

ЕКОНОМІКА ECONOMICS

УДК 621

В.В. ГУЗ, аспірант

ОПЕРАТИВНЕ ПРОГНОЗУВАННЯ ЦІН ПОСТАВОК ЕЛЕКТРОЕНЕРГІЇ НА ОПТОВОМУ РИНКУ

В даній роботі було проаналізовано доступні методи для процесу оперативного прогнозування цін на оптовому ринку електроенергії України та запропоновано використання Методу Групового Урахування Аргументів (МГУА). МГУА дозволяє знаходити оптимальну, з точки зору складності, структури та точності передбачення, модель, що описує динаміку зміни цін та дозволяє враховувати нестационарність вхідних даних. Більше того, запропоновано використання зовнішніх доповнень, що базуються на вхідних даних при утворенні таких моделей.

Ключові слова: Метод Групового Урахування Аргументів, планування поставок, оптовий ринок, електроенергія, комп'ютерна модель, прогнозування

Вступ

Наразі існує широкий набір математичних методів для вирішення задач прогнозування. Серед них присутні Генетичні Алгоритми (ГА), Штучні Нейронні Мережі (ШНМ), класичний регресійний аналіз, дистрибутивно-лагові моделі.

Останнім часом, моделювання з використанням методів машинного навчання, як то нейронні мережі, нечітка логіка й генетичні алгоритми, стало популярним підходом для вирішення складних нелінійних задач, що не вимагає вичерпних теоретичних знань з явищ, що моделюються. Такі підходи залежать, насамперед, від експериментальних даних про процес, які зазвичай доступні у великих кількостях, а не від точних теоретичних знань. Модель, що описує явище, створюється на основі даних, які подаються у вигляді достатньої кількості зразків для навчання. Після того, як модель синтезується, вона може бути використана для передбачення результатів на основі даних, яких не було в зразках для навчання.

Метод має ряд переваг у порівнянні з традиційними підходами, у тому числі підвищену стійкість до шуму й невизначеності, зниженню потреби в даних про явище, і відносній легкості розробки та оновлення моделі. В наступних розділах буде розглянуто використання методів ГА, регресії, дистрибутивно-лагових моделей для завдань прогнозування та їх теоретична застосовність для даної предметної області.

Мета і завдання дослідження

Метою роботи є підвищення точності прогнозу цін постачання електроенергії на оптовому ринку електроенергії України, за рахунок математичного визначення і формалізації факторів впливу та тренд зміни цін за допомогою застосування методів групового урахування аргументів.

Для досягнення цієї мети були поставлені наступні основні завдання:

- провести аналіз сучасного стану проблеми прогнозування цін постачання електроенергії на оптовому ринку;
- зробити вибір та обґрунтувати засоби економії обчислювальних ресурсів для точного прогнозування цін постачання електроенергії на ОПЕ.

Матеріал і результати досліджень

Використання генетичних алгоритмів для завдань оптимізації

Одним з методів знаходження екстремумів складних функцій сьогодні є генетичні алгоритми (ГА).

ГА є однією зі складових еволюційного моделювання як наукового напрямку, що базується на принципах природного відбору за Ч. Дарвіном. Генетичний алгоритм вперше був запропонований в 1975 році в Мічиганському університеті Джоном Холландом [1]. Він отримав назву репродуктивного плану Холланда і надалі активно використовувався в якості базового алгоритму в еволюційних обчисленнях. Подальший розвиток ГА, як власне і свою назву, отримали в роботах Goldberg D.E. (Гольдберга) в лабораторії ГА Іллінойсу, De Jong K. A. (Кеннет Де Йонг) в університеті Джорджа Мейсона, Вірджинія і їх учнів.

Розглянемо базовий ГА. [1], [2]

Нехай S - деяка система або процес. Її атрибутами є X - вектор вхідних параметрів, Y - вектор вихідних змінних.

Нехай закон $Y = f(X)$ ідентифікований і залежність f досить складна. Відомі також межі можливих змін складових вектора X . Необхідно знайти такі значення вектора X , щоб значення Y було оптимальним.

Функція f називається фітнес-функцією. Можливі значення елемента вектора X є його фенотипом. Двійкове подання фенотипу є генотип (напр. 12345 \rightarrow 100011). Генотип має певну кількість елементів. Один або кілька генотипів (за кількістю елементів X) представляють хромосому. Кросовером називають поділ двох хромосом та обмін частинами (напр. 1100 і 1010 \rightarrow 1110 і 1000). Мутація - інвертування одного з елементів хромосоми (напр. 0000 \rightarrow 0100). Інверсія - зміна порядку проходження частин хромосоми (напр. 1100 \rightarrow 0011).

1) Ініціювати початковий момент часу $t = 0$. Випадковим чином сформувати початкову популяцію, що складається з k особин. $B_0 = \{A_1, A_2, \dots, A_k\}$.

2) Обчислити пристосованість кожної особини $F_{A_i} = \text{fit}(A_i), i = 1 \dots k$ і популяції в цілому $F_t = \text{fit}(B_t)$ (фітнес-функція).

Значення цієї функції визначає наскільки добре підходить особина, описана даною хромосомою, для вирішення завдання.

3) Вибрати особину A_c з популяції. $A_c = \text{Get}(B_t)$

4) З певною вірогідністю (ймовірністю кросовера P_c) вибрати другу особину з популяції $A_{c1} = \text{Get}(B_t)$ і провести оператор кросовера $A_c = \text{Crossing}(A_c, A_{c1})$.

5) З певною вірогідністю (ймовірністю мутації P_m) виконати оператор мутації $A_c = \text{mutation}(A_c)$.

6) З певною вірогідністю (ймовірністю інверсії P_i) виконати оператор інверсії $A_c = \text{inversion}(A_c)$.

7) Помістити отриману хромосому в нову популяцію $\text{insert}(B_{t+1}, A_c)$.

8) Виконати операції, починаючи з пункту 3, k разів.

9) Збільшити номер поточної епохи $t = t + 1$.

10) Якщо виконана умова зупинки, то завершити роботу, інакше перехід на крок 2.

ГА використовуються тоді, коли комбінаторна складність методу повного перебору робить його непридатним через час, необхідний для знаходження розв'язку.

ГА відрізняються інструментами, що використовуються при його використанні. Так, наприклад, кросовер може бути як однокрапковим так і двокрапковим.

Ймовірності вибору батьківських пар теж можуть визначатися по-різному. Відомі такі способи [3]:

- панміксія, коли батьки вибираються з популяції випадковим чином, таким чином, що один батько може бути в парі з самим собою або брати участь в декількох парах;
- селекція, коли значення функції пристосованості батьків вище середнього значення по популяції;
- інбридинг, коли перший батько обирається випадковим чином, а другий батько з більшою ймовірністю є найближчий до першого в популяції;
- аутбридинг, коли перший батько обирається випадковим чином, а другим батько з більшою ймовірністю є найвіддаленіший від першого в популяції;
- пропорційний, коли батьки вибираються з імовірностями, пропорційними їх значенням функції пристосованості.

Інбридинг і аутбридинг буває генотипним та фенотипним. Існує також два механізми відбору членів нової популяції: елітний та відбір з витісненням. У першому випадку нова популяція складається з найкращих членів репродукційної групи, яка об'єднує в собі батьків та дітей. При відборі з витісненням факт внесення члена репродукційної групи в нову популяцію буде залежати не тільки від величини її пристосованості, а й тим, чи є в новій популяції особина з аналогічним набором хромосом. Варіантів алгоритмів ГА існує безліч, і щороку з'являються все нові й нові.

Існує кілька приводів для критики на рахунок використання генетичного алгоритму в порівнянні з іншими методами оптимізації:

- повторна оцінка функції пристосованості (фітнес-функції) для складних проблем, часто є чинником, що обмежує використання алгоритмів штучної еволюції. Пошук оптимального рішення для складної задачі високої розмірності часто вимагає дуже витратною оцінки функції пристосованості. У реальних задачах, таких як задачі структурної оптимізації, єдиний запуск функціональної оцінки вимагає від кількох годин до декількох днів для проведення необхідних обчислень.

- генетичні алгоритми погано масштабуються під складність розв'язуваної проблеми. Це означає, що число елементів, які будуть мутувати дуже велике, якщо великий розмір області пошуку рішень. Це робить використання наявної обчислювальної техніки надзвичайно складним при вирішенні таких проблем, як, наприклад, проектування двигуна, будинку або літака. Для того щоб зробити так, щоб такі проблеми піддавалися еволюційним алгоритмам, вони повинні бути розділені на найпростіші представлення даних проблем. Таким чином, еволюційні обчислення використовуються, наприклад, при розробці форми лопатей, замість всього двигуна, форми будівлі, замість докладного будівельного проекту та форми фюзеляжу, замість розробки вигляду всього літака. Друга проблема полягає в питанні, як захистити частини, які еволюціонували і є високопридатними від подальшої руйнівної мутації, зокрема тоді, коли від них вимагається хороша сумісність з іншими частинами в процесі оцінки придатності.
- рішення є більш придатним лише в порівнянні з іншими рішеннями. В результаті умова зупинки алгоритму не є чіткої для кожної окремої проблеми.
- у багатьох задачах генетичні алгоритми мають тенденцію сходитися до локального оптимуму або навіть до спірних точок, замість глобального оптимуму даної задачі.

ШНМ

ШНМ для оперативного прогнозування цін поставок електроенергії на оптовому ринку було запропоновано використовувати в роботі [4]. Не дивлячись на очевидні переваги ШНМ, моделі, що отримуються на виході, при навчанні є надто складними або часто недоступними. Додатковими перевагами ШНМ є такі властивості як здатність до узагальнень та хороша апроксимація нелінійних систем. Загалом, підхід передбачення за допомогою нейронних мереж страждає одразу від певного числа обмежень, включаючи складність визначення оптимальної топології мережі та параметрів тренування [5]. У визначенні численних критичних параметрів є широкий вибір та мало доступних рекомендацій [6], із-за чого розробники часто схиляються до підходу проб та помилок, що може бути громіздким та затратним [7], [8]. В ці параметри входять число та кількість прихованих шарів, тип порогової функції нейрона для різних шарів, коефіцієнт навчання та коефіцієнт моменту, критерій зупинки навчання для того щоб уникнути перенавчання та забезпечити адекватний рівень узагальнення для нових даних. Ще одним обмеженням є природа «чорного ящика» нейромережових моделей, що дає низький рівень розуміння змодельованих відношень та відносної значущості різних входів, таким чином не забезпечуючи, в достатньому рівні, засобів для пояснення моделі, що була збудована [9]. Прийнятність, впевненість у результатах засобів автоматичного передбачення в таких областях як прогнозування електричного навантаження, контроль за викидами та медична діагностика пов'язані з їх прозорістю та змогою виправдати результати людських операторів та експертів [10]. Для того, щоб подолати дані обмеження пропонується використати підхід з використанням Методу Групового Урахування Аргументів (МГУА) [11], [12], [13]. Порівнюючи з нейромережами, МГУА пропонує переваги швидшого знаходження моделі, потребуючи при цьому мінімального втручання зі сторони користувача, швидшої збіжності під час синтезування моделі без проблем застрягання у локальних мінімумах, ефективного вибору вхідних змінних та автоматичного вибору структури моделі [5]. З моделями, що представлені як ієрархія поліноміальних виразів, відношення результатної аналітичної моделі можуть забезпечити розуміння змодельованого явища, підкреслюючи внесок різних змінних та дозволяють виконувати порівняння з попередніми емпіричними або статистичними моделями. Підхід автоматично уникає перенавчання за допомогою використання критерію регуляризації, що заснований на униканні складних моделей [13] без обов'язкової додаткової валідації даних під час навчання, так само як і у випадку нейромережових парадигм.

Використання регресійного аналізу для завдань прогнозування

Під терміном «регресія» розуміють рух назад, повернення до попереднього стану. Названий термін ввів у кінці XIX ст. Френсіс Галтон. В результаті проведеного ним дослідження зв'язку між зростом батьків і дітей, виявилось, що наявна обернена залежність. Так, у батьків з дуже високим зростом діти мають менший зріст порівняно з середнім зростом батьків. І, навпаки, у дуже низьких батьків середній зріст дітей вищий. В одному і другому випадку середній зріст дітей прямує (повертається) до середнього зросту населення певної місцевості. Саме такою залежністю і пояснюють вибір терміна «регресія» [14].

Регресійний аналіз - статистичний метод дослідження впливу однієї або декількох незалежних змінних на залежну змінну. Незалежні змінні інакше називають регресорами або предикторами, а залежні змінні - критеріальними. Термінологія залежних і незалежних змінних відображає лише математичну залежність змінних, а не причинно-наслідкові зв'язки. Регресійний аналіз проводиться на основі побудованого рівняння регресії і визначає внесок кожної незалежної змінної у варіацію досліджуваної (прогнозованої) залежної змінної величини. Основним завданням регресійного аналізу є визначення впливу факторів на результативний показник (в абсолютних показниках). Передусім для цього необхідно підібрати та обґрунтувати рівняння зв'язку, що відповідає характеру аналітичної стохастичної залежності між досліджуваними ознаками. Рівняння регресії показує як в середньому

змінюється результативна ознака Y_x під впливом зміни факторних ознак x_i . У загальному вигляді рівняння регресії можна представити так:

$$Y_x = f(x_1, x_2, \dots, x_n),$$

де Y_x – залежна змінна величина; x – незалежні змінні величини (фактори).

Залежно від кількості змінних величин виділяють різні види регресійного аналізу. Якщо змінна величина завжди одна, то змінних може бути як одна, так і декілька. Виходячи з цього, виділяють два види регресійного аналізу: парний (простий) регресійний аналіз і регресійний аналіз на основі множинної регресії, або багатофакторний. Парний регресійний аналіз – вид регресійного аналізу, що включає у себе розгляд однієї незалежної змінної величини, а багатофакторний – відповідно дві величини і більше.

Зважаючи на характер зв'язку, в регресійному аналізі можуть використовуватися лінійні та нелінійні функції. Для визначення характеру залежності та, відповідно, побудови рівняння регресії доцільно застосувати графічний метод, порівняння рівнобіжних рядів вихідних даних, табличний метод. Так, графічний метод дає найбільш наочну картину розміщення крапок на графіку, завдяки чому можна виявити напрям і вид залежності між досліджуваними показниками: прямолінійна чи криволінійна.

За допомогою порівняння рівнобіжних рядів ознак можна спостерігати за рівномірністю їх взаємних змін. Якщо зміна факторної ознаки x призводить до відносно рівномірної зміни результативної Y_x , тоді використовується лінійна функція (наприклад, залежність між урожайністю культур і кількістю внесених добрив).

Найпростішим рівнянням парної регресії, що описує лінійну залежність між факторною і результативною ознаками, є рівняння прямої, яке має такий вигляд:

$$Y_x = a_0 + a_1 x$$

де Y_x – залежна змінна, яка оцінюється або прогнозується (результативна ознака);

a_0 – вільний член рівняння;

a_1 – коефіцієнт регресії;

x – незалежна змінна (факторна ознака), яка використовується для визначення залежної змінної.

Параметри рівняння обчислюються на основі системи нормальних рівнянь методом найменших квадратів:

$$\begin{cases} \sum y = na_0 + a_1 \sum x \\ \sum xy = a_0 \sum x + a_1 \sum x^2 \end{cases}$$

Звідси

$$a_1 = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - \sum x \sum x}$$

$$a_0 = \frac{\sum x^2 \sum y - \sum x \sum xy}{n \sum x^2 - \sum x \sum x}$$

або

$$a_0 = \bar{y} - a_1 \bar{x}$$

Основне змістовне навантаження в рівнянні регресії несе коефіцієнт регресії. Найчастіше застосовуються лінійні рівняння або приведені до лінійного вигляду. Коефіцієнт регресії – це кутовий коефіцієнт у прямолінійному рівнянні кореляційного зв'язку. У лінійній функції рівняння регресії він показує на скільки одиниць в середньому зміниться результативна ознака y при зміні факторної ознаки x на одиницю свого натурального виміру. Тобто, коефіцієнт регресії – це варіація y , яка припадає на одиницю варіації x . Коефіцієнт регресії має одиницю виміру результативної ознаки. За наявності прямого зв'язку коефіцієнт регресії є додатною величиною, а за зворотного зв'язку – від'ємною.

Параметр a_0 – вільний член рівняння регресії, тобто це значення y при $x = 0$. Цей показник має тільки розрахункове значення у випадках, коли x не має нульових значень. У разі, коли зі зміною факторної ознаки результативна змінюється нерівномірно, використовуються нелінійні функції. Так,

якщо зміна факторного показника сприяла прискореній динаміці результативного показника (наприклад, вплив обсягу грошової маси на рівень інфляції), доцільно використати степеневу функцію:

$$Y_x = ax^b$$

У випадку, коли під впливом факторної ознаки результативна змінюється нерівномірно, причому з уповільненням, використовується рівняння гіперболи:

$$Y_x = a + \frac{b}{x}$$

Прикладом такої залежності є залежність рівня продуктивності праці робітників від рівня їх заробітної плати.

Якщо зміна факторної ознаки супроводжується нерівномірною варіацією факторної ознаки із зміною напрямку зв'язку, нелінійна регресія описується рівнянням параболи:

$$Y_x = a + bx + cx^2$$

Так, за допомогою функції параболи можна виразити залежність урожайності культур від кількості опадів. Безумовно, що при застосуванні регресійного аналізу дотримуються деяких умовностей. Так, попередньо обумовлюється, що дія інших факторів, крім взятого за факторну ознаку, залишиться незмінною, а в дослідженні взято тільки один фактор. Проте останню умовність можна усунути за допомогою застосування множинної регресії та кореляції, за яких підбирається значна кількість факторів. Аналіз на основі множинної регресії (анг. multiple regression analysis) – вид регресійного аналізу, який ґрунтується на використанні в рівнянні регресії більше, ніж однієї незалежної змінної.

Так, його застосовують при прогнозуванні попиту. Причому спочатку ідентифікуються фактори, що визначають попит, потім встановлюються наявні між ними взаємозв'язки та прогнозується ймовірні майбутні їх значення. На основі отриманих даних виводиться прогнозне значення попиту.

Багатофакторне рівняння множинної регресії при лінійній залежності має такий вигляд:

$$Y_{x_1-x_n} = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n,$$

де a_0 – вільний член рівняння;

a_1, a_2, \dots, a_n – коефіцієнти регресії;

x_1, x_2, \dots, x_n – незалежні змінні (факторні ознаки);

n – кількість незалежних змінних.

Визначення параметрів множинної регресії вимагає трудомістких розрахунків із застосуванням комп'ютерних інформаційних систем. Однак одержані результати будуть достовірними і можуть широко використовуватися в економічній та управлінській діяльності насамперед для складання довгострокових прогнозів. Відомо, що однофакторна модель придатна для короткострокових прогнозів (на 2–3 роки).

Недоліки регресійного аналізу [15]: моделі, що мають занадто малу складність, можуть виявитися неточними, а моделі, що мають надлишкову складність, можуть виявитися перенавчання. Крім того даний метод є статистичним, отже таким, що не враховує аргумент часу і не годиться для випадку прогнозування цін постачання електроенергії.

Використання дистрибутивно-лагових моделей для завдань прогнозування

Дистрибутивно - лагові моделі (ДЛМ) - це регресійні моделі з присутнім тимчасовим лагом, тобто вихідна характеристика залежить також і від значень вхідного фактора в попередні періоди.

Загальний вигляд нескінченної ДЛМ:

$$y_t = const + \beta_0x_t + \beta_1x_{t-1} + \beta_2x_{t-2} + \dots + \varepsilon_t$$

ДЛМ з кінцевим лагом у k періодів:

$$y_t = \alpha + \beta_0x_t + \beta_1x_{t-1} + \beta_2x_{t-2} + \dots + \beta_kx_{t-k} + \varepsilon_t,$$

де t - поточний період,

$x_{t-i}, i = 1, 2, 3, \dots$ - сумарне значення фактора в період $t - i$, $\beta_i, i = 1, 2, 3, \dots$ - коефіцієнти впливу i -

го часового лага, ε_t - похибка, β_0 - короткостроковий мультиплікатор, $\sum_i \beta_i = \beta$ - довгостроковий

мультиплікатор. Тоді $\beta_i^* = \frac{\beta_i}{\beta}$ - показує частину загального впливу лага на i -ому проміжку часу.

Враховуючи важливість використання ДЛМ в економіці та в інших галузях життєдіяльності людини, є затребуваною адекватна оцінка параметрів такої моделі. Маємо завдання оцінки параметрів i , як найбільш значимих.

1. Послідовна оцінка.

Припустимо, що x_i - не стохастичні, тоді маємо, що $x_{t-i}, i = 1, 2, 3, \dots$ також не стохастичні, тому можна застосувати метод найменших квадратів (МНК). На цій підставі Ф. Альт і Дж. Тінберген запропонували алгоритм послідовної оцінки параметрів за допомогою МНК: будемо регресію вигляду $y_t(x_t)$ і оцінюємо параметри, далі $y_t(x_t, x_{t-1})$ - і оцінюємо параметри і так далі, поки коефіцієнти стануть незначущими або змінять знак. Беручи до уваги проблеми мультиколінеарності і невизначеності максимальної довжини лага даний метод не рекомендується для використання на практиці.

2. Підхід Койка. Припустимо, що коефіцієнти для моделі з невизначеним лагом мають однаковий знак і змінюються у геометричній прогресії, тоді $\beta_k = \beta_0 \lambda^k$, причому $0 < \lambda < 1$ - темп зменшення, і модель стає кінцевою, тобто:

$$\sum_{k=0}^{\infty} \beta_k = \beta_0 \frac{1}{1-\lambda}$$

Вплив лага на з часом зменшується і модель можна записати, як

$$y_t = \alpha + \beta_0 x_t + \beta_0 \lambda x_{t-1} + \beta_0 \lambda^2 x_{t-2} + \dots + \varepsilon_t \quad (1)$$

Далі в модель вводиться затримка на один період і множиться на λ , тоді маємо

$$\lambda y_{t-1} = \lambda \alpha + \beta_0 \lambda x_{t-1} + \beta_0 \lambda^2 x_{t-2} + \beta_0 \lambda^3 x_{t-3} + \dots + \lambda \varepsilon_{t-1} \quad (2)$$

Віднімаючи 2 від 1 отримаємо

$$y_t - \lambda y_{t-1} = \alpha(1-\lambda) + \beta_0 x_t + \varepsilon_t - \lambda \varepsilon_{t-1}$$

або

$$y_t = \alpha(1-\lambda) + \beta_0 x_t + \lambda x_{t-1} + v_t \quad (3)$$

де $v_t = \varepsilon_t - \lambda \varepsilon_{t-1}$

Залежність 3 отримала назву перетворення Койка. Відзначимо особливості перетворення Койка:

1) заміна нескінченної кількості параметрів на всього три: α, β_0, λ , які необхідно оцінити, і відповідно зникла проблема мультиколінеарності;

2) модель з дистрибутивно-лагової перетворилася на авторегресійну, оскільки залишається незалежна змінна;

3) введення ще однієї стохастичною змінної породжує проблему серійної кореляції відносно y_{t-1} . Для використання МНК всі незалежні змінні повинні бути нестохастичні, або не залежати від похибки v_t , а v_t - серійно кореляційні. Для визначення серійної кореляції використовується h -тест Дарбіна.

На практиці для визначення швидкості з якою y реагує на зміну x застосовують медіанний та середній лаги.

1) Медіанний лаг - це час, необхідний для 50%-ої зміни y при одиничній зміні x . Для моделі Койка медіанний лаг дорівнює:

$$\frac{\log 2}{\log \lambda}$$

2) Середній лаг \bar{k} , по суті, зваженим середнім всіх лагів моделі і розраховується за формулою:

$$\frac{\sum_{k=0}^{\infty} k \beta_k}{\sum_{k=0}^{\infty} \beta_k}$$

Фактично, модель Койка є послідовною моделлю, яка не має чіткого теоретичного обґрунтування.

Дані методи виявились корисними в емпіричній економіці, оскільки вони перетворюють моделі, які б залишились статистичними, на динамічні, за допомогою фактору часу, але значним недоліком, який не дозволяє їх використовувати у прогнозуванні цін постачання електроенергії є той, що дослідник повинен перед початком дослідження задавати як ступінь поліному, так і максимальну довжину лагу.

МГУА

МГУА (Метод Групового Урахування Аргументів) індуктивний інструмент машинного навчання для автоматичного синтезування моделей згідно бази даних входів та відповідних їм виходів, що представляють навчальний набір прикладів. Інструмент може автоматично синтезувати адекватні моделі, що реалізують природну структуру складних та високо нелінійних систем. Автоматизація синтезування моделі не тільки зменшує складність аналізу, але й запобігає впливу невірних людських суджень та невиправданому відданню переваги тим, чи іншим моделям при генеруванні. Підхід МГУА це

формалізована парадигма багатозадової поліноміальної регресії, що здатний продукувати поліноміальні моделі високих ступенів в якості ефективних предикторів. Процес по суті є «еволюційним», використовуючи на початку прості регресійні відношення, для отримання більш точних представлень на наступній ітерації. Для того, щоб запобігти експоненційному росту та обмежити складність моделей, алгоритм обирає тільки відношення, що мають хорошу здатність до коректних передбачень на кожному кроці. Процес ітерації зупиняється коли наступне покоління регресійних рівнянь починає набувати гіршу здатність до передбачення порівнюючи з попередниками. У цей момент модель починає ставати надто спеціалізованою і тому, навряд чи покаже себе добре з новими даними. Алгоритм містить три основні елементи: представлення, вибір та зупинку. Він застосовує евристики для здійснення рішень стосовно одного або всіх цих трьох аспектів.

Для того, щоб проілюструвати ці три кроки, візьмемо оціночний набір даних спостережень (рядків) та $m+1$ стовпців для m незалежних змінних (x_1, x_2, \dots, x_m) та одну залежну змінну y . На першій ітерації ми припускаємо, що наші предиктори це, власне, вхідні змінні. Початкові рівняння отримуються взяттям кожної пари вхідних змінних $(x_i, x_j; i, j = 1, \dots, m)$ разом з виходом y для обрахування квадратичного регресійного поліному [12]:

$$y = A + Bx_i + Cx_j + Dx_i^2 + Ex_j^2 + Fx_ix_j \quad (4)$$

Кожний з результуючих $m(m-1)$ поліномів оцінюється використовуючи дані для пари змінних x , таким чином продукуючи нові змінні $(z_1, z_2, \dots, z_{\frac{m-1}{2}})$, які мають описувати y краще, ніж початкові змінні. Результуючі змінні z відбираються відповідно до певного критерію селекції і тільки ті, що будуть мати найкращу здатність до передбачення залишаються. МГУА застосовує додаткові та незалежні набори n_s спостережень для цих цілей та використовує критерій регулярності, що базований на корені середньоквадратичної похибки r_k для набору даних, де:

$$r_k^2 = \frac{\sum_{l=1}^{n_s} (y_l - z_{kl})^2}{\sum_{l=1}^{n_s} y_l^2}; k = 1, 2, \dots, \frac{m-1}{2} \quad (5)$$

Тільки ті поліноми (і зв'язані з ними змінні z), що мають r_k менше визначеного ліміта залишаються, також зберігається мінімальне значення, r_{\min} , отримане для r_k . Обрані змінні z представляють нову базу даних, що використовується для повторних кроків оцінки та вибору на наступній ітерації, для того, щоб утворити змінні вищого рівня. На кожній ітерації r_{\min} порівнюється з його попереднім значенням і процес продовжується до тих пір, доки r_{\min} зменшується або поки не буде досягнуто заданий рівень складності моделі. Якщо r_{\min} збільшується, то це вказує на надто велике ускладнення моделі, таким чином модель стає перенавченою і буде погано спрацьовувати на нових даних. Перевірка складності моделі також є важливим аспектом МГУА, слідкуючи за тим, щоб була досягнена кінцева ціль створення моделі, для використання з новими даними, які не використовувались для навчання. Найкращою моделлю буде та, що надає найкоротший опис для доступних даних [13]. Чисельно, кінцева МГУА модель виглядає як багаторівнева мережа часткових квадратичних поліномів, кожен рівень якої представляє результат ітерації.

Згідно принципу зовнішнього доповнення модель, що синтезується за допомогою МГУА повинна бути заснована на використанні зовнішніх критеріїв. У цьому випадку мінімум квадратичної похибки на даних, що використовувались для навчання не годиться. Таким чином потрібно запропонувати критерій, які б давали оптимальну складність обраної моделі. Першим таким критерієм може служити, так званий, «критерій регулярності» і є середньоквадратичною помилкою, що знаходиться на перевірочній послідовності. Також можна оптимізувати процес розділення даних на навчальну та перевірочну ввівши нове зовнішнє доповнення, яке є помилкою на екзаменаційних даних.

В даному завданні запропоновано зовнішнє доповнення засноване на екзаменаційних даних, що представлені значеннями вартості електроенергії зсунутими в часі на певну величину.

Висновки

В роботі було виконано аналіз існуючих методів прогнозування, в результаті якого виявлено, що жоден з розглянутих методів не є оптимальним для завдань прогнозування цін на електроенергію оптового ринку електроенергії України. Оскільки у кожного з вказаних методів присутні окремі недоліки, що не дозволяють використовувати їх в даній задачі (ГА має тенденцію сходиться до локального мінімуму, структуру ШНМ необхідно задавати із самого початку, класичний регресійний аналіз не враховує змінну часу, недолік дистрибутивно-лагової моделі схожий на недоліки ШНМ і полягає в тому, що перед початком дослідження потрібно задавати максимальну довжину лагу), потрібно обрати один з методів, що позбавлений цих недостатків. Одним з таких методів є МГУА. Він дозволяє отримувати структуру об'єкта із даних, на яких навчається, тобто можна обійтись без апіорних даних, він враховує змінну часу і дозволяє робити як короткострокові прогнози, так і довгострокові, в залежності від обраного критерію селекції моделей.

Таким чином, доцільно використати МГУА для прогнозу цін постачання електроенергії на ОРЕ.

Література

1. Holland, J.H. Adaptation in natural and artificial systems [Text] / J.H. Holland. — Ann Arbor : Univ. of Michigan Press, 1975. — 183 p.
2. Генетические алгоритмы, искусственные нейронные сети и проблемы виртуальной реальности [Текст] / Г.К. Вороновский, К.В. Махотило, С.Н. Петрашев, С.А. Сергеев. — Х. : Основа, 1997. — 112 с.
3. Батищев, Д.И. Оптимизация многоэкстремальных функций с помощью генетических алгоритмов [Текст] / Д.И. Батищев, С.А. Исаев // Сборник статей ВГТУ. — 1997.
4. Борукаев З.Х., Остапченко К.Б. Лисовиченко О.И. Компьютерное моделирование задач планирования поставок электроэнергии на оптовом рынке с применением искусственных нейронных сетей [Текст] / Остапченко К.Б. Лисовиченко О.И. Борукаев, З.Х. // Збірник наукових праць, Інститут проблем моделювання в енергетиці ім. Г.С.Пухова. — 2012. — No 65. — С. 140–152.
5. A.P. Alves da Silva U.P. Rodrigues, A.J. Rocha Reis L.S. Moulin. NeuroDem - a neural network based short term demand forecaster [Text] / A.J. Rocha Reis L.S. Moulin A.P. Alves da Silva, U.P. Rodrigues // IEEE Power Technical Conference. — 2001.
6. H.S. Hippert C.E. Pedreira, R.C. Souza. Combining neural networks and ARIMA models for hourly temperature forecasts [Text] / R.C. Souza H.S. Hippert, C.E. Pedreira // Proceedings of the IEEE International Joint Conference on Neural Networks. — 2000. — P. 414–419.
7. W. Charytoniuk, M.S. Chen. Neural network design for short-term load forecasting. [Text] / M.S. Chen W. Charytoniuk.
8. I Tassadduq, S. Rehman K. Bubshait. Application of neural networks for the prediction of hourly mean surface temperatures in Saudi Arabia [Text] / S. Rehman K. Bubshait I, Tassadduq // Renewable Energy. — 2002. — Vol. 25. — P. 545–554.
9. T. Matsui T. Iizaka, Y. Fukuyama. Peak load forecasting using analyzable structured neural network [Text] / Y. Fukuyama T. Matsui, T. Iizaka // Proceedings of the IEEE Power Engineering Society Winter Meeting. — 2001. — P. 405–410.
10. III, H.W. Lewis. Intelligent hybrid load forecasting system for an electric power company [Text] / H.W. Lewis III // Proceedings of the IEEE Mountain Workshop on Soft Computing in Industrial Applications. — 2001. — P. 23–27.
11. Ивахненко, А. Г. Индуктивный метод самоорганизации моделей сложных систем [Текст] / А. Г. Ивахненко. — Киев : Наук. думка, 1981. — 296 с.
12. Farlow, S.J. The GMDH algorithm, In: Farlow SJ, ed. Self-Organizing Methods in Modeling: GMDH Type Algorithms [Text] / S.J. Farlow. — New York : Marcel-Dekker, 1984. — P. 1–24.
13. Barron, A.R. Predicted squared error- a criterion for automatic model selection. In: S.J. Farlow, ed. Self-Organizing Methods in Modeling: GMDH Type Algorithms [Text] / A.R. Barron. — New York : Marcel-Dekker, 1984. — P. 87–103.
14. Купалова, Г.І. Теорія економічного аналізу: навч. посіб [Текст] / Г.І. Купалова. — К. : Знання, 2008. — 639 p.
15. Дрейпер, Н. Прикладной регрессионный анализ [Текст] / Н. Дрейпер, Г. Смит. — М. : Издательский дом «Вильямс», 2007.

V.V. HUZ, postgraduate student

Operational forecasting of prices for electric power supply in the wholesale market

In this paper we analyzed the available methods for the process of operational forecasting prices in the wholesale electricity market of Ukraine and suggested to use of Group Method of Data Handling (GMDH). GMDH allows you to find the optimum in terms of complexity, structure and accuracy of the prediction model that describes the dynamics of change prices and takes into account stationary ergodic input. Moreover, was proposed to use external additions, based on the inputs for creating such models.

Keywords: Group Method of Data Handling, planning of supplies, wholesale market, electricity, computer model, prediction

В.В. ГУЗ, аспірант кафедри

Оперативное прогнозирование цен поставок электроэнергии на оптовом рынке

В данной работе были проанализированы доступные методы для процесса оперативного прогнозирования цен на оптовом рынке электроэнергии Украины и предложено использование Метода Группового Учета Аргументов (МГУА). МГУА позволяет находить оптимальную, с точки зрения сложности, структуры и точности предсказания, модель, описывающая динамику изменения цен и позволяет учитывать нестационарность входных данных. Более того, предложено использование внешних дополнений, основанных на входных данных при образовании таких моделей.

Ключевые слова: Метод Группового Учета Аргументов, планирование поставок, оптовый рынок, электроэнергия, компьютерная модель, прогнозирование

УДК 621.31

ПАВЛОВА Ю.С., аспірант,

КОЦАР О.В., канд.техн.наук, доцент

Національний технічний університет України «Київський політехнічний інститут»

КЕРУВАННЯ ЕНЕРГОЕФЕКТИВНІСТЮ В ЛІБЕРАЛІЗОВАНИХ РИНКАХ ЕЛЕКТРИЧНОЇ ЕНЕРГІЇ

В роботі обґрунтована доцільність визначення механізмів та інструментів ринків електричної енергії, які б дозволяли вирішувати не лише головні завдання енергоринку України – забезпечення справедливого доступу споживачів до енергоресурсів на конкурентних засадах за справедливою ціною, – а й стимулювали б учасників ринку до їхнього ефективного використання. В роботі представлений аналіз результатів функціонування світових лібералізованих ринків електричної енергії, досліджено способи та результати врахування показників енергоефективності під час розрахунків за електроенергію в конкурентних ринках. За результатами аналізу зроблено висновок щодо необхідності дослідження шляхів та визначення методів залучення енергоефективності до асортименту товарів в ринку двосторонніх договорів та балансуєчому ринку, який сьогодні запроваджується в Україні. Визначено напрямки подальших досліджень та завдання, які необхідно вирішити.

Ключові слова: енергоефективність, ринок електроенергії, лібералізовані ринки, керування.

Вступ

На сьогоднішній день в Україні проявляється негативна світова тенденція щодо неухильного збільшення шкідливого техногенного впливу на навколишнє середовище, зокрема під час виробництва, розподілу та використання енергоресурсів. Дані державного комітету статистики України за 2010-2012 роки свідчать про те, що серед викидів забруднюючих речовин в атмосферне повітря від стаціонарних джерел забруднення близько 52-54 % на рік приходить на енергетику в цілому, що досягає 2300 тис. тон на рік [1, 2]. Такі кількісні значення зумовлені, переважно, електроенергетичним сектором, оскільки лише він є джерелом майже 39,5 % загального обсягу викидів забруднюючих речовин по Україні щорічно (близько 1700 тис. т.) [1, 2]. Крім того, викиди діоксиду вуглецю (CO₂) під час виробництва та розподілу електроенергії, газу та води становлять близько 47 % від загальних викидів CO₂ на рік [3, 4]. До того ж негативний вплив доповнюється вичерпністю запасів використовуваного на виробіток енергії викопного палива (природний газ, вугілля, нафта тощо). Оскільки більшість способів виробітку