

ФОРМУВАННЯ НАБОРУ ІНФОРМАТИВНИХ ВИЗНАЧАЛЬНИХ ЗМІННИХ БАЗОВОГО РІВНЯ ЕЛЕКТРОСПОЖИВАННЯ НАСОСНОЇ СТАНЦІЇ ВОДОПОСТАЧАННЯ

Давиденко Н. В.

Національний університет водного господарства та природокористування (м. Рівне)

Сформовано на основі статистичного аналізу експериментальних даних набір змінних, що мають вплив на ефективність електроспоживання, придатний для визначення базового рівня електроспоживання.

Постановка проблеми. Побудова базового рівня енергоспоживання (БРЕ), унормованого до визначальних змінних, що впливають на енергоспоживання, є основою контролю ефективності енергоспоживання. Отже, об'єктивність контролю залежить від повноти врахування під час побудови математичної моделі електроспоживання всіх чинників, що суттєво впливають на витрату електроенергії на об'єкті.

Аналіз останніх досліджень та публікацій. Більшість параметрів технологічного процесу водопостачання мають ймовірнісний характер, оскільки водоспоживання формується під впливом чинників, багато з яких складно формалізувати. Зважаючи на значну кількість чинників, що впливають на витрату електроенергії, взаємозв'язки між ними та складність математичного опису [1], для вирішення задач контролю ефективності електроспоживання насосної станції (НС) актуальним є використання методів математичного моделювання на основі натурних вимірів значень електроспоживання та визначальних змінних [1], що характеризують технологічний процес водоподачі, зареєстрованих у відповідні моменти часу. Для розробки математичної моделі необхідно мати достатній обсяг експериментальних даних. Впровадження системи моніторингу ефективності функціонування об'єктів водопостачання забезпечує можливість створення великих баз даних, що містять інформацію про об'єми добового водоспоживання, електроспоживання, режими роботи НС та насосних агрегатів (НА), їх технічні та технологічні параметри, показники енергоефективності. Слід зазначити, що прагнення врахувати найбільше потенційно корисної інформації призводить до появи надлишкових (шумових) ознак. Якщо ознака не інформативна, тобто не впливає на відповідь, тоді включення в модель може тільки погіршити її якість. Отже, оскільки на етапах постановки завдання і формування набору даних ще не ясно, які ознаки несуттєві або дублюють один одного, то виникає проблема відбору інформативних ознак.

Мета статті. Підвищення ефективності процедур контролю енергоефективності НС шляхом формування набору інформативних визначальних змінних БРЕ, які мають суттєвий вплив на електроспоживання.

Основні матеріали дослідження. Передумовою процедури моделювання електроспоживання та визначення БРЕ є отримання інформації щодо електроспоживання та набору визначальних змінних БРЕ [1] (технологічних параметрів режиму роботи НС) для створення інформаційної бази даних (БД) та статистичний аналіз отриманих даних. До складу інформаційної БД входить добова статистика ретроспектив-

них виробничих показників, тобто фактично відпрацьованих показників попереднього періоду. Ці показники дозволяють врахувати дійсну зміну обсягів виробництва, характеристики технологічного процесу та фактичні витрати електроенергії на НС. Результатом формування інформаційної БД є формування вибірок ретроспективних даних для вихідної (електроспоживання) та вхідних змінних з урахуванням виявлених в [2] закономірностей у водоспоживанні: для вихідних та робочих днів кожного сезону. Статистична вибірка з інформаційної БД використовується для розробки математичної моделі електроспоживання для типового дня відповідного сезону.

Статистичний аналіз експериментальних даних передбачає визначення описових (дескриптивних) статистик вибірок та має на меті характеристику розподілу імовірності випадкової величини, зокрема, центру, ширини, симетрії розподілу, варіабельності величин та однорідності вибірки тощо [3]. Центр статистичного розподілу характеризують його середнє значення \bar{x} , мода Mo , медіана Me , середнє геометричне \bar{x}_{geom} ; розкид значень у вибірці – найбільше Max та найменше Min значення вибірки, нижній (перший - Q_1) та верхній (третій - Q_3) квартилі розподілу, розмах R , квартильний розмах R_Q , дисперсія σ^2 , середньоквадратичне відхилення σ ; ступінь зміщення вибіркової середньої - коефіцієнт асиметрії (вибірковий коефіцієнт скісності) A_s , який характеризує міру відхилення симетричного розподілу відносно максимальної ординати; кількісну оцінку ступеня відхилення емпіричної кривої розподілу від теоретичної - коефіцієнт ексцесу (вибірковий коефіцієнт гостроверхості) E_k . Для характеристики варіації даних у вибірці використано максимальне та мінімальне значення, дисперсію, середньоквадратичне відхилення; для характеристики міри однорідності вибірок - коефіцієнт варіації V .

З урахуванням результатів класифікації [2] для характерних днів кожного сезону отримано вибірки значень електроспоживання та визначальних змінних [1] (подачі Q , надлишкового тиску ΔH , коефіцієнту ефективності режиму водоподачі K_{ef}^{HNA} та морфометричних показників M_1, M_2, M_3, M_4) і виконано статистичний аналіз отриманих даних. Розрахунок проводиться з використанням стандартних процедур модуля Basic Statistics/Tables (Базові статистики) - Descriptive statistics (Описові статистики) та Correlation matrices (Парні кореляції) пакету STATISTICA StatSoft, Inc. (2001). Результати розрахунку описових статистик для вхідних змінних та електроспоживання НС для робочого дня сезону «Зима» представлено в табл. 1.

Таблиця 1 – Результати розрахунку описових статистик для робочого дня сезону «Зима»

Змінна	\bar{x}	Довірчий інтервал		\bar{x}_{geom}	Me	Mo	Min	Max	Квартиль		R	R_Q	σ^2	σ	s_σ	A_s	s_{A_s}	E_k	s_{E_k}	$V, \%$
		-95%	+95%						Q_1	Q_3										
W	20,434	19,528	21,351	19,94	19,1	19,151	9,84	32,74	16,91	24,71	22,9	7,80	20,887	4,572	0,459	0,385	0,242	-0,661	0,481	22,37
Q	47,971	47,645	48,296	47,943	47,919	47,374	44,73	50,550	46,94	49,304	5,82	2,364	2,667	1,633	0,164	-0,184	0,242	-0,867	0,481	3,40
K_{ef}	0,838	0,832	0,844	0,838	0,837	0,827	0,781	0,883	0,82	0,861	0,101	0,041	0,001	0,028	0,003	-0,184	0,242	-0,867	0,481	3,34
ΔH_Σ	2,592	2,476	2,708	2,529	2,422	2,412	1,248	4,151	2,144	3,133	2,904	0,989	0,336	0,579	0,058	0,385	0,242	-0,661	0,481	22,34
M_1	0,307	0,296	0,313	0,302	0,30	0,311	0,1	0,53	0,27	0,34	0,39	0,07	0,003	0,056	0,006	0,403	0,242	2,062	0,481	18,24
M_2	0,592	0,581	0,603	0,59	0,61	0,612	0,43	0,73	0,56	0,63	0,3	0,07	0,003	0,056	0,006	-0,554	0,242	0,24	0,481	9,46
M_3	0,547	0,526	0,568	0,531	0,53	0,541	0,37	0,91	0,47	0,59	0,54	0,12	0,011	0,106	0,011	1,5	0,242	2,59	0,481	19,38
M_4	0,833	0,826	0,84	0,832	0,84	0,842	0,73	0,92	0,81	0,86	0,19	0,05	0,001	0,035	0,004	-0,394	0,242	0,492	0,481	4,20

Аналогічні розрахунки виконано для типових днів інших сезонів.

Дескриптивні (описові) статистики свідчать про типовість та однорідність даних спостереження для кожної змінної. Коефіцієнт асиметрії та ексцес є відмінними від нуля для всіх вибірок, проте ця відмінність у більшості випадків не є істотною (значення коефіцієнтів за абсолютною величиною не перевищують більше, ніж у 1,5-2 рази значення стандартних відхилень). Це дозволяє зробити висновок про близькість закону розподілу імовірностей вибірок до нормального. Для розподілів, близьких до нормального вважають вибірку однорідною, якщо значення коефіцієнта варіації не перевищує 33%. Отримані результати підтверджують однорідність інформації у сформованих вибірках.

Для побудови математичної моделі електроживлення важливим є аналіз зв'язків між вихідною та вхідними змінними. Для оцінки взаємозв'язку між змінними використовувався метод кореляційного аналізу. Вибірковий коефіцієнт парної кореляції, який є математичною мірою зв'язку між двома факторами

$$r_{ij} = \frac{\sum_{l=1}^n (x_{il} - \bar{x}_i) \cdot (x_{jl} - \bar{x}_j)}{\sqrt{\sum_{l=1}^n (x_{il} - \bar{x}_i)^2} \cdot \sqrt{\sum_{l=1}^n (x_{jl} - \bar{x}_j)^2}}, \quad (1)$$

де x_i, x_j – значення ознак;
 \bar{x}_i, \bar{x}_j – їх середнє значення.

Значення коефіцієнтів кореляції лежать в межах $-1 < r_{ij} < +1$ і характеризують залежність між змінними. Значення $r_{ij} = 0$ свідчить про відсутність статистичного зв'язку між змінними [3]. Наявність великих за модулем (вище 0,7-0,8) значень коефіцієнтів парної кореляції між визначальними змінними свідчить про лінійну залежність між факторами (часткову колінеарність - мультиколінеарність).

Коефіцієнти парної кореляції розраховано для вхідних та вихідних змінних кожної з отриманих ви-

бірок – для характерних днів кожного сезону. В табл.2 наведено результати розрахунку коефіцієнтів кореляції між вихідною змінною (електроспоживанням), та вхідними змінними, щопливають на витрату електроенергії. Аналіз результатів кореляційної залежності дозволив зробити наступні висновки: для всіх вибірок є наявним зв'язок між вихідною та вхідними змінними, проте ступінь тісноти зв'язку є різним. Між електроживленням W та змінною подача Q , та надлишковий тиск ΔH спостерігається сильний та дуже сильний кореляційний зв'язок (згідно шкали Чеддока [4]). Між електроживленням W та коефіцієнтом ефективності режиму водоподачі K_{ef}^{HHA} та морфометричними показниками в основному спостерігається слабкий (в окремих випадках помірний) зв'язок.

Таблиця 2 - Коефіцієнти кореляції

Сезон	Тип дня	W	Q	ΔH	K_{ef}^{HHA}	M_1	M_2	M_3	M_4
Зима	Робочий	1,00	0,81	0,89	-0,19	0,17	0,14	0,33	-0,14
	Вихідний	1,00	0,76	0,88	-0,16	0,23	0,23	0,08	-0,28
Весна-літо-осінь	Робочий	1,00	0,89	0,91	-0,13	0,14	0,15	0,21	-0,18
	Вихідний	1,00	0,71	0,78	-0,37	0,31	0,15	0,22	-0,12
Літо	Робочий	1,00	0,85	0,92	-0,15	0,27	0,14	0,11	-0,12
	Вихідний	1,00	0,71	0,95	-0,23	0,45	0,28	0,13	-0,35

Отримані результати є основою для прийняття рішення щодо формування набору інформативних змінних, що суттєво впливають на витрату електроенергії, для побудови моделі електроживлення. Слід зазначити, що врахування в моделі електроживлення вхідних змінних, що мають незначний вплив, ускладнює процедуру моделювання, проте їх неврахування унеможливило відслідковування причин, що зумовили неефективне електроживлення. Зокрема: 1) за умови вдалого підбору, оптимального режиму роботи НА та їх регулювання повинна виконуватись

умова $\Delta H=0$. Однак, випадковий характер і нерівномірність водоспоживання, інерційність процесу водоподачі зумовлює появу надлишкового тиску в диктуючих точках мережі. Отже, врахування морфометричних показників нерівномірності ГВВ сприятиме виявленню причини виникнення таких ситуацій; 2) в нічну пору доби для подачі води можливе використання лише нерегульованих НА (як базових НА), що зумовлює виникнення надлишкових тисків, а також нераціональні витрати електроенергії за рахунок регулювання водоподачі засувками. Тому, для виявлення причин неефективного електроспоживання необхідним є врахування коефіцієнту K_{ef}^{HNA} ; 3) випадкові зміни водоспоживання зумовлюють неспівпадання часу ввімкнення/вимкнення НА (в тому числі, регульованих НА) згідно запланованого режиму водоподачі з фактичними моментами зростання або зменшення витрати води з мережі. Як наслідок, має місце неефективний режим водоподачі та поява надлишкових тисків. Отже, для моніторингу таких ситуацій необхідним є врахування коефіцієнту енергоефективності режиму водоподачі та морфометричних показників нерівномірності ГВВ.

Таким чином, для остаточного формування набору визначальних змінних, що мають вплив на ефективність електроспоживання, необхідним є логічний аналіз взаємозв'язків, на основі якого буде прийнято остаточне рішення про формування набору інформативних змінних. Разом з тим, слід враховувати, що керовані параметри (наприклад, коефіцієнт енергоефективності режиму водоподачі), які віднесено до визначальних змінних, що впливають на величину витрати електроенергії, внаслідок забезпечення максимально можливого ефективного режиму водоподачі, будуть зведені до оптимальних значень, що забезпечують мінімальні витрати енергії. При цьому варіація цих параметрів знизиться, їх дисперсія зменшиться, що призведе до зниження дисперсії в рівні споживання енергії, обумовленої впливом цих параметрів, тобто в цьому випадку враховувати керовані параметри в моделі енерговитрат недоцільно [5]. У разі залучення до аналізу малокерованого параметра з істотним розкидом його значень, слід чекати значних коливань енергоспоживання, отже, такі параметри слід враховувати при побудові математичної моделі.

Слід також враховувати, що наявність сильної кореляції між окремими визначальними змінними може зумовити виникнення проблем з якістю одержуваних оцінок параметрів моделі, зокрема, нестійкості оцінок, яка виражається в збільшенні статистичної невизначеності (дисперсії оцінок) і, як наслідок, відмінності конкретних результатів оцінки для різних вибірок незважаючи на їх однорідність. Тому, змінну коефіцієнт ефективності режиму водоподачі K_{ef}^{HNA} , яка сильно корелює з подачею та належить до керованих, доцільно виключити з розгляду.

Висновки. На основі статистичного аналізу експериментальних даних підтверджено близькість закону розподілу імовірностей вибірок до нормального та однорідність інформації у сформованих вибірках, що забезпечує можливість використання статистичних методів моделювання електроспоживання та контролю

його ефективності, достовірність результатів яких обумовлена підпорядкованістю закону розподілу імовірностей спостережених змінних нормальному закону. Логічний аналіз взаємозв'язків між електроспоживанням та визначальними змінними дозволив прийняти остаточне рішення про формування набору інформативних змінних, які характеризують технологічний процес водоподачі, мають вплив на ефективність електроспоживання НС та є малокерованими змінними.

Список використаних джерел

1. Давиденко Н. В. Формування кортежу визначальних змінних базового рівня електроспоживання насосної станції водопостачання / Н. В. Давиденко // Вісн. ХНТУСГ ім. П. Василенка. Технічні науки. – 2017. – Вип. 187. – С. 23-25.
2. Давиденко Н. В. Формування груп однотипних добових графіків витрати води з мережі водопостачання методом К-середніх / Н. В. Давиденко // Науково-технічний збірник "Комунальне господарство". – 2018. – Вип. 142. – С. 8-13.
3. Шашков В. Б. Обработка экспериментальных данных и построение эмпирических формул / В. Б. Шашков. – Оренбург: ГОУ ОГУ, 2005.
4. Дубров А. М. Многомерные статистические методы / А. М. Дубров, В. С. Мхитарян, Л. И. Трошин. – Москва: Финансы и статистика, 2003.
5. Півняк Г. Г. Системи енергоменеджменту та їх математичне забезпечення / Г. Г. Півняк та ін.. – Донецьк: Національний гірничий університет, 2013.

Аннотация

ФОРМИРОВАНИЕ НАБОРА ИНФОРМАТИВНЫХ ВЛИЯЮЩИХ ПЕРЕМЕННЫХ БАЗОВОГО УРОВНЯ ЭЛЕКТРОПОТРЕБЛЕНИЯ НАСОСНОЙ СТАНЦИИ ВОДОСНАБЖЕНИЯ

Давыденко Н. В.

Сформирован на основе статистического анализа экспериментальных данных набор переменных, влияющих на эффективность электропотребления, пригодный для определения базового уровня электропотребления.

Abstract

FORMATION OF A SET OF INFORMATIVE RELEVANT VARIABLES OF ENERGY BASELINE OF THE PUMPING WATER SUPPLY STATION

N. Davydenko

The set of variables that influence on the power consumption efficiency, is suitable for determining the energy baseline, has been formed based on the statistical analysis of experimental data.