу дослідження.

Режим електроспоживання є відкритою системою на яку здійснює позитивний чи негативний вплив низки факторів. Для дослідження тенденцій режиму електроспоживання об'єктів енергоринку необхідно виділити домінуючі електричні навантаження з ГЕН і також визначити причинно-наслідкові зв'язки між ними. Деякі із цих навантажень можуть бути спільними і їх доцільно об'єднувати в одну головну компоненту. Це дозволить зменшити об'єм статистичних даних, які будуть включати в себе основну частину електричних навантажень та визначити структуру ГЕН – часів максимума, напівпіку та провалу електричного навантаження. Найменш впливові складові ГЕН які не ввійдуть до основної частини не роблять внеску в кореляційний зв'язок між ознаками. Спільні електричні навантаження, кількість яких є меншою за число досліджуваних електричних навантажень, роблять внесок у матрицю парних кореляцій.

Під час аналізу використовувався МГК. Алгоритм якого має вид:

Етап 1. Формування матриці вихідних даних, яка інтерпретується як реалізація n -мірного випадкового вектору (кількість експертів) при N спостережень. Рядок це добові електричні навантаження за 24 години. Стобець є погодинні електричні навантаження за період одного місяця

$$Y_n = \begin{bmatrix} y_{11} & \dots & y_{1n} \\ \dots & \dots & \dots \\ y_{m1} & \dots & y_{mn} \end{bmatrix} \quad (1)$$

Етап 2. На цьому етапі властивості даного випадкового вектору з достатньою точністю описуються вектором математичних сподівань

$$Y = (y_1 \dots y_n) \quad (2)$$

та кореляційною матрицею електричних навантажень

$$K_{[n]} = \|K_{ij}\|_n^n \quad (3)$$

або матрицею коефіцієнтів кореляцій

$$R_{[n]} = \|r_{ij}\|_n^n \quad (4)$$

де $r_{ij} = \frac{k_{ij}}{\sigma_i \sigma_j}$ – коефіцієнт кореляції між i -м та j -м фактором.

Етап 3. Параметри кореляційної матриці перед обчисленням нормуються:

$$z_{ij} = \frac{y_{ij} - y_{cpi}}{\sigma_j} \quad i = 1, \dots, N, j = 1, \dots, n; \quad (5)$$

де y_{ij} – значення випадкової величини з матриці електричних навантажень;

y_{cpi} – середнє значення випадкової величини y_{ij} за результатами N спостережень;

σ_j – середнє квадратичне відхилення;

Середнє значення випадкової величини визначається за виразом:

$$y_{cpi} = \frac{1}{N} \sum_n y_{ij} \quad (6)$$

а середнє квадратичне відхилення

$$\sigma_i = \sqrt{\frac{\sum_{i=1}^N (y_{ij} - y_{cpi})^2}{N - 1}} \quad (7)$$

Етап 4. Представлення результатів спостережень у формі:

$$z_{[N,n]} = \|z_{ij}\|_N^n \quad (8)$$

Етап 5 розрахунку побудований на відомій процедурі обчислення головних компонент, яка в подальшому деяким чином модифікується для отримання інваріантного тривимірного зображення.

4. Результати дослідження визначення вкладу кожного з компонентів графіка електричних навантажень на загальний результат.

Аналіз базується на статистичних даних графіків електричних навантажень об'єкта енергетичного ринку в період з 01.01.2011р по 31.01.2011р.(в табл. 1).

Таблиця 1.

Статистичні дані графіків електричних навантажень за січень 2011 року об'єкта енергетичного ринку, кВт*год

Го- дина	00:00	01:00	02:00	...	21:00	22:00	23:00
Дата							
01.01 .2011	3223 32,00	2910 88,00	2745 60,00	...	299746 ,00	286124 ,00	2667 03,00
02	2582 04,00	2470 91,00	2407 44,00	...	300571 ,00	296212 ,00	2700 29,00
...
29	2639 46,00	2539 81,00	2494 78,00	...	320987 ,00	305956 ,00	2826 89,00
30	2668 47,00	2543 50,00	2510 32,00	...	306038 ,00	299979 ,00	2714 37,00
31.01 .2011	2548 83,00	2490 54,00	2406 80,00	...	316565 ,00	298830 ,00	2618 01,00

Розглянемо перший крок алгоритму розрахунку, який здійснюється на основі вихідних даних кореляційної матриці, і відображає тісноту зв'язку між відібраними електричними навантаженнями (табл. 2). Обчислення проводилися за допомогою стандартного модуля *Factor Analysis*, запропонованого пакетом *STATISTICA*[4].

Таблиця 2.

Кореляційна матриця

Показники	Коефіцієнти кореляції						
	X1	X2	X3	...	X22	X23	X24
X1	1,00	0,94	0,90	...	0,63	0,44	0,31
X2	0,94	1,00	0,98	...	0,67	0,48	0,34
X3	0,90	0,98	1,00	...	0,68	0,50	0,34
X4	0,81	0,92	0,94	...	0,65	0,47	0,33
...
X28	0,05	0,13	0,18	...	0,95	0,83	0,59
X29	0,24	0,30	0,36	...	1,00	0,92	0,71
X30	0,39	0,47	0,52	...	0,92	1,00	0,76
X31	0,34	0,41	0,48	...	0,71	0,76	1,00

Коректний розв'язок задач за допомогою методу головних компонент передбачає підтвердження значимості вихідної матриці парних кореляцій (коваріацій) та достатньої кількості узагальнених факторних ознак під час аналізуванні. Подальший розрахунок спирається на відому процедуру обчислення головних компонент, яка в подальшому деяким чином модифікується для отримання інваріантного тривимірного зображення.

Раціональну кількість головних компонент дозволяє вибрати критерій відсіювання (*scree-test*), який був розроблений Каттеллом (*Cattell*) [12]. Послідовність виділення головних компонент доцільно представити у вигляді графіка «кам'янистого осипу» (рис. 3), який вказує на

доцільність формування трьох головних факторів.

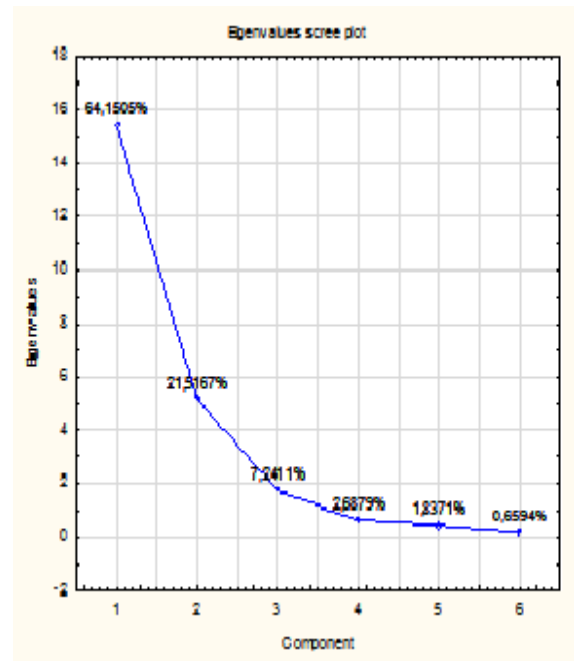


Рис. 1. Графік «кам'янистого осипу»

Формування головних компонент із сукупності результатів спостережень зводиться до аналізування матриці кореляційних моментів та обчислення власних чисел та власних векторів цієї матриці.

Результати застосування методу головних компонент представляються даними матриці факторних навантажень (табл. 3). Коефіцієнти цієї матриці a_{ij} дозволяють провести чисельно-формальне пояснення коефіцієнтів кореляції, а це дає нам можливість припускати, що за кореляціями стоїть фактор, який міг би причинно обумовлювати ці кореляції.

Як видно з матриці факторних навантажень (табл. 3), сформовані три головні компоненти електричних навантажень які описують 92,9% загальної дисперсії.

Таблиця 3.

Матриця факторних навантажень

Показники	F1	F2	F3
1	-0,00639	-0,908566	-0,181002
2	-0,11564	-0,947181	-0,235642
3	-0,15099	-0,958006	-0,177748
4	-0,34344	-0,902493	-0,158560
5	-0,59438	-0,706934	-0,262129
6	-0,87465	-0,224996	-0,341872
7	-0,90316	0,148638	-0,260846
8	-0,92247	0,217665	-0,233321
9	-0,94200	0,240377	-0,178771
10	-0,95456	0,221947	-0,139371

11	-0,96304	0,173390	-0,117272
12	-0,96058	0,175114	-0,090245
13	-0,96081	0,201513	-0,102544
14	-0,92578	0,271847	-0,199226
15	-0,90810	0,255872	-0,164235
16	-0,90072	0,150367	-0,137461
17	-0,91353	0,091272	-0,057443
18	-0,94940	0,126267	0,220166
19	-0,90687	0,036801	0,334697
20	-0,87192	0,014420	0,429990
21	-0,87920	-0,124071	0,423645
22	-0,80114	-0,335554	0,457538
23	-0,64921	-0,544818	0,472277
24	-0,48713	-0,517904	0,393548
Власні числа, λ_j	15,39614	5,164002	1,737876
Вага факторів, %	0,64151	0,215167	0,072411

Перша головна компонента F_1 (64,15% від загальної дисперсії) включає такі електричні навантаження пікових та напівпікових годин споживання електроенергії з 06:00 до 22:00 год. Усі ці електричні навантаження відображають головне споживання електроенергії.

Друга головна компонента F_2 (21,51% від загальної дисперсії) включає електричні навантаження нічних годин споживання електроенергії з 01:00 до 05:00 год. Усі ці електричні навантаження відображають низьке споживання електроенергії.

Третя головна компонента F_3 (7,24% від загальної дисперсії) включає електричні навантаження яких не ввійшли до жодного з перших двох головних компонент.

Графічну візуалізацію приналежності вихідних факторів до певної головної компоненти забезпечує побудова тримірної діаграми факторних навантажень головних компонент (рис. 4).

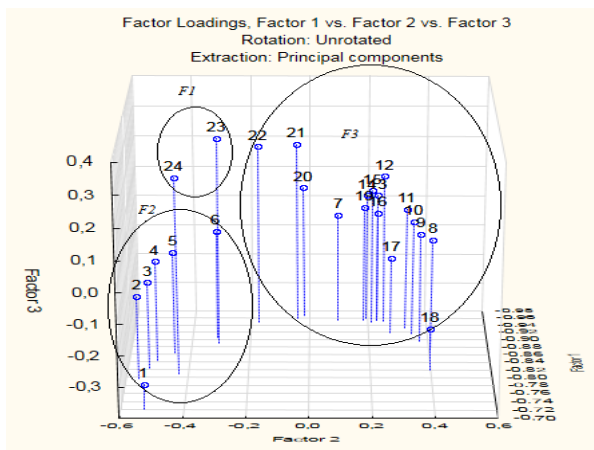


Рис. 2. Діаграма факторних навантажень головних компонент

Дослідження залишків.

Питання дослідження, як змінилася якість даних після перетворення. Для розрахунку TRV і

ERV можна отримати матрицю остатків E для кожного числа головних компонентів і вирахувати потрібні показники.

Таблиця 4

Аналіз залишків

Годи-на	01:00	02:00	03:00	...	22:00	23:00	
Дата							
01.01.2011	0,145524	-0,066999	-0,081695	...	0,097943	0,052710	-0,007031
02	0,087487	0,048546	-0,077573	...	-0,072428	0,481550	-0,056387
...
30	0,069849	-0,013904	-0,076076	...	-0,020512	0,187145	-0,046065
31	0,074595	-0,156343	-0,075968		-0,357125	0,122548	0,047137

Після отримання матриці залишків розраховуємо по кожній складовій середнє квадратичне відхилення

$$v_0 = \frac{1}{I} \sum_{i=1}^I v_i \tag{9}$$

$$TVR = \frac{v_0}{J} \tag{10}$$

$$TRV = 0,081$$

$$ERV = 1 - \sum_{ij}^{auto} \frac{Iv_0}{2} \tag{11}$$

$$ERV = 0,929$$

З результатів дослідження (табл.5), робимо висновок найбільший вклад в загальний результат графіків електричних навантажень вносить перший головний компонент.

Таблиця 5

Результати дослідження

№	Особ. значення	% Зальної дисперсії	Кумулятивн. загал. знач.	Кумулят %
1	15,39612	64,15048	15,39612	64,15048
2	5,16402	21,51674	20,56013	85,66722
3	1,73787	7,24114	22,29801	92,90836

Інші компоненти не такі важливі, але для більшої загальної дисперсії беремо 3 компоненти в які не входить жоден із складових показників.

Висновки

1. Метод МГК дозволив отримати об'єднаний та скорочений об'єм вихідної інформації яка включає в себе головну складову пока-

зників які моделюють 92,9% даних, або іншими словами , шум який залишився після проєкції залишає лише 8,1% вихідних даних.

2. За рахунок методу головних компонент ГЕН розділено на три головні групи пікові, напівпікові та години провалу електричного навантаження. Перший головний компонент включив в себе пікові та напівпікові години споживання електричної енергії що складає 64.15% від загальної дисперсії. Другий головний компонент включив в себе години провалу електричного навантаження (нічні години) споживання електричної енергії що складає 21.51% від загальної дисперсії. Та третя головна компонента включаєв себе лише 7.24% від загальної дисперсії.

3. Виходячи із інтерпретації головних чинників, отримано основні компоненти які впливають на загальний результат електроспоживання. Зокрема, це пікові та напівпікові години електроспоживання. Нічні години показують, що споживання в ці часи є мінімальним і тому дозволять за рахунок них вирівнювати електроспоживання об'єктів.

Список використаної літератури

- Івахненко, В. М. Курс економічного аналізу. Навчальний посібник [Текст] / В. М. Івахненко. – 3-тє видання перероблене та доповнене. – Київ: Знання-Прес, 2002. – 190с.
- Толкачев, Д. О. Застосування методу головних компонент для оцінки конкурентоспроможності підприємства [Текст] / Д. О. Толкачев. – Донецький національний університет. – Донецьк.
- Загреба, М. М. Теоретичні аспекти використання факторного аналізу при аналізі динаміки фінансового стану підприємства [Текст] / М. М. Загреба. – Наукові праці Кіровоградського національного технічного університету. Економічні науки. – Кіровоград, 2014. – 455–461с.
- Халафян, А. А. STATISTICA 6. Статистический анализ данных: Учебник [Текст] / А. А. Халафян [3-е изд.]. – М.: ООО "Бином-Пресс", 2007. – 512 с.
- Андриенко, В. М. Статистические методы анализа экономических показателей [Электронный ресурс] / В. М. Андриенко, М. Г. Самисько. – Режим доступа: http://www.rusnauka.com/11_EISN_2008/Economics/30352.doc.htm.
- Ким, Дж.-О. Факторный, дискриминантный и кластерный анализ [Текст] / Дж.-О. Ким, Ч. У. Мьюллер, У. Р. Клекка; пер с англ. – М.: Финансы и статистика, 1989. – 608с
- Янковой, А. Г. Многомерный анализ в системе STATISTICA [Текст] / А. Г. Янковой. – Вып. 2. – Одесса: Оптимум, 2001. – 216 с.
- Сошникова, Л. А. Многомерный статистический анализ в экономике [Текст] / Л. А. Сошникова, В. Н. Томашевич, Г. Уебе, М. Шефер; под ред. проф. В. Н. Томашевича. – М.: ЮНИТИ-ДАНА, 1999. – 598 с.
- Дубров, А. М. Многомерные статистические методы [Текст] / А. М. Дубров, В. С. Мхитарян, Л. И. Трошин – М.: Финансы и статистика, 1998
- Главные компоненты и факторный анализ [Электронный ресурс] // StatSoft, Inc.. – 2003. – Режим доступа до ресурсу: <http://statsoft.ru/home/textbook/default.htm>.
- Метод Главных Компонент (PCA) [Электронный ресурс] // chemometrics.ru. – 2012. – Режим доступа до ресурсу: <http://www.chemometrics.ru/materials/textbooks/pca.htm>.
- Иберла, К. Факторный анализ [Текст] / К. Иберла. – М.: Статистика, 1980. – 398 с.
- Добровольська, Л. Н. Моніторинг та планування енергоефективних режимів електроспоживання вугільних шахт: монографія [Текст] / Л. Н. Добровольська, В. П. Розен, В. І. Волинець. – Луцьк: луцький НТУ, 2014. – 152с

References

- Ivakhnenko, V. M., (2002), The course of economic analysis. [Kurs ekonomichnoho analizu. Navchal'nyy posibnyk] 3rd edition revised and enlarged. – Kyiv: Knowledge Press, –190S.
- Tolkachev, D. O., Application of principal component for assessing the competitiveness of enterprises [Zastosuvannya metodu holovnykh komponent dlya otsinky konkurentospromozhnosti pidpryyemstva] / Donetsk National University. – Donetsk.
- Zagreb, N. M. (2014), Theoretical aspects of factor analysis in analyzing the dynamics of the financial condition of the company [Teoretychni aspekty vykorystannya faktornoho analizu pry analizi dynamiky finansovoho stanu pidpryyemstva] / Proceedings Kirovograd National Technical University. Economic science. - Kirovograd :, – 455–461s.
- Halafyan, A. A. (2007) STATISTICA 6. Statystichesky data analysis: Textbook [Statystichesky analiz dannykh: Uchebnyk] / [3rd ed.]. – М.: signature "Binom-Press", – 512 p.

5. Andryenko, V. N. (2008), Statistical methods of analysis of economic indicators [Statysticheskye metody analiza ékonomycheskykh pokazateley] / – Access: http://www.rusnauka.com/11_EISN_2008/Economics/30352.doc.htm.
6. Who, Dzh.-O. (1989), Faktor, dyskrymynant klaster analysis and [Faktornyy, dyskrymynantnyy y klasternyy analiz] / Lane with English. – M.: Finance and Statistics, – 608s
7. Yankovoy, A. G. (2001), Mnohomer analysis in the system STATISTICA [Mnohomernyy analiz v systeme STATISTICA] / Odessa: Optimum, – 216 p.
8. Soshnikova, L. A. (1999), Mnohomer statystichesky analysis in economy [Mnohomernyy statystichesky analiz v ékonomyke] / ed. prof.. – Moscow: UNITY-DANA, – 598 p.
9. Dubrov, A. M. (1998) Multidimensional statistical methods [Mnohomernye statysticheskye metody] – M.: Finance and Statistics,
10. (2003) Key Komponenty faktor analysis // StatSoft, Inc [Hlavnye komponenty y faktornyy analiz] – Access to the property: <http://statsoft.ru/home/textbook/default.htm>.
11. (2012) The method Hlav component (PCA) [Metod Hlavnykh Komponent (PCA)] // chemometrics.ru. – Access to the property: <http://www.chemometrics.ru/materials/textbooks/pca.htm>.
12. K. Yberla (1980), Faktor analysis [factor analysis] / M. Statistics, – 398 p.
13. Dobrovolska, L. N. (2014), Monitoring and planning of energy efficient modes elektrospozhyvannyavuhilnyh mines: monograph [Monitoring and planning of energy efficient mode power consumption of coal mines: monograph] / Lutsk Lutsk NTU, 152pages

APPLICATION OF PRINCIPAL COMPONENT INDICES FOR IDENTIFICATION OF INFLUENCE ON LEVEL POWER CONSUMPTION

V. P. Rozen, J. M. Demchyk

National Technical University of Ukraine «Kyiv Polytechnic Institute. I.Sikorskoho»

Abstract. *The article describes the application of the main component method for estimating the contribution of the hourly electrical load to the configuration of the daily schedule of electrical loads of power market objects. The results of the article can be used to determine the tariff zones of electric load graphs that affect the overall result of the analysis of energy market objects. To do this, it is suggested to avoid evaluating experts, to take into account as much as possible the number of indicators influencing the decision-making and apply the data obtained from the systems of automated metering of power consumption. As a mathematical tool for the analysis of the GEN, the method of principal components can be used as a composite factor analysis. The main component method reveals k component factors that explain the entire variance and correlation of the output k random variables; While the components are constructed in decreasing order of the fraction, which explains them, the total variance of the output quantities, allows us to often confine ourselves to a few first components. To study trends in the energy consumption regime of energy market objects, it is necessary to identify the dominant electrical loads from the GEN and also to determine the cause-effect relationships between them. Some of these loads can be common and it is advisable to combine them into one main component. This will reduce the amount of statistical data that will include the bulk of electrical loads and determine the structure of the GEN - peak time, half-peak and electrical load failure. The least influential components of the GEN that do not enter the main part do not contribute to the correlation relationship between the symptoms.*

Key words: *factor analysis, principal component method, system power consumption, schedule power loads.*

ПРИМЕНЕНИЕ МЕТОДА ГЛАВНЫХ КОМПОНЕНТ ДЛЯ ИДЕНТИФИКАЦИИ ВЛИЯНИЯ ПОКАЗАТЕЛЕЙ НА УРОВЕНЬ ЭЛЕКТРОПОТРЕБЛЕНИЯ

В. П. Розен, Я. М. Демчик

Национальный технический университет Украины «Киевский политехнический институт им. И.Сикорского»

Аннотация. В статье описано применение метода главных компонент для оценки вклада часовой электрической нагрузки в формирование конфигурации суточного графика электрических нагрузок объектов энергетического рынка. Результаты статьи можно использовать для определения тарифных зон графиков электрических нагрузок, влияющих на общий результат анализа объектов энергорынка.

Ключевые слова: факторный анализ, метод главных компонент, система электроснабжения, график электрических нагрузок.

Отримано 26.04.2017



Розен Віктор Петрович, доктор технічних наук, професор, завідувачий кафедрою автоматизації електротехнічних комплексів Національного технічного університету України «Київський політехнічний інститут ім. Ігоря Сікорського». Просп. Перемоги, 37, Київ, Україна

Victor. Rosen, Dr. of Science, Professor, Head of the Department of electrical automation systems, National Technical University of Ukraine "Kyiv Polytechnic Institute. Igor Sikorsky. " Peremohy ave, 37, Kyiv, Ukraine

ORCID ID 0000-0002-0440-4251



Демчик Ярослав Михайлович, аспірант кафедри автоматизації електротехнічними комплексами Національного технічного університету України «Київський політехнічний інститут ім. Ігоря Сікорського». Просп. Перемоги, 37, Київ, Україна

Demchik Yaroslav Mikhailovich, graduate student the Department of electrical automation systems, National Technical University of Ukraine "Kyiv Polytechnic Institute. Igor Sikorsky. " Peremohy ave, 37, Kyiv, Ukraine

ORCID ID 0000-0001-6936-2366