

І.В. Блінов¹, д-р. техн. наук, проф. ORCID 0000-0001-8010-530П.В. Шиманюк¹, д-р. філософії ORCID 0000-0002-7585-7493В.В. Сичова¹, д-р. філософії ORCID 0000-0001-7385-1680В.О. Мірошник¹, к.т.н. ORCID 0000-0001-9036-7268¹Інститут електродинаміки НАН України

ВПЛИВ АНОМАЛЬНИХ ЗНАЧЕНЬ НА ТОЧНІСТЬ ПРОГНОЗУВАННЯ ВТРАТ В РОЗПОДІЛЬЧИХ МЕРЕЖАХ

Актуальність дослідження обумовлена сучасними трендами в управлінні режимами роботи розподільних електричних мереж із використанням технологій Smart Grid, а також необхідністю зниження витрат операторів систем розподілу на закупівлю електроенергії. Для цього потрібні точні результати прогнозування навантажень у вузлах мережі на різних горизонтах прогнозування. Різкі зміни топології мережі можуть збільшувати похибки прогнозу витрат як єдиного часового ряду, що негативно впливає на ефективність керування мережею та підвищує витрати на закупівлю електроенергії для покриття витрат.

У роботі було розглянуто методи прогнозування на основі штучних нейронних мереж для розрахунку та прогнозування витрат електричної енергії, а також проведено порівняння цих методів між собою. Розрахунки виконано на основі даних одного з українських операторів систем розподілу, а тестова електрична мережа адаптована на основі схеми CIGRE для моделювання витрат електричної енергії. Оскільки дані вузлового навантаження містили пропуски та аномалії, було використано двоетапний алгоритм аналізу даних із застосуванням методу кластеризації DBSCAN для їх виявлення та корекції. В результаті проведених обчислень витрат на основі достовірних даних, похибка витрат була зменшена втричі порівняно з розрахунками, що базуються на коефіцієнтах навантаження. Використання методів аналізу даних та прогнозування на основі штучних нейронних мереж значно підвищує точність розрахунків витрат і мінімізує похибки.

Ключові слова: витрати, вузлове навантаження, прогнозування, нейронні мережі, CIGRE, LSTM.

Вступ.

На ринку електроенергії України набуває все більшої актуальності задача прогнозування витрат електричної енергії в розподільних мережах. Точність прогнозування витрат електроенергії має безпосередній вплив на вартість електроенергії для споживачів. А саме, на формування кінцевих тарифів впливають витрати на придбання електроенергії операторами системи розподілу (ОСР) та операторами системи передачі (ОСП) для компенсації витрат у власних мережах [1]. Цей вплив здійснюється через складові тарифів ОСР та ОСП [2–4]. Крім того, неточне прогнозування може спричинити збільшення значень небалансів електроенергії. Це створює потребу в закупівлі додаткового обсягу електроенергії на балансуєчому ринку [5]. Внаслідок чого утворюється додатковий фактор зростання вартості електроенергії для ОСР і споживачів.

Здебільшого задля визначення витрат у мережах використовують узагальнені коефіцієнти витрат, які розраховані на основі сезонних пікових навантажень. Нестабільність топології мереж також ускладнює прогнозування витрат, що знижує ефективність управління енергосистемами [6].

Наявність точних прогнозів навантаження у вузлах мережі для різних періодів дозволяє не лише зменшити витрати на придбання електроенергії, а й оптимізувати управління режимами мереж, що особливо важливо в умовах впровадження технологій Smart Grid в Україні.

Розвиток методів машинного навчання та нейронних мереж для прогнозування витрат в електричних мережах є важливим кроком у вдосконаленні управління сучасними енергосистемами [7]. Традиційні методи розрахунку витрат, засновані на фізичних моделях, мають обмежену точність через складність врахування багатьох змінних, таких як динаміка навантаження, погодні умови, структурні особливості мережі та її стан. Використання машинного навчання дозволяє подолати ці обмеження завдяки здатності алгоритмів виявляти складні нелінійні залежності між параметрами, що важко формалізувати вручну. Нейронні мережі особливо ефективні завдяки їх гнучкості та здатності до навчання на великих масивах даних, що дає можливість прогнозувати витрати з високою точністю навіть у ситуаціях, коли початкові дані містять значну кількість невизначеності.

Наприклад, у роботі [8] було розроблено комбіновану нейронну мережу, в якій поєднано багатоварний перцептрон із алгоритмом авторегресії. Модель використовувала алгоритм авторегресії для попередньої обробки даних, а її ефективність перевірено на даних енергосистеми PJM у США за 2014–2015 роки. Інша комбінована модель, описана в [9], поєднує кілька нейронних мереж із модулем нечіткої

логіки (модуль PROTREN) для виявлення трендів на основі даних енергосистеми острова Крит, враховуючи інформацію про температуру повітря.

Метод Support Vector Machine (SVM) також широко застосовується для прогнозування вузлових навантажень. Як описано в [10], цей підхід використовується для управління енергосистемою провінції Шаньдун у Китаї. Модель враховує співвідношення між активною потужністю вузлів і системи загалом, а також взаємозв'язок активної та реактивної потужностей. Для оцінки ефективності SVM порівнювали з нелінійною авторегресійною нейронною мережею [11] та адаптивним фільтром Калмана, який використовується для прогнозування коефіцієнтів потужності вузлів.

Для зниження вартості втрат запропоновано використання сучасних методів розрахунку та прогнозування втрат електроенергії на основі методів штучної нейронної мережі глибокого навчання. Для перевірки ефективності прогнозування втрат електроенергії використовувалися різні архітектури штучних нейронних мереж глибокого навчання.

Виявлення та заміна аномальних значень у даних є надзвичайно важливими для забезпечення точності та стабільності прогнозування втрат в електричних мережах за допомогою штучних нейронних мереж. Аномальні значення можуть виникати через апаратні збої, помилки в записах даних або вплив зовнішніх факторів, таких як погодні аномалії чи надзвичайні стани в мережі. Якщо такі дані не обробляти, вони здатні значно вплинути на результати моделі, спотворюючи її прогнози та знижуючи здатність до узагальнення. Це особливо критично, адже нейронні мережі чутливі до якості вхідної інформації: навіть невелика кількість некоректних значень може викликати значне викривлення ваг моделі, змушуючи її адаптуватися до помилкових залежностей, які не відображають реальних фізичних процесів у мережі. При прогнозуванні часових рядів це може відбуватись через вплив аномальних значень на цільову функцію при навчанні нейронних мереж, наприклад, середньоквадратичну похибку, яка є чутливою до великих відхилень.

Обробка аномальних значень також сприяє покращенню стабільності роботи нейронної мережі. У випадку складних енергосистем, де багато параметрів взаємозалежні, некоректні дані можуть створювати нелогічні зв'язки між змінними, що суперечать фізиці процесу. Це не лише погіршує прогноз, але й може призводити до некоректних висновків про стан мережі. Заміна або видалення аномалій допомагає нейронній мережі навчатися на коректних даних і створювати моделі, які відображають реальну поведінку енергосистем.

Таким чином, вчасне виявлення та заміна аномальних значень є ключовим етапом у процесі підготовки даних для прогнозування втрат. Це забезпечує не лише більш точні та надійні результати, а й підвищує довіру до прогнозу, адже модель базується на даних, які відповідають реальним умовам і фізичним закономірностям.

Для оцінки впливу аномальних значень на точність прогнозування втрат було розроблено два підходи до прогнозування втрат:

- Прогнозування навантаження кожного вузла окремо з подальшим розрахунком втрат на основі прогнозованих даних.
- Розрахунок втрат електричної енергії з наступним її прогнозом.

Для прогнозування всіх вузлів навантаження використовувалась рекурентна нейронна мережа типу LSTM, яка описана в [12]. Для індивідуального прогнозування вузлів навантаження та втрат електроенергії використовувалася мережа eResNet, розроблена в Інституті електродинаміки. На рисунку 1 показано блок-схеми алгоритмів прогнозування та розрахунку втрат електроенергії.

Обидва методи прогнозування були адаптовані до ретроспективних даних «Вінницяобленерго». Навчання обох алгоритмів проводиться за допомогою алгоритму оптимізації ADAM.

Розрахунок втрат та прогнозування електричної енергії виконувалось на ретроспективних даних «Вінницяобленерго» які включають в себе 16 вузлів навантаження за період з 2017 по 2019 роки з 30-хвилинною дискретністю. Загальний обсяг даних становить 48 048 значень.

Для моделювання електричної мережі та розрахунку втрат була використана мережа середнього рівня напруги CIGRE [13] як основа для тестової електричної мережі.

Мова програмування Python та бібліотека для аналізу даних Pandarower були використані для побудови та аналізу тестової системи живлення. Бібліотека Pandarower являє собою окремий набір інструментів для розробки та аналізу електричних мереж. Бібліотека охоплює широкий спектр різноманітних моделей електричних мереж, включаючи численні тестові системи та приклади енергосистем CIGRE [14].

Тестова електрична мережа складається з двох трансформаторів потужністю 40 МВА 110/20 кВ, 15 вузлів, 18 джерел навантаження, з вилученими двома джерелами для узгодження з ретроспективними даними, 14 ліній, у тому числі 12 кабельних і дві повітряні, вимикач. Усі елементи, які використовуються в цій мережі, є елементами бібліотеки Pandarower.

Дана тестова мережа була розроблена для моделювання режимів роботи електричної мережі та визначення втрат електричної енергії, дана мережа представлена в статті [15]. Оскільки ця мережа

побудована на основі реальної європейської мережі, вона суттєво відрізняється за складом навантаження та параметрами від української мережі. Тому для забезпечення належного функціонування алгоритму розрахунку збитків схему розрахунку було адаптовано з використанням наявних ретроспективних даних.

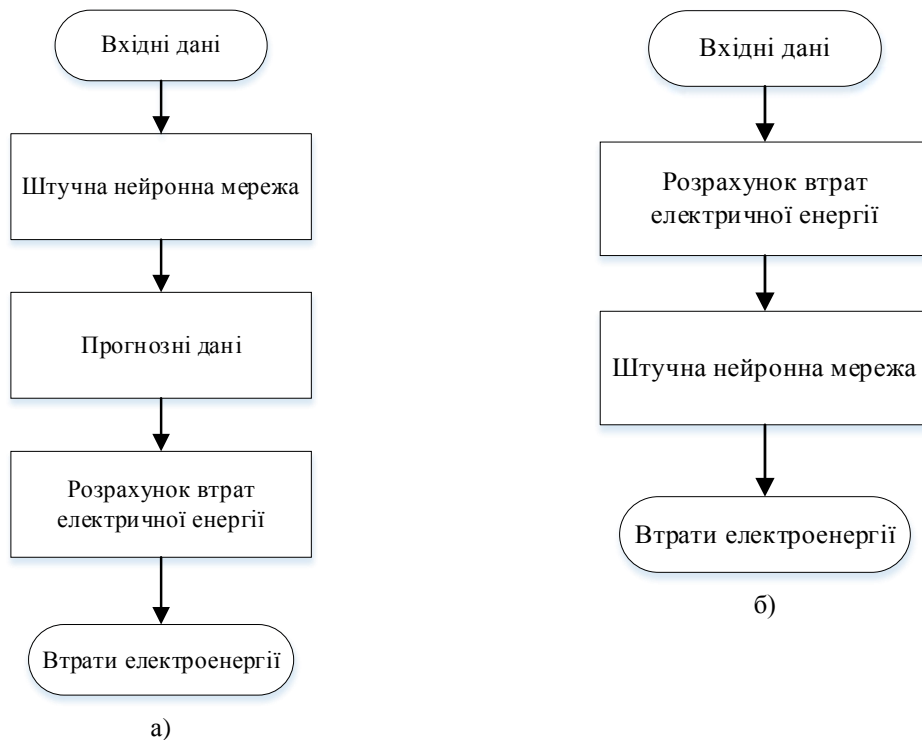


Рисунок 1 - а) Розрахунку втрат за прогнозними значеннями; б) Розрахунку та прогнозування втрат

Враховуючи різницю у величині навантаження між енергосистемою CIGRE та даними ОСП, кабельні та повітряні лінії були замінені на лінії більшого перерізу для належного функціонування тестової енергосистеми. Крім того, дані про навантаження вузлів від ОСП розподілялися відповідно до величин навантаження вузлів енергосистеми CIGRE.

Після проведення попереднього аналізу даних було виявлено, що дані містять пропуски значень та аномальні значення, які значно відрізняються від нормальних значень.

Для виявлення та заміни відсутніх значень і аномальних даних був застосований двоетапний алгоритм перевірки даних з використанням методу кластеризації DBSCAN [16].

Запропонований алгоритм достовірності для окремого вузла складається з наступних етапів:

1. Вибір часових зрізів із безперервного часового ряду навантаження $R^{n \times 1} \rightarrow R^{n/24 \times 24}$.

2. Виявлення грубих аномальних значень у часових зрізах за допомогою методу кластеризації DBSCAN.

Значення, які не належать до першого кластера, вважаються аномальними.

3. Заміна аномальних значень за допомогою лінійної інтерполяції.

4. Розгортка часових зрізів у безперервний ряд навантажень $R^{n/24 \times 24} \rightarrow R^{n \times 1}$.

5. Декомпозиція часового ряду на трендову, сезонну та залишкову складові.

6. Виявлення аномальних значень (п. 2) у часовому ряді залишкової складової.

7. Заміна виявлених значень за допомогою лінійної інтерполяції.

Ретроспективні дані вузлового навантаження на були розділені на вибірки для навчання та тестування, для даних з аномальними значеннями та даних без аномальних значень (після перевірки). Навчальні набори охоплюють усі значення навантаження вузла, за винятком 336 значень, що представляють один тиждень (7 днів), які використовувалися як тестовий набір.

Результати прогнозування втрат за день та тиждень наведені на рисунках 2-3.

Для оцінки точності представлених підходів була використана функція середньої абсолютної відсоткової помилки у відсотках (MAPE). У таблиці 1 наведені похибки розрахунку втрат електричної енергії.

Провівши прогнозування та розрахунок втрат з використанням усіх підходів, ми спостерігаємо, що, як правило, прогнозування навантаження з подальшими розрахунками втрат мають меншу похибку втрат. Також варто зазначити, що використання методів аналізу даних для виявлення та заміни аномальних значень дозволяє покращити результати прогнозування.

Отримані результати дослідження є складовою роботи з підготовки даних для моделі Smart Grid для оперативного управління розподільними мережами на основі методів штучного інтелекту. Ця модель розробляється авторами публікації в рамках спільного українсько-литовського проекту «Спільні українсько-литовські науково-дослідні проекти у 2024 – 2025 роках».



Рисунок 2 - Півгодинні значення втрат за добу з використанням різних методів прогнозування

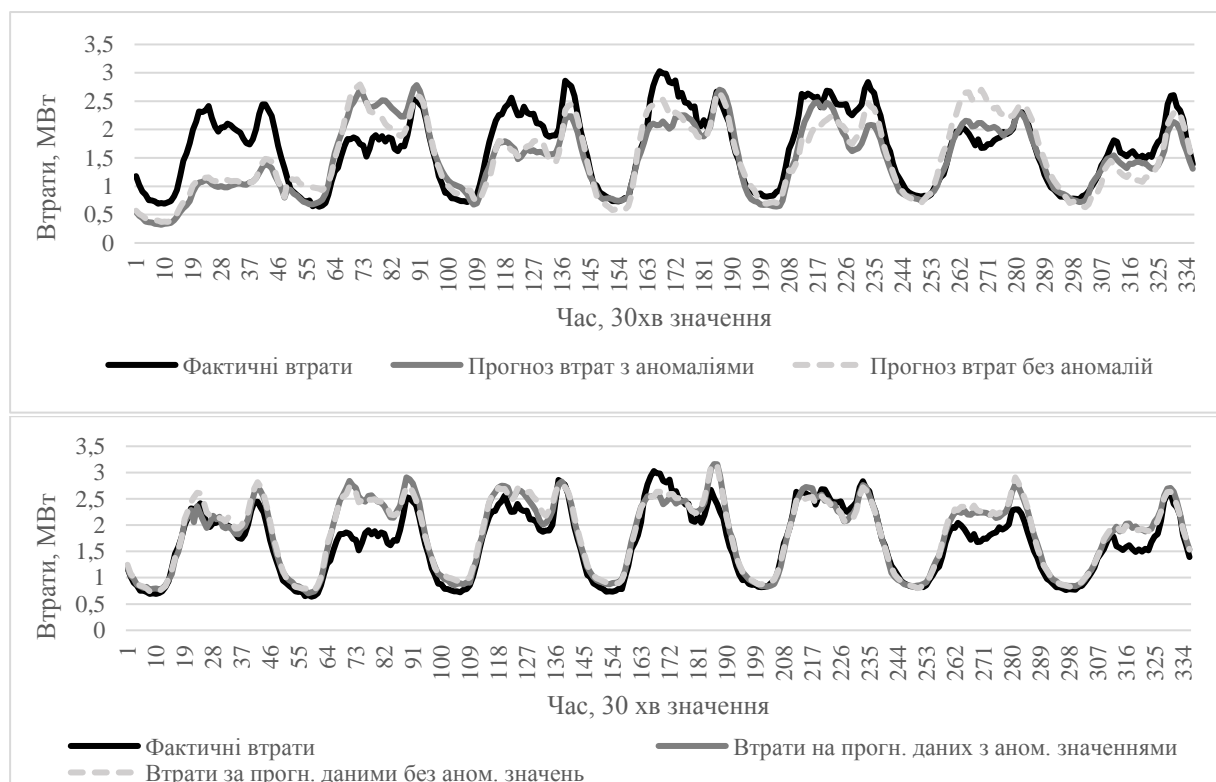


Рисунок 3 - Півгодинні значення втрат за місяць з використанням різних методів прогнозування

Метод прогнозування втрат	Величина похибки MAPE, %
Прогноз втрат з аномаліями	20.08
Прогноз втрат без аномалій	20.24
Втрати за прогнозними даними з аномальними значеннями	12.78
Втрати за прогнозними даними без аномальних значень	3.29

Висновки.

Розглянута методика короткострокового прогнозування величин сумарних втрат електричної енергії в мережі на основі ретроспективних даних з використанням штучних нейронних мереж глибокого навчання дозволила підвищити точність визначення втрат без виконання їх розрахунку порівняно з існуючими методами.

Крім того, запропоновано методи короткострокового прогнозування вузлових навантажень на основі штучних нейронних мереж глибокого навчання. Це дозволило підвищити точність визначення розрахункових значень втрат в мережі за рахунок їх врахування прогнозованих значень навантажень вузлового навантаження під час розрахунку.

Результати прогнозування втрат електроенергії показують, що застосування методів аналізу даних і методів прогнозування дозволяє знизити величину втрат порівняно з прогнозними даними фактичних втрат з наявними аномальними значеннями.

Проведення додаткового аналізу даних для виявлення та заміни аномальних значень підвищило точність прогнозів як для вузлів навантаження, так і для втрат електричної енергії.

Похибка прогнозування втрат з використанням штучних нейронних мереж дозволяє отримати похибку MAPE в межах 3,29% для прогнозованих достовіризованих даних вузлів навантаження, які використовуються для розрахунку втрат. Тоді як прогнозування втрат достовіризованих даних мають похибку 20,24%. Похибка втрат для неперевіраних даних становить 20,08%.

Використання методів прогнозування найбільш ефективно продемонстровано при прогнозуванні достовіризованих даних вузлів навантаження, які використовуються для розрахунку втрат.

На основі представлених результатів у подальших дослідженнях планується удосконалити методи прогнозування та розрахунку втрат електричної енергії, зокрема з можливістю використання додаткових коефіцієнтів. Крім того, для отримання більш точних значень втрат необхідно будувати власні тестові мережі на основі реальних схем електричних мереж.

Публікація підготовлена за результатами виконання спільного українсько-литовського науково-дослідного проєкту «Smart Grid модель для оперативного керування розподільними мережами на основі методів штучного інтелекту».

I. Blinov¹, Dr. Sc. (Eng.), Prof., ORCID 0000-0001-8010-5301

P. Shymaniuk¹, Ph. D., ORCID 0000-0002-7585-7493

V. Sychova¹, Ph. D., ORCID 0000-0001-7385-1680

V. Mirosnyk¹, Cand. Sc. (Eng.), ORCID 0000-0001-9036-7268

¹Institute of Electrodynamics of the National Academy of Sciences of Ukraine

INFLUENCE OF ANOMALOUS VALUES ON THE ACCURACY OF FORECASTING LOSSES IN DISTRIBUTION NETWORKS

The relevance of the study is determined by modern trends in managing the operation modes of distribution electrical networks using Smart Grid technologies, as well as the need to reduce the costs incurred by distribution system operators for purchasing electricity. Accurate load forecasting results at network nodes for different forecasting horizons are crucial for this purpose. Sudden changes in network topology can increase the errors in loss forecasts as a single time series, negatively impacting network management efficiency and increasing the costs of electricity procurement to cover losses.

The study proposes using forecasting methods based on artificial neural networks for the calculation and prediction of electricity losses, along with a comparison of these methods. The calculations were performed using data from one of Ukraine's distribution system operators, and the test electrical network was adapted based on the CIGRE scheme for modeling electricity losses.

Since the nodal load data contained gaps and anomalies, a two-step data analysis algorithm was employed using the DBSCAN clustering method for detection and correction. As a result of loss calculations based on cleaned data, the error was reduced threefold compared to calculations based on load factors. Applying data analysis methods and forecasting methods based on artificial neural networks significantly improves the accuracy of loss calculations and minimizes errors.

Keywords: losses, node load, forecasting, neural networks, CIGRE, LSTM.

References

1. Miroshnyk V Blinov I and Shymaniuk P. "The cost of error of day ahead forecast of technological losses of electrical energy". In: *Tekhnichna elektrodynamika* 5 (2020), pp. 70–73. doi: <https://doi.org/10.15407/techned2020.05.070>. Blinov I. Problems of functioning and development of a new electricity market model in Ukraine. *Visn. Nac. Acad. Nauk Ukr*, 2021, vol. 3. pp. 20–28. DOI: <https://doi.org/10.15407/visn2021.03.020>.
2. On the electricity market: Law of Ukraine No 2019-VIII of 13.04.2017.
3. Resolution 14.03.2018 No. 307 On approval of the Market Rules.
4. NEURC Resolution "On Approval of Distribution Network Code" No. 310 of 14. 03.2018.
5. Miroshnyk V Blinov I and Sychova V. Short-term forecasting of electricity imbalances using artificial neural networks. *IOP Conf. Series: Earth and Environmental Science* (2023), pp. 1–7. DOI:10.1088/1755-1315/1254/1/012029.
6. Blinov I., Trach I., Parus Y., Khomenko V., Kuchanskyy V., Shkarupylo V. Evaluation of The Efficiency of The Use of Electricity Storage Systems in The Balancing Group and The Small Distribution System. 2021 IEEE 2nd KhPI Week on Advanced Technology (KhPIWeek), 2021, pp. 262–265, doi: 10.1109/KhPIWeek53812.2021.9569981
7. ENTSO-E RDI Roadmap 2024–2034. Innovation Missions to build the power system for a Carbon-Neutral Europe. https://eepublicdownloads.blob.core.windows.net/public-cdn-container/clean-documents/Publications/RDC%20publications/entso-e_RDI_roadmap_2024-2034_240710.pdf
8. Hou G et al. "A novel algorithm for multi-node load forecasting based on big data of distribution network". In: *International Conference on Advanced Electronic Science and Technology (AEST 2016)* (2016), pp. 655–667.
9. Hatziaargyriou N Wang X and Tsoukalas L H. "A new methodology for nodal load forecasting in deregulated power systems". In: *IEEE Power Engineering Review* (2020), pp. 48–51. doi: <http://hdl.handle.net/123456789/14566>.
10. Hatziaargyriou N Wang X and Tsoukalas L H. "A new methodology for nodal load forecasting in deregulated power systems". In: *IEEE Power Engineering Review* (2020), pp. 48–51. doi: <http://hdl.handle.net/123456789/14566>.
11. Bezerra U H Falcao D M. "Short-term forecasting of nodal active and reactive load in electric power system". In: *IEE Conf. Publication* (1986), pp. 18–22.
12. Miroshnyk V Chernenko P and Shymaniuk P. Univariable short-term forecast of nodal electrical loads of energy systems. 2020. doi: <https://doi.org/10.15407/techned2020.02.067>.
13. "Benchmark Systems for Network Integration of Renewable and Distributed Energy Resources". In: *ELECTRA* (2014). doi: http://e-cigre.org/publication/ELT_273_8-benchmark-systems-for-network-integration-of-renewable-and-distributed-energy-resources.
14. Styczynski Z A Rudion K Orths O and Strunz K. "Design of benchmark of medium voltage distribution network for investigation of DG integration". In: *IEEE Power Engineering Society General Meeting, Montreal, Canada* (2006). doi: DOI : 10 . 1109 / PES . 2006 .1709447.
15. P. Shymaniuk, V. Miroshnyk, I. Blinov Determination of electrical losses based on nodal electrical loadforecasts. *Energy: economics, technology, ecology*. 2022. Vol. 3. Pp.38-43. DOI 10.20535/1813-5420.3.2022.271484
16. Sander J Ester M Kriegel H P and Xu X. "A density-based algorithm for discovering clusters in large spatial databases with noise". In: *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96)* (1996), pp. 226–231.

Надійшла: 26.11.2024
Received: 26.11.2024