

МОДЕЛЮВАННЯ ЕЛЕКТРОСПОЖИВАННЯ НАСОСНОЇ СТАНЦІЇ ВОДОПОДАЧІ**Л. В. Давиденко, Н. В. Давиденко**

Луцький національний технічний університет

вул. Львівська, 75, м. Луцьк, 43018, Україна.

E-mail: L.Davydenko033@gmail.com; ninadavydenko1992@gmail.com

Обґрунтовано необхідність застосування для побудови моделі електроспоживання насосної станції методів математичного моделювання на основі натурних вимірів електроспоживання та технологічних змінних процесу водопостачання. На основі аналізу взаємозв'язку електроспоживання та чинників, що мають вплив на його ефективність, сформовано сукупність визначальних змінних. Обґрунтовано доцільність застосування методів самоорганізації моделей складних систем для моделювання електроспоживання за експериментальними даними. Для структурно-параметричної ідентифікації математичної моделі електроспоживання, структура і механізм взаємодії спостережуваних параметрів якої є невідомими, застосовано комбінаторний алгоритм методу групового урахування аргументів. Пошук моделей оптимальної складності виконано в шести класах опорних функцій. Для кожного з них вибрано кращу модель, яка характеризується мінімумом критеріїв відбору. Сформовано сукупність моделей-кандидатів дає змогу вибирати моделі певної складності для забезпечення ефективного контролю ефективності електроспоживання.

Ключові слова: математична модель електроспоживання, метод групового урахування аргументів.

МОДЕЛИРОВАНИЕ ЭЛЕКТРОПОТРЕБЛЕНИЯ НАСОСНОЙ СТАНЦИИ ВОДОПОДАЧИ**Л. В. Давыденко, Н. В. Давыденко**

Луцкий национальный технический университет

ул. Львовская, 75, г. Луцк, 43018, Украина.

E-mail: L.Davydenko033@gmail.com; ninadavydenko1992@gmail.com

Обоснована необходимость применения для построения модели электропотребления насосной станции методов математического моделирования на основе натурных измерений электропотребления и технологических параметров процесса водоснабжения. На основе анализа взаимосвязи электропотребления и факторов, влияющих на его эффективность, сформирована совокупность определяющих переменных. Обоснована целесообразность применения методов самоорганизации моделей сложных систем для моделирования электропотребления по экспериментальным данным. Для структурно-параметрической идентификации математической модели электропотребления, структура и механизм взаимодействия наблюдаемых параметров которой неизвестны, применен комбинаторный алгоритм метода группового учета аргументов. Поиск моделей оптимальной сложности выполнен в шести классах опорных функций. Для каждого из них выбрана модель, которая характеризуется минимумом критериев отбора. Полученная совокупность моделей-кандидатов позволяет выбирать модели определенной сложности для обеспечения эффективного контроля эффективности электропотребления.

Ключевые слова: математическая модель электропотребления, метод группового учета аргументов.

АКТУАЛЬНІСТЬ РОБОТИ. Енергоефективність нині є одним із пріоритетів держав та окремих підприємств. Підвищення рівня енергоефективності потребує впровадження енергозберігаючих технологій, підвищення результативності та ефективності виробничих процесів і енергоспоживання, розвитку системи управління енергоефективністю. Складовою процесу такого управління є контроль ефективності енергоспоживання, що вимагає постійного аналізу енергоспоживання, динаміки показників енергоефективності та виявлення тенденцій до погіршення (покрашення). Початковим етапом організації контролю є виявлення залежності обсягу енергоспоживання на об'єкті від значень низки виробничих і технологічних параметрів (показників), тобто побудова базового рівня енергоспоживання (БРЕ), порівняно з яким й визначається рівень досяжної/досягнутої енергоефективності на об'єкті. БРЕ являє собою математичну модель енергоспоживання, побудовану з урахуванням визначальних змінних - чинників, які суттєво впливають на величину витрати енергії при здійсненні відповідних технологічних процесів [1]. Під час побудови та функціонування систем контролю енергоефективності важливим є коректне вирішення питання побудови математичної моделі, що

використовується для визначення БРЕ на об'єкті, якими методами будувати такі моделі. Від якості математичної моделі, ступеню її адекватності залежить коректність та обґрунтованість результатів контролю ефективності енергоспоживання [2].

Більшість параметрів технологічного процесу водопостачання мають ймовірнісний характер, зумовлений тим, що водоспоживання формується під впливом великого числа чинників, багато з яких складно формалізувати [3]. Тому для побудови математичної моделі електроспоживання насосної станції водопостачання (НС) необхідним є використання методів математичної статистики та математичного моделювання на основі натурних вимірів електроспоживання та змінних, що характеризують технологічний процес водопостачання.

Комплексний підхід до управління ефективністю електроспоживання заснований на багатофакторній моделі витрати електроенергії [4]. Суть задачі моделювання: за результатами n спостережень отримано значення вихідної змінної (електроспоживання) y^j та низки вхідних параметрів x_i^j : $\{y^j, x^j\}$, $\{i=1, \dots, m; j=1, \dots, n\}$; на підставі отриманих даних необхідно визначити залежність $Y = f(x_1, x_2, \dots, x_m)$, яка найкраще описує взаємозв'язок між вхідними парамет-

рами та вихідною змінною. Для побудови математичної моделі слід визначити її структуру та оцінити параметри, щоб вибрати найбільш значимі, тобто, виконати структурно-параметричну ідентифікацію моделі за вибіркою експериментальних даних.

Отже, вирішення проблеми моделювання електроспоживання потребує одночасного розв'язку низки задач: формування достатнього набору статистичних даних, відбору інформативних змінних, вибору структури моделі, забезпечення адекватності моделі.

Метою роботи є побудова адекватної математичної моделі електроспоживання НС шляхом структурно-параметричної ідентифікації залежності електроспоживання від визначальних змінних на основі експериментальних даних, придатної для контролю ефективності електроспоживання.

МАТЕРІАЛ І РЕЗУЛЬТАТИ ДОСЛІДЖЕНЬ. Передумовою моделювання електроспоживання є створення бази даних (БД) про електроспоживання та технологічні параметри режиму роботи НС. До складу БД входить добова статистика ретроспективних виробничих показників - фактично відпрацьованих показників попереднього періоду, які дозволяють врахувати дійсну зміну обсягів виробництва, характеристики технологічного процесу та фактичні витрати електроенергії НС на перекачування води.

Аналіз чинників, що впливають на ефективність електроспоживання [5], дозволяє представити модель електроспоживання НС у вигляді залежності:

$$W_{нс} = f\{Q, K_{ef}^{HHA}, \Delta H_{\Sigma}, M_1, M_2, M_3, M_4\}, \quad (1)$$

де Q – добова подача води в мережу; K_{ef}^{HHA} – коефіцієнт енергоефективності режиму водоподачі, що враховує нераціональні витрати електроенергії за рахунок регулювання водоподачі засувками або байпасуванням (для нерегульованих насосних агрегатів (ННА)); ΔH_{Σ} – сумарний добовий надлишковий тиск; $M_1 - M_4$ – морфометричні параметри, що описують нерівномірність графіка витрати води (ГВВ) [6].

Для побудови математичної моделі електроспоживання важливим є аналіз зв'язків між вихідною та вхідними змінними. Для оцінки взаємозв'язку між змінними використано метод кореляційного аналізу (вибірковий коефіцієнт парної кореляції - табл. 1).

Таблиця 1 – Коефіцієнти кореляції

W	Q	ΔH	K_{ef}^{HHA}	M_1	M_2	M_3	M_4
1,00	0,81	0,89	-0,19	0,17	0,14	0,33	-0,14

Аналіз результатів кореляційної залежності між вихідною змінною (електроспоживанням), та вхідними змінними, що впливають на витрату електроенергії, дозволяє стверджувати: між електроспоживанням (W) та змінною подача (Q) і надлишковий тиск (ΔH) спостерігається сильний кореляційний зв'язок; між електроспоживанням (W) та коефіцієнтом ефективності режиму водоподачі K_{ef}^{HHA} та мор-

фометричними показниками спостерігається слабкий та помірний зв'язок.

Слід зазначити, що врахування в моделі електроспоживання вхідних змінних, що мають незначний вплив, ускладнює процедуру моделювання, проте їх неврахування унеможливує відслідковування причин, що зумовили неефективне електроспоживання. 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 причин виникнення таких ситуацій; 2) в нічну пору доби для подачі води можливе використання лише ННА (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}^{HHA} ; 3) випадкові зміни водоспоживання зумовлюють неспівпадання часу ввімкнення/вимкнення НА 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 K_{ef}^{HHA} та морфометричних показників нерівномірності ГВВ.

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 враховувати керовані параметри в моделі електроспоживання недоцільно [7]. У разі залучення до аналізу малокерованого параметра з істотним розкидом його значень, слід чекати значних коливань енергоспоживання, отже, такі параметри слід ураховувати 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}^{HHA} , яка

сильно корелює з подачею та належить до керованих, доцільно виключити з розгляду. Таким чином, модель електроспоживання НС можна представити у вигляді залежності:

$$W_{nc} = f\{Q, \Delta H_{\Sigma}, M_1, M_2, M_3, M_4\}. \quad (2)$$

Причому структура моделі $f(X)$ невідома.

Одним із методів побудови багатofакторних математичних моделей енергоспоживання є класична процедура побудови рівнянь багатofакторної регресії [1]. Під час її застосування значною проблемою є вибір структури моделі. Мета регресійного аналізу – досягнення мінімуму середньоквадратичної похибки на всіх експериментальних точках для заданого виду рівняння регресії. Проте критерій середньоквадратичної похибки є внутрішнім, оскільки розрахований по тій же інформації, по якій отримана сама модель. Прагнення врахувати найбільше інформації призводить до появи надлишкових ознак. Якщо ознака не інформативна, то її врахування може погіршити якість моделі [1]. Крім того, регресійні моделі, навіть якщо вони є статистично значимими, недостатньо точно відображають прогнозовані тенденції. Побудова складних регресійних моделей, що містять велику кількість незалежних змінних, потребує наявності значних обсягів статистичних даних, збір яких є досить трудомістким [2]. Крім того, сама процедура побудови багатofакторних рівнянь регресії є складною і трудомісткою, потребує відповідних знань та практичних навичок, участі дослідника на етапі визначення структури моделі та під час аналізу мультиколінеарності незалежних змінних, значимості параметрів моделі тощо [2].

Проблема вибору структури моделі стає неістотною у випадку застосування індуктивного підходу, який ґрунтується на принципі «від емпіричних даних до побудови моделей» із застосуванням інформаційних технологій «видобування» знань із даних [8–10], а сучасні обчислювальні можливості дають змогу застосовувати алгоритми, які не тільки замінюють попередній глибокий аналіз інформативності предикторів, а й сприяють більш стійким рішенням. Одним з таких індуктивних алгоритмів є метод групового урахування аргументів (МГУА), заснований на принципі евристичної самоорганізації математичних моделей складних систем [8–10]. Згідно МГУА оптимальна модель в сенсі достовірності та ступеня складності знаходиться за мінімальним значенням низки зовнішніх критеріїв. МГУА є одним з методів розв'язання задач структурно-параметричної ідентифікації моделей складних об'єктів, процесів і систем, їх моделювання за експериментальними даними в умовах невизначеності та неповноти інформації.

МГУА застосовується для аналізу даних, прогнозування та моделювання систем у випадках, коли відсутня або майже відсутня апріорна інформація про структуру моделі та розподіл її параметрів, недостатньо даних спостережень (навіть коли кількість параметрів моделі перевищує кількість спостережень).

МГУА передбачає, що вся інформація про систему, що моделюється, закладена у вихідних даних і не вимагає додаткової інформації про причинно-

наслідкові зв'язки [10]. Побудова моделей за експериментальними даними здійснюється автоматично. В процесі моделювання ставиться завдання виявлення форми залежності і вибору факторів, що мають істотний вплив на залежну змінну. На відміну від звичайних методів регресії, методи самоорганізації моделей дозволяють враховувати множину факторів-претендентів замість фіксованого переліку факторів. При цьому, структура моделі, на відміну від регресійного аналізу, заздалегідь не фіксується, а вибирається з множини варіантів за заданими критеріями. У МГУА здійснюється цілеспрямований перебір структур моделей, що поступово ускладнюються. Застосовування зовнішніх критеріїв дозволяє вибрати єдину модель оптимальної складності. Як спосіб подолання неповноти інформації застосовується принцип зовнішнього доповнення (додаткова перевірна вибірка, точки якої не використовувалися при навчанні системи). Оцінювання параметрів і перевірка якості моделей виконується на різних підвбірках: по одній частині вихідних даних визначають коефіцієнти моделей-претендентів, а за іншою - оцінки моделей за обраними критеріями селекції. Це дозволяє обійтися без апріорних припущень, оскільки поділ вибірки дозволяє автоматично врахувати різні види апріорної невизначеності під час побудови моделі [8]. МГУА забезпечує побудову несуперечливих моделей, що дають стійкий результат незалежно від вихідних даних.

Отже, МГУА є методом пошуку закономірностей (математичного опису закономірних причинно-наслідкових зв'язків) з автоматичним вибором структури і параметрів моделі на основі вибірки статистичних даних, в основу якого покладено принципи зовнішнього доповнення, автоматичної генерації та послідовної селекції ускладнюваних структур моделей [9]. Основний результат застосування теорії МГУА полягає в тому, що для коротких вибірок, неточних або зашумлених даних мінімум критерію дозволяє знайти оптимальну нефізичну модель, точність якої вища і структура якої простіша ніж повної фізичної моделі. Отже, індуктивні алгоритми МГУА дають можливість автоматично знайти знання про об'єкт безпосередньо по вибірці даних [8, 9], виявити взаємозалежності та закономірності, неявно відображені в даних, і представити їх у вигляді математичних моделей оптимальної складності та структури. Підхід індуктивного моделювання за допомогою перебору різних рішень намагається мінімізувати роль дослідника в результатах моделювання.

Моделі, побудовані за алгоритмом МГУА, за своїми прогнозуючими властивостями значно ефективніші, ніж регресійні, оскільки автоматично відбираються аргументи, які є найбільш інформативними та значущими для моделювання. Моделі, одержані методами самоорганізації, мають оптимальну структуру, дають можливість враховувати вплив на процес, що досліджується, великої кількості змінних і потребують для визначення параметрів моделі значно менших обсягів статистичних даних, ніж у випадку застосування класичних алгоритмів регресійного аналізу. Висока точність в МГУА досягається за рахунок оптимізації складної моделі [8].

Оскільки процес електроспоживання НС належить до класу моделей, в яких невідома структура і механізм взаємодії елементів спостережуваних параметрів, то підхід самоорганізації є єдиним надійним засобом моделювання. Широка можливість застосування МГУА та наведені переваги в побудові математичних моделей підтверджують доцільність його використання для побудови математичної моделі електроспоживання НС.

Задача індуктивного моделювання зводиться до вибору оптимальної за заданим критерієм моделі з множини генерованих для об'єкта з одним виходом.

Нехай є n спостережень за поведінкою об'єкту, тобто задана матриця даних $W = (X; y)$, що містить інформацію про зміну m вхідних змінних $X [n \times m]$ і однієї вихідної $y [n \times 1]$, тобто, в просторі змінних розмірністю m сформована навчальна послідовність, в якій кожному рядку матриці X поставлено у відповідність певне значення відгуку Y . Необхідно отримати закон зміни відгуку Y залежно від конкретних значень незалежних змінних X .

Задача побудови залежності між статистичними показниками полягає в побудові функції [9]:

$$y = f(\theta, x) = \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_m x_m, \quad (3)$$

де y – вихідна змінна; $x_i, i = \overline{1, m}$ – m вхідних змінних, які впливають на вихідну змінну - відгук.

Під факторами $x_i, i = \overline{1, m}$ розуміють також функції від вимірюваних незалежних показників (квадрати, змішані добутки, нелінійні перетворення).

Задача полягає у тому, щоб знайти вектор коефіцієнтів θ , тобто підмножину факторів (від 1 до m), яка найкраще описує цільовий показник. Але які саме фактори входять до цієї підмножини, наперед невідомо, що може бути зумовлено такими причинами [8,9]: коефіцієнти парної кореляції між цільовим показником та окремими факторами не дають підстав для виявлення характеру залежності, оскільки істотний вплив має багатовимірна взаємодія між факторами; показники можуть бути тісно пов'язані між собою, тому внутрішні взаємозв'язки між факторами теж впливають на оптимальну структуру моделі.

Виходячи з вигляду загальної моделі, можна сконструювати багато варіантів моделей, які мають різну складність: різний склад і кількість факторів. При цьому, необхідно: задати обсяг вихідної інформації; вказати клас базисних функцій, з яких формується множина кращих моделей; визначити спосіб генерації моделей разом з методом оцінювання параметрів; вибрати критерій порівняння моделей; вказати метод мінімізації критерію якості моделей.

Для побудови індуктивних моделей запропоновані різні алгоритми МГУА. Більшість алгоритмів МГУА використовують поліноміальну базисну функцію. Загальний зв'язок між вхідними і вихідними змінними може бути виражений у вигляді функціонального ряду Вольтера, дискретним аналогом якого є поліном Колмогорова-Габора [9, 11]:

$$y = a_0 + \sum_{i=1}^M a_i x_i + \sum_{i=1}^M \sum_{j=1}^M a_{ij} x_i x_j + \sum_{i=1}^M \sum_{j=1}^M \sum_{k=1}^M a_{ijk} x_i x_j x_k, \quad (4)$$

де $X = (x_1, x_2, \dots, x_m)$ - вхідний вектор змінних; $A = (a_1, a_2, \dots, a_m)$ - вектор коефіцієнтів або ваг.

Компонентами вектора X є незалежні змінні. Генеруючи часткові поліноми із загального полінома (8) з послідовним збільшенням складності моделі S можна вибрати модель оптимальної складності.

Алгоритми МГУА різняться за способом генерування множини моделей-кандидатів, що підлягають перебору за заданим зовнішнім критерієм. У структурі кожного з алгоритмів МГУА виділяють три основні блоки [11]: перетворення вихідних даних відповідно до обраної системи опорних функцій, в якій шукається модель; генерування в обраному базисі множини моделей-претендентів (у порядку підвищення їх складності); обчислення значень критерію селекції, що має властивості зовнішнього доповнення, та послідовного відбору кращих моделей. Компроміс між точністю моделі і її складністю, типовою для критеріїв структурної ідентифікації, може бути виражений явно або неявно. Для зовнішніх критеріїв неявний штраф досягається через поділ вибірки даних. Метод дозволяє одночасно отримати оптимальну структуру моделі і залежність вихідних параметрів від найбільш значимих вхідних змінних.

Для визначення параметрів моделі використовується внутрішній критерій – найчастіше, відносна середньоквадратична похибка [10, 11]:

$$\Delta^2(A) = \sum_{i=1}^{N_A} (y_i - \hat{y}_i)^2 / \sum_{i=1}^{N_A} y_i^2 \rightarrow \min, \quad (5)$$

де y_i – табличні значення вихідної змінної для навчальної вибірки; \hat{y}_i – значення, розраховані за моделлю за змінними x_m навчальної вибірки; N_A – кількість точок навчальної вибірки.

Оптимальність моделі визначається критеріями селекції моделі [10]: регулярності та незміщеності.

Критерій регулярності – відносна середньоквадратична похибка, розрахована на точках, не використаних для побудови моделі [10]:

$$\Delta^2(B) = \sum_{i=1}^{N_B} (y_i - \hat{y}_i)^2 / \sum_{i=1}^{N_B} y_i^2 \rightarrow \min, \quad (6)$$

де y_i – табличні значення вихідної змінної для перевіркової вибірки; \hat{y}_i – значення, розраховані за моделлю для незалежних змінних перевіркової вибірки; N_B – кількість точок перевіркової вибірки (приблизно 30% вибірки), причому, $N = N_A + N_B$.

Як критерії регулярності також використовують коефіцієнт кореляції:

$$R = \sum_{i=1}^{N_B} (y_i \cdot \hat{y}_i) / \sqrt{\sum_{i=1}^{N_B} y_i^2 \cdot \sum_{i=1}^{N_B} \hat{y}_i^2} \rightarrow \max. \quad (7)$$

Чим менше значення середньоквадратичної похибки та більше значення коефіцієнту кореляції для тестової вибірки тим більш регулярна модель.

Критерій незміщеності (мінімуму зсуву) вимагає максимального збігу значень вихідної величини двох моделей, отриманих на двох різних частинах вихідної вибірки даних. Необхідно, щоб моделі, які отримані на підвбірках А і В, якомога менше відрізнялися одна від одної на інтервалі екстраполяції (x_1 ; $x_{\alpha N}$). Близькість цих моделей оцінюється за величиною середньоквадратичної розбіжності їх виходів по всіх точках послідовності [10]:

$$n_{зм}^2 = \sum_1^{\alpha N} (\hat{y}_{Ai} - \hat{y}_{Bi})^2 / \left(\alpha \sum_{i=1}^N y_i^2 \right) \rightarrow \min, \quad (8)$$

де \hat{y}_{Ai} – значення, розраховані за моделлю, отриманою на частині вибірки А; \hat{y}_{Bi} – значення, розраховані за моделлю, отриманою на частині вибірки В; α – коефіцієнт екстраполяції, $\alpha = 1,5 \div 3,0$; y – табличні значення вихідної змінної для точок вихідної вибірки; N – всі точки таблиці вихідних даних.

Критерій незміщеності дозволяє вибрати модель, найменш чутливу до зміни множини експериментальних точок, за якими вона отримана, та встановити закон зв'язку, прихований в зашумлених даних.

Для зменшення обсягу перебору у вигляді обмеження застосовується допоміжний критерій точності (критерій варіації похибки прогнозу) [11]:

$$\delta^2 = \sum_{i=1}^N (y_i - \hat{y}_i)^2 / \sum_{i=1}^N (y_i - \bar{y})^2 \rightarrow \min, \quad (9)$$

де y_i – фактичне значення вихідної змінної; \hat{y}_i – значення вихідної змінної, розраховане по моделі; \bar{y} – середнє значення вихідної змінної.

Якщо: $\delta^2 < 0,5$ - результати моделювання хороші, модель перспективна й може застосовуватися далі; $0,5 \leq \delta^2 < 0,8$ - результати моделювання задовільні, модель слід використовувати обережно; $0,8 \leq \delta^2 < 1$ - використання моделі небажане; $\delta^2 > 1,0$ – модель приносить дезінформацію. Цей критерій рекомендується для оцінки успіху апроксимації.

Особливістю алгоритмів МГУА є способи генерації множини моделей-кандидатів. Серед параметричних алгоритмів більш відомими є комбінаторний алгоритм і багатощаровий ітераційний алгоритм.

Комбінаторний алгоритм – найпростіший з базових алгоритмів МГУА. Ідея алгоритму полягає в організації повного перебору всіх можливих комбінацій змінних в моделі в рамках заданого класу з урахуванням обмежень з метою знаходження структури і параметрів, які мінімізують значення зовнішнього критерію якості моделей. Основний комбінаторний алгоритм МГУА має багаторядну ітераційну структуру. Особливість СОМВІ-алгоритму [11]: правило ітерації не залишається постійним, а розширюється з кожним новим рядом, тобто, при підборі моделей їх складність поступово збільшується. Загальна схема СОМВІ-алгоритму містить операції: визначення коефіцієнтів окремих моделей при складності $s = 1$,

$s = 2, \dots, s = n$, де s – кількість коефіцієнтів, що оцінюються (за методом найменших квадратів); розрахунок значення зовнішнього індивідуального або комбінованого критерію селекції; вибір моделі оптимальної складності за мінімумом критерію.

Комбінаторний алгоритм МГУА базується на методі повної математичної індукції, оскільки не пропускається жоден із можливих варіантів моделі, вказаних у вихідному повному базисі. Організація комбінаторного перебору здійснюється заданням опорної функції, з якої способом занулення коефіцієнтів виходить певна кількість часткових описів. Для кожної пари аргументів методом найменших квадратів знаходяться коефіцієнти рівнянь всіх часткових описів. Оптимальна модель знаходиться з умови мінімуму зовнішніх критеріїв. Модель вибирається з множини моделей-претендентів послідовним відбором спочатку за критерієм регулярності, потім – за критерієм мінімуму зсуву [10] (зовнішніми критеріями якості моделі), що підвищує стійкість моделі. Оцінка коефіцієнтів на навчальній вибірці і вибір оптимальної структури моделі на екзаменаційної частини вибірки здійснюється багаторазово, тобто для всієї множини моделей-кандидатів.

Моделювання електроспоживання НС виконувалося в програмному середовищі GMDH Shell DS 3.8.2 [12]. Застосовано комбінаторний алгоритм. Пошук моделей оптимальної складності виконувався в таких класах опорних функцій:

1) клас моделей лінійного опису – лінійні функції від n змінних (клас № 1):

$$y = a_0 + \sum_{i=1}^M a_i x_i. \quad (10)$$

2) клас моделей лінійного опису з ефектом взаємодії (клас № 2):

$$y = a_0 + \sum_{i=1}^M a_i x_i + \sum_{i=1}^M \sum_{j=1}^M a_{ij} x_i x_j. \quad (11)$$

3) клас моделей нелінійного опису – нелінійні функції від n змінних з обмеженням ступеня p : $p=2$ – квадратичні функції від n змінних (клас № 3); $p=3$ – кубічні функції від n змінних (клас № 5):

$$y = a_0 + \sum_{i=1}^M a_i x_i^p. \quad (12)$$

4) клас поліноміальних моделей з ефектом взаємодії – поліноміальні функції від n змінних з обмеженням ступеня p з ефектом взаємодії: $p=2$ – квадратичні функції від n змінних (клас № 4); $p=3$ – кубічні функції від n змінних (клас № 6):

$$y = a_0 + \sum_{i=1}^M a_i x_i^p + \sum_{i=1}^M \sum_{j=1}^M a_{ij} x_i^p x_j^q. \quad (13)$$

Оскільки індуктивний підхід має справу з фіксованим класом моделей, то відшукується оптимальна модель лише в заданому класі. Одним з етапів є перевірка її достовірності та апробація на реальних

даних. Про ступінь валідації моделі свідчить її здатність правильно визначати вихідну величину. Стратегія валідації використовується для вибору методу перевірки моделі. Одна з них - спосіб формування навчальної вибірки (для побудови моделі) і тестової (для її перевірки). Хороша модель дає прийнятні результати точності на навчальній і контрольній вибірці. Як стратегію валідації моделей (алгоритму використання «зовнішнього доповнення») вибрано перехресну перевірку (крос-валідацію) по К-блоках [14]. Суть стратегії: деяку фіксовану частку (перші 10% зразків) виключають з вихідного набору даних, будують модель (використовуючи 90% даних, що

залишилися) і застосовують її до виключеного набору; на наступному циклі видалені дані повертаються, видаляються інші 10% даних, знову будується модель, яка застосовується до виключених даних. Процедура повторюється доти, доки всі дані не побувають в числі виключених.

У результаті використання МГУА отримано модель оптимальної складності для кожного класу опорних функцій. Отримані моделі забезпечують мінімум внутрішнього критерію $\Delta^2(A)$, критерію регулярності $\Delta^2(B)$, незміщеності $n^2_{зм}$ та варіації похибки прогнозу δ^2 , а також максимум коефіцієнта кореляції R в обраному класі функцій (табл. 2).

Таблиця 2 – Математичні моделі електроспоживання та їх характеристика

Клас опорної функції	Вид моделі	Характеристика математичної моделі					
		Складність	$\Delta^2(A)$	$\Delta^2(B)$	$n^2_{зм}$	δ^2	R
№ 1	$W = 2,409 + \Delta H \cdot 0,976 + M_3 \cdot (-0,211)$	3	0,00091	0,0016	0,00059	0,00001	0,9998
№ 2	$W = 8,204 + Q \cdot 0,516 + Q \cdot \Delta H \cdot 0,334 + Q \cdot M_4 \cdot 0,492 + \Delta H \cdot 2,394 + M_1 \cdot (-4,558) + M_1 \cdot M_3 \cdot 9,527 + M_2 \cdot M_3 \cdot (-5,892) + M_2 \cdot M_4 \cdot 6,92 + M_4 \cdot (-30,21)$	10	0,00019	0,0052	0,00023	0,0011	0,9942
№ 3	$W = 8,704 + \Delta H^2 \cdot 0,0372 + M_3^2 \cdot (-0,239)$	3	0,00089	0,0027	0,00058	0,000027	0,9992
№ 4	$W = 8,807 + Q \cdot M_3 \cdot 0,007 + \Delta H^2 \cdot 0,037$	3	0,00093	0,0029	0,0185	0,000031	0,9992
№ 5	$W = 10,804 + \Delta H^3 \cdot 0,002 + M_3^3 \cdot (-0,321)$	3	0,0012	0,0043	0,00087	0,00021	0,9982
№ 6	$W = -131,776 + Q \cdot M_2^2 \cdot 1,563 + Q^2 \cdot M_2 \cdot (-0,011) + Q^2 \cdot M_4 \cdot (-0,001) + \Delta H^3 \cdot 0,002 + M_1 \cdot M_3 \cdot (-1,047) + M_2 \cdot 703,867 + M_2^2 \cdot (-1146,05) + M_2^3 \cdot 566,246$	9	0,00018	0,0099	0,00017	0,0018	0,9978

Отже, ускладнення моделі та підвищення степеню полінома не призводить до значного підвищення якості результатів моделювання та точності прогнозування. Всі моделі характеризуються мінімумом критеріїв відбору. Це дає змогу вибирати моделі невеликої складності для забезпечення ефективного контролю ефективності електроспоживання.

ВИСНОВКИ. Застосування комбінаційного алгоритму МГУА до побудови математичної моделі електроспоживання НС (структура і механізм взаємодії спостережуваних параметрів якої є невідомими) дало змогу шляхом перебору усіх можливих моделей із заданого базису вибрати форму залежності, тобто, здійснити структурно-параметричну ідентифікацію моделі за експериментальними даними, та сформувані сукупність оптимальних в своєму класі опорних функцій нефізичних моделей-кандидатів.

Автори висловлюють подяку дирекції та персоналу КП «Луцькводоканал» за сприяння у проведенні досліджень та надання необхідної інформації.

ЛІТЕРАТУРА

- Находов В. Ф., Бориченко О. В., Іванько Д. О., Єгорова І. О. Комплексний підхід до визначення складу чинників, що впливають на величину енергоспоживання при впровадженні систем оперативного контролю енергоефективності. *Енергетика: економіка, технології, екологія*. 2014. № 2. С. 68–77.
- Находов В. Ф., Стеценко І. В., Бедерак Я. С. Застосування методів самоорганізації математичних моделей енергоспоживання для встановлення «стандартів» в системах оперативного контролю енергоефективності. *Енергосбереження. Енергетика. Енергоаудит*. 2012. № 5 (99). С. 23–33.
- Якимчук Н. М. Моделювання системи автоматичного керування з нечіткою логікою об'єктами

водопостачання. *Вісник Інженерної академії України*. 2013. № 3-4. С. 145–150.

- Грунтович Н. В., Мороз Д. Р., Капанський А. А. Развитие методического обеспечения диагностирования и прогнозирования энергоэффективности технологических систем водоснабжения и водоотведения. *Енергоефективність*. 2015. № 1. С. 20–23.
- Давиденко Н. В. Формування кортежу визначальних змінних базового рівня електроспоживання насосної станції водопостачання. *Вісник ХНТУСГ ім. П. Василенка. Технічні науки*. 2017. Вип. 187. С. 23–25.
- Давиденко Л. В., Давиденко Н. В. Побудова правил дискримінації добових графіків витрати води з мережі водопостачання з урахуванням сезонних та соціальних чинників. *Вісник Кременчуцького національного університету імені Михайла Остроградського. Випуск 3/2018 (110)*. С. 20–25.
- Системи енергоменеджменту та їх математичне забезпечення. Півняк Г. Г., Випанасенко С. У. та ін. Дніпропетровськ: *Національний гірничий університет*, 2013. 214 с.
- Anastasakis L., Mort N. The development of self-organization techniques in modelling: A review of the Group Method of Data Handling (GMDH). *The University of Sheffield. United Kingdom*, 2001. Research Report No.813. 39p.
- Степашко В. С., Єфіменко С. М., Савченко Є. А. Комп'ютерний експеримент в індуктивному моделюванні. Київ: Наукова думка. 2014. 222 с.
- Стеценко І. В. Моделювання систем. Черкаси: ЧДТУ, 2010. 399 с.
- Івахненко О. Г., Івахненко Г. О. Індуктивні методи прогнозування та аналізу складних економічних систем. *Економіст*. 1998. №.5. С. 88–96.
- GMDH Shell for Data Science: GMDH Shell documentation. New York: GMDH LLC, 2017. [Online]. Available: <https://gmdhsoftware.com/>.

MODELING OF POWER CONSUMPTION OF WATER SUPPLY PUMPING STATION

L. Davydenko, N. Davydenko

Lutsk National Technical University

vul. Lvivska, 75, Lutsk, 43018, Ukraine. E-mail: L.Davydenko033@gmail.com; ninadavydenko1992@gmail.com

Purpose. The article aims to construct a mathematical model by structural and parametric identification of the dependence of the pumping station power consumption on relevant variables based on the experimental data that is suitable for controlling the power consumption efficiency. **Methodology.** Correlation analysis and logical analysis have been used to characterize the interconnection between power consumption and factors that influence on its efficiency. The method of self-organization of models of complex systems, in particular, the combinatorial algorithm of the group method of data handling, has been used for modeling of power consumption by experimental data. Cross-validation on K-blocks has been used to verify the reliability and approbation of the power consumption model. **Results.** The necessity of mathematical modeling of pumping station power consumption based on the experimental data on power consumption and technological parameters of water supply process has been substantiated. The set of relevant variables has been formed based on the analysis of the interconnection of power consumption and factors that influence on its efficiency. The obtained set contains the main parameters of the water supply process and the morphometric parameters that characterize the unevenness of the water supply graph. The expediency of application self-organization methods of models of complex systems for modeling of power consumption by experimental data has been substantiated. The structure and parameters of the mathematical model of power consumption have been selected using the combinatorial algorithm of the group method of data handling. The search for optimal complexity models has been performed in six classes of basic functions. The best model for each class, which is characterized by a minimum of selection criteria, has been selected. **Originality.** The power consumption model of the water supply pumping station, which is based on the group consideration of the technological parameters of the water supply process and the characteristics of the unevenness of the water supply graph from the water supply network, has been constructed. **Practical value.** The formed set of candidate models allows choosing models of some complexity to ensure effective control of the power consumption efficiency. **Conclusions.** The use of the combinatorial GMDH algorithm has been enabled to automatically accomplish the structural and parametric identification of the power consumption model of the pumping station according to experimental data and form a set of non-physical candidate models that are optimal in their class of basic functions. References 12, tables 2.

Key words: mathematical model of power consumption, group method of data handling.

REFERENCES

1. Nahodov, V., Borychenko, O., Ivanko, D., Yegorova, I. (2014), «An integrated approach to determine the composition of the factors that affect the amount of energy in the implementation of operational control efficiency», *Energy: Economics, Technology, Ecology*, № 2, pp. 68-77.
2. Nahodov, V., Stetsenko, I., Bedarak, Ya. (2012), «Application of mathematical models for the installation “standard” in the systems of operational control of energy efficiency», *Energy saving. Power engineering. Energy audit*, no. 5 (99), pp. 23-33.
3. Yakymchuk, N. (2013) «Modeling of the automatic control system with fuzzy logic of water supply facilities», *Bulletin of Engineering Academy of Ukraine*, no. 3-4, pp. 145-150.
4. Gruntovych, N., Moroz, D., Kapanskyi, A. (2015), «Development of methodical support of diagnostics and prediction of energy efficiency of technological systems of water supply and sanitation», *Energy Efficiency*, no. 1, pp. 20-23.
5. Davydenko, N. (2017), «Formation of the relevant variables set of the energy baseline of the water supply pumping station», *Transactions of KhNTUSH after P. Vasylenko, Tehnichal Sciences*, Vol. 187, pp. 23-25.
6. Davydenko, L., Davydenko, N. (2018), «Construction of discrimination rules of daily water consumption graphs from the water supply network with consideration of seasonal and social factors», *Transactions of Kremenichuk Mykhailo Ostrohradskyi National University*, Issue 3, no. 110, pp. 20-25.
7. Pivnyak, G., Vipanassenko, S. (2013), *Systemy enerhomenedzhmentu ta ikh matematychnye zabezpechennia* [Energy management systems and their mathematical support], National Mining University, Dnipropetrovsk, Ukraine.
8. Anastasakis, L., Mort, N. (2001), «The development of self-organization techniques in modelling: A review of the Group Method of Data Handling, (GMDH)», *The University of Sheffield*. United Kingdom, Research Report No.813, 39p.
9. Stepashko, V., Yefimenko, S., Savchenko, E. (2014), *Komputernyi eksperyment v induktyvnomu modeliuvanni* [Computer experiment in inductive modeling], Scientific Thought, Kyiv, Ukraine.
10. Stetsenko, I. (2010), *Modeliuvannia system* [Modeling of systems], Cherkasy State Technological University, Cherkasy, Ukraine.
11. Ivakhnenko, O., Ivakhnenko, G. (1998), «Inductive methods of forecasting and analysis of complex economic systems», *Ekonomist*, no. 5, pp. 88-96.
12. (2017) GMDH Shell for Data Science: GMDH Shell documentation. New York: GMDH LLC. Available: <https://gmdhsoftware.com/>.

Стаття надійшла 24.01.2019.