

УДК 519.216.3:621.317.38

**Левченко Сергій Андрійович**, кандидат технічних наук, доцент

Запорізька державна інженерна академія. Кафедра електротехніки та енергетичного менеджменту, м. Запоріжжя, Україна. *Пр. Леніна 226, м. Запоріжжя, 69006. Тел.: +38 (063) 6045676.*

*E-mail: levchenko\_s@rambler.ru*

### ПРОГНОЗУВАННЯ ВИТРАТ ЕЛЕКТРИЧНОЇ ЕНЕРГІЇ В ЗАЛЕЖНОСТІ ВІД ОБ'ЄМІВ ВИГОТОВЛЕНОЇ ПРОДУКЦІЇ

*В статті розглянуто прогнозування електроспоживання за допомогою кореляційно-регресійного аналізу та теорії ймовірностей на перспективу на прикладі киснево-компресорного цеху ПАТ "Запоріжсталь-АГ".*

**Ключові слова:** рівняння регресії, довірчі інтервали, нормальний розподіл.

**Левченко Сергей Андреевич**, кандидат технических наук, доцент

Запорожская государственная инженерная академия. Кафедра электротехники и энергетического менеджмента, г. Запорожье, Украина. *Пр. Ленина 226, г. Запорожье, 69006. Тел.: +38 (066) 6045676.*

*E-mail: levchenko\_s@rambler.ru*

### ПРОГНОЗИРОВАНИЕ ЗАТРАТ ЭЛЕКТРИЧЕСКОЙ ЭНЕРГИИ В ЗАВИСИМОСТИ ОТ ОБЪЕМОВ ИЗГОТОВЛЕННОЙ ПРОДУКЦИИ

*В статье рассмотрено прогнозирование электропотребления с помощью корреляционно-регрессионного анализа и теории вероятностей на перспективу на примере кислородно-компрессорного цеха ПАО "Запорожсталь-АГ".*

**Ключевые слова:** уравнение регрессии, доверительные интервалы, нормальное распределение.

**Sergiy Levchenko, Ph. D.** in Engineering Science

Zaporizhzhya State Engineering Academy. Department of Electrotechnics and energy management, Zaporizhzhya, Ukraine. *Lenina prosp. 226, Zaporizhye, 69006. Tel. +38 (063) 6045676. E-mail: levchenko\_s@rambler.ru*

### PREDICTION POWER CONSUMPTION DEPENDING ON VOLUME OF MANUFACTURED PRODUCTION

*This article considers forecasting power consumption with using correlation regression analysis and probability theory on perspective on the example of the oxygen compressor plant of PSC "Zaporizhstal-AG".*

**Keywords:** regression equation, confidential intervals, normal distribution.

#### Вступ

При нестабільності економічних умов практична потреба в об'єктивному співставленні обсягів виробництва та питомих витрат електроенергії на підприємствах велика, оскільки обумовлена необхідністю, по-перше, прогнозувати електроспоживання при значних змінах обсягів виробництва для формування заявки в енергосистему і, по-друге, оцінювати раціональність використання електроенергії.

В умовах функціонування оптового ринку електроенергії (ОРЕ), для великих і середніх споживачів електричної енергії гостро стоїть питання точності планування електроспоживання. Складання достовірної заявки на споживання електроенергії повинно дозволити знизити витрати на придбання електричної енергії.

Практика показує, що для прогнозування споживання електроенергії не існує загального, єдиного методу: кожне виробництво містить індивідуальні технологічні цикли, які, сумуючись, утворюють унікальний часовий процес. Однак у всіх виробничих циклах споживання енергії можна знайти спільні риси, тим самим утворюючи методичну базу для виконання точного прогнозу.

При наявності великої статистичної бази з електроспоживання та об'ємів виготовленої продукції найбільш доцільним методом прогнозування буде кореляційно-регресійний аналіз з можливістю визначити ймовірність такого прогнозу.

#### Аналіз літературних даних і постановка проблеми

Питаннями прогнозування електроспоживання займалися такі вчені, як Праховник А. В., Пашук В. Я., Копцев Л. А., Філіпов С. П., Кулік М. М. та інші.

Перед тим, як подати заявки з енергоспоживання на ОРЕ на перспективний період, підприємства мають дати попередню оцінку величині електроспоживання. Проте, найчастіше прогнозовані величини не відповідають дійсним фактичним витратам.

Для того, щоб застрахувати підприємство від можливих збитків, пов'язаних з понад договірними витратами електроенергії або перевищенням розміру замовлення, необхідно підвищити точність побудови прогнозу енергоспоживання.

### Мета і задачі дослідження

При достатньому об'ємі статистичних даних, проблему підвищення точності прогнозування графіків електричного навантаження доцільно вирішити шляхом виявлення залежності між об'ємами виготовленої продукції та відповідними витратами електроенергії за допомогою кореляційно-регресійного аналізу.

Не дивлячись на те, що методи кореляційно-регресійного аналізу для прогнозування електричного навантаження вже було застосовано раніше, виникає проблема достовірності такого прогнозу. Для того, щоб підтвердити результати розрахунків, пропонується використовувати теорію ймовірностей.

### Прогнозування електроспоживання за допомогою методів кореляційно-регресійного аналізу та теорії ймовірностей

В якості експериментальних даних, необхідних для проведення дослідження, використано статистичні дані киснево-компресорного цеху ПАТ “Запоріжсталь-АГ” з об'ємів виготовленого кисню та іншої продукції — стисненого повітря, азоту і аргону, та відповідних витрат електричної енергії за період з 2009 до вересня 2013 року за місяцями.

За допомогою засобів MS Excel в якості регресійного рівняння для математичного опису залежності між об'ємами виготовленої продукції та відповідними витратами електричної енергії було підібрано степеневу функцію, яка найбільш точно відповідає статистичним даним:

$$\hat{y}_x = a \cdot x^b \cdot \varepsilon, \quad (1)$$

де  $x$  – об'єм виготовленої продукції;

$y$  – витрати електричної енергії на виготовлення  $x$  об'єму продукції;

$a$  – коефіцієнт рівняння регресії, який показує витрати електроенергії на холостому ході;  $b$  – коефіцієнт рівняння регресії, який показує середню зміну витрат електроенергії із зміною об'єму виготовленої продукції на одну одиницю;

$\varepsilon$  – випадкова помилка, що підлягає мінімізації.

Для оцінки параметрів регресійного рівняння використано підхід, заснований на лінеаризації моделі, який полягає в тому, що за допомогою належних перетворень вихідних змінних досліджувану нелінійну залежність представляють у вигляді лінійного співвідношення між перетвореними змінними, а саме — про логарифмувавши обидві частини рівняння та замінив змінні:

$$y^* = a^* + b^* \cdot x^* + \varepsilon^*, \quad (2)$$

де  $y^* = \lg \hat{y}_x$ ;  $x^* = \lg x$ ;  $a^* = \lg a$ ;  $b^* = b$ ;  $\varepsilon^* = \lg \varepsilon$ .

Далі за методом найменших квадратів даються оцінки параметрів даної лінійної функції:

$$a^* = \frac{\sum (\lg y_i) \cdot \sum (\lg x_i)^2 - \sum (\lg x_i \cdot \lg y_i) \cdot \sum (\lg x_i)}{n \cdot \sum (\lg x_i)^2 - (\sum (\lg x_i))^2},$$

$$b^* = \frac{\sum (\lg x_i \cdot \lg y_i) - a^* \cdot \sum (\lg x_i)}{\sum (\lg x_i)^2} \quad (3)$$

Параметри степеневі функції знаходяться відповідно, як  $10^{a^*}$  та  $10^{b^*}$ .

Для оцінки суттєвості рівняння регресії в цілому розраховується F-критерій Фішера, необхідний для перевірки нульової гіпотези  $H_0$ : факторна і залишкова дисперсії, обумовлені рівнянням регресії, не відрізняються одна від одної.

$$F = \frac{D_{\text{факт}}}{D_{\text{зал}}} = \frac{\sum (\hat{y}_x - \bar{y})^2 \cdot (n-2)}{\sum (y - \hat{y}_x)^2} . \quad (4)$$

Якщо розрахований критерій більше табличного, гіпотеза відхиляється.

Для оцінки тісноти зв'язку між енерговитратами та об'ємами продукції розраховується лінійний показник – індекс кореляції, значення якого знаходяться в межах  $0 \leq R \leq 1$ :

$$R = \sqrt{1 - \frac{\sum (y - \hat{y}_x)^2}{\sum (y - \bar{y})^2}} . \quad (5)$$

Для оцінки долі дисперсії, поясненої рівнянням регресії, розраховується коефіцієнт детермінації, який свідчить про правильність підібраної моделі регресії і дорівнює квадрату індексу кореляції  $R^2$ .

Якість вибраної моделі характеризує середня помилка апроксимації, значення якої мають знаходитися в межах 5–7%:

$$A = \frac{1}{n} \cdot \sum \left| \frac{(y - \hat{y}_x)}{y} \right| \cdot 100 . \quad (6)$$

Для оцінки суттєвості коефіцієнта регресії його величина порівнюється з його стандартною помилкою, тобто визначається фактичне значення t-критерію Стюдента:

$t_b = \frac{b}{m_b}$ , яке потім порівнюється з табличним значенням при певному рівні значущості  $\alpha$  і числі ступенів свободи  $(n - 2)$ .

Величина стандартної помилки спільно з t-розподілом Стюдента при  $n - 2$  ступенях свободи застосовується для перевірки суттєвості коефіцієнта регресії та для розрахунку його довірчих інтервалів:

$$m_b = \sqrt{\frac{\sum (y - \hat{y}_x)^2}{(n-2) \cdot \sum (x - \bar{x})^2}} = \sqrt{\frac{S^2}{\sum (x - \bar{x})^2}} . \quad (7)$$

Якщо фактичне значення t-критерію перевищує табличне, тоді гіпотезу про незначимість коефіцієнту регресії можна відхилити.

Довірчий інтервал для коефіцієнту регресії визначається, як  $b \pm t \cdot m_b$ .

Стандартна помилка параметру  $a$  визначається, як:

$$m_a = \sqrt{\frac{\sum (y - \hat{y}_x)^2}{n-2} \cdot \frac{\sum x^2}{n \cdot \sum (x - \bar{x})^2}} = \sqrt{S^2 \cdot \frac{\sum x^2}{n \cdot \sum (x - \bar{x})^2}} , \quad (8)$$

Процедура оцінки значимості даного параметру така саме, як і для параметру  $b$ ; обчислюється t-критерій:  $t_a = \frac{b}{m_a}$ , його величина порівнюється з табличним значенням при  $n - 2$  ступінях свободи.

Стандартна помилка прогнозованого за лінією регресії значення визначається за формулою, де  $x_k$  — значення змінної  $x$ :

$$m_{\hat{y}_x} = S \cdot \sqrt{\frac{1}{n} + \frac{(x_k - \bar{x})^2}{\sum (x - \bar{x})^2}} , \quad (9)$$

Для прогнозного значення  $\hat{y}_{x_k}$  довірчі інтервали визначаються наступним чином:

$$\hat{y}_{x_k} - t_a \cdot m_{\hat{y}_x} \leq y^* \leq \hat{y}_{x_k} + t_a \cdot m_{\hat{y}_x} . \quad (10)$$

Далі за теорією ймовірностей визначається достовірність даного прогнозу, для чого в якості випадкової величини береться значення питомого електроспоживання.

Для знаходження ймовірності потрапляння величини питомого енергоспоживання в довірчі інтервали, розраховані за кореляційно-регресійним методом, спочатку за допомогою критерію  $\chi^2$  ("хі квадрат") Пірсона перевіряється гіпотеза про нормальний розподіл даної величини.

Перевірка виконується за наступним алгоритмом.

1. Методом добутоків, розділивши генеральну сукупність значень питомого електроспоживання на інтервали, обчислюються вибіркова середня  $\bar{x}_B$  та вибіркове середнє квадратичне відхилення  $\sigma_B$ :

$$\bar{x}_B = M_1^* \cdot h + C ; \quad (11)$$

$$D_B = [M_2^* - (M_1^*)^2] \cdot h^2 , \quad (12)$$

де  $h$  – крок (різниця між двома сусідніми варіантами);

$C$  – хибний нуль (варіанта, яка розміщена приблизно в середині варіаційного ряду);

$M_1^* = \frac{\sum n_i \cdot u_i}{n}$  – умовний момент першого порядку;  $M_2^* = \frac{\sum n_i \cdot u_i^2}{n}$  – умовний момент другого порядку;

$u_i = \frac{x_i - C}{h}$  – умовна варіанта;

$n$  – емпірична частота потрапляння питомого енергоспоживання в інтервал.

Для зменшення помилки, викликані групуванням даних за інтервалами, при розрахунку дисперсії і відповідно середнього квадратичного відхилення роблять поправку Шепарда:

$$\sigma_B = \sqrt{D_B - \left(\frac{1}{12}\right) \cdot h^2} . \quad (13)$$

2. Обчислюються теоретичні частоти потрапляння питомого електроспоживання в інтервал:

$$n_i = \frac{n \cdot h}{\sigma_B} \cdot \varphi(u_i) , \quad (14)$$

де  $n$  – об'єм вибірки (сума всіх частот);

$h$  – крок (різниця між двома сусідніми варіантами);

$$u_i = \frac{x_i - \bar{x}_B}{\sigma_B} ; \quad (15)$$

$$\varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2} . \quad (16)$$

3. Емпіричні та теоретичні частоти порівнюються між собою, для чого:

а) знаходять спостережуване значення критерію Пірсона:

$$\chi^2_{\text{спост}} = \sum \frac{(n_i - n'_i)^2}{n'_i} ; \quad (17)$$

б) за таблицею критичних точок розподілу  $\chi^2$ , за заданим рівнем значущості  $\alpha$  і числу ступенів свободи  $k=S-3$ , де  $S$  – число груп вибірки, знаходять критичну точку  $\chi_{кр}^2(\alpha; k)$  області прийняття нульової гіпотези.

Якщо – немає підстав відхилити гіпотезу про нормальний розподіл генеральної сукупності, тобто, емпіричні та теоретичні частоти відрізняються несуттєво (випадково).

Після підтвердження гіпотези, знаючи густину та функцію розподілу, можна визначити ймовірність потрапляння величини питомого електроспоживання в заданий довірчий інтервал  $(\alpha; \beta)$ :

$$P(\alpha < X < \beta) = \frac{1}{\sigma\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{(x-a)^2}{(2\sigma)^2}} dx = \Phi\left(\frac{\beta-a}{\sigma}\right) - \Phi\left(\frac{\alpha-a}{\sigma}\right), \quad (18)$$

де  $a$  – вибіркова середня  $\bar{x}$ ;

$\Phi\left(\frac{\beta-a}{\sigma}\right)$  і  $\Phi\left(\frac{\alpha-a}{\sigma}\right)$  – функції Лапласа, значення яких знаходяться за таблицею.

Отже, для киснево-компресорного цеху було виявлено наступні залежності витрат електричної енергії на виготовлення об'ємів кисню та іншої продукції відповідно:  $y = 0,345x^{1,047}$  та  $y = 0,186x^{0,972}$ .

Розрахунок необхідних коефіцієнтів та перевірка рівняння регресії в цілому показали правильність підбору функції та її параметрів.

Довірчі інтервали для даних рівнянь відповідно виглядають, як:

$$\hat{y}_{x_k} - 2,005 \cdot m_{\hat{y}_x} \leq y^* \leq \hat{y}_{x_k} + 2,005 \cdot m_{\hat{y}_x};$$

де стандартні помилки прогнозованих значень:

$$m_{\hat{y}_x} = 0,024 \cdot \sqrt{\frac{1}{56} + \frac{(x_k - 4,55)^2}{0,342}};$$

$$m_{\hat{y}_x} = 0,011 \cdot \sqrt{\frac{1}{56} + \frac{(x_k - 5,125)^2}{0,167}}.$$

Графік залежності прогнозного енергоспоживання за даними Бізнес плану на 2013 – 2016 роки від необхідних для виготовлення об'ємів кисню з урахуванням планових витрат та довірчих інтервалів за місяцями зображено на рис. 1, а для іншої продукції – на рис. 2.

Розраховані ймовірності потрапляння прогнозного питомого енергоспоживання в отримані довірчі інтервали: 0,87 та 0,92 відповідно підтверджують високу якість прогнозу.

Аналіз обох графіків показав, що планові витрати електричної енергії наближаються до верхнього довірчого інтервалу прогнозного значення. Проте, ймовірність їх потрапляння в даний інтервал буде значно меншою, що свідчить про доцільність використання прогнозних значень.

### Висновки

Завдяки застосованому кореляційно-регресійному аналізу та теорії ймовірностей прогнозовані значення енергоспоживання максимально наближено до фактичних, що, в свою чергу, призведе до мінімізації грошових витрат, пов'язаних з неточністю прогнозування.

### Список використаних джерел

1. Єлісєєва, І. І. Економетрика: навч. посібник / С. В. Куришева, Т. В. Костєєва та ін.; під заг. ред. І.І. Єлісєєвої. – 2-е вид. – М.: Фінанси і статистика, 2007. – 576 с.
2. Гмурман В. Є. Теорія ймовірностей і математична статистика: навч. посібник для вузів / В.Є. Гмурман – М., Вища школа, 2003. – 479 с.
3. Уварова Л. Г. Прогнозування графіків електричних навантажень підприємства / Л. Г. Уварова, С. А. Левченко//Вісник національного технічного університету «ХПІ». – 2011. – N 53. – С. 122–128.

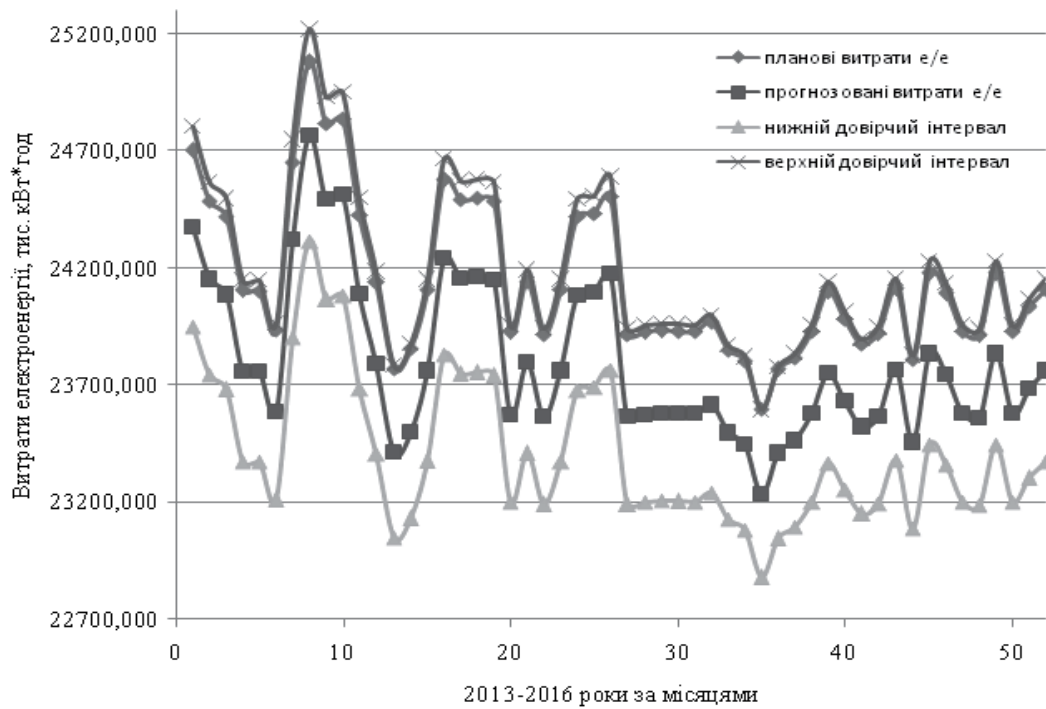


Рис. 1. Графік електричного навантаження в залежності від об'ємів виготовленого кисню на 2013–2016 роки

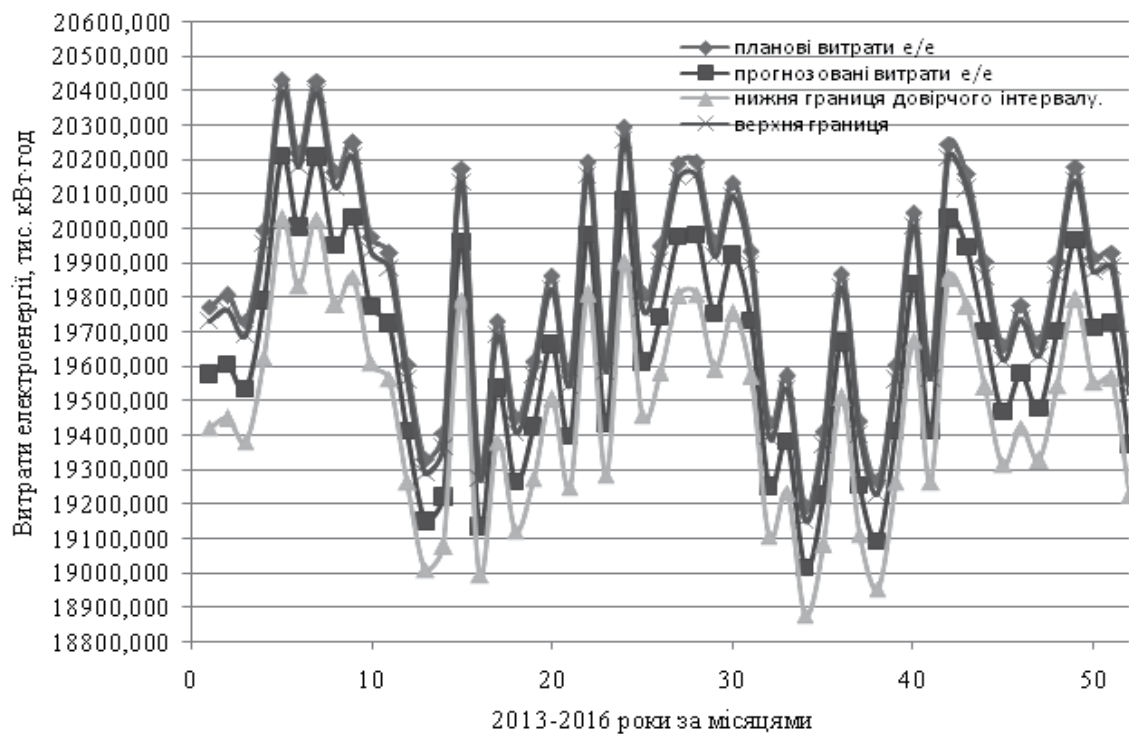


Рис. 2. Графік електричного навантаження в залежності від об'ємів виготовленої іншої продукції на 2013–2016 роки

#### References

1. Eliseeva, I. I. Econometrics [Ekonometrika]. 2007, 576 p.
2. Gmurman, V. E. Probability theory and mathematical statistics [Teoriya ymovirnostey i matematychna statystyka]. 2003, 479 p.
3. Uvarova, L. G. & Levchenko S. A. (2011), Prediction charts electrical loads enterprise [Prognozuvannya grafikiv elektrichnih navantazhen pidpriyemstva], *Proceedings of the National Technical University "HPI"*, no. 53, P. 122-128.

Поступила в редакцию 23.09 2014 г.