UDC 004.42:004.85
DOI https://doi.org/10.32689/maup.it.2025.2.12

Artem KOLOMYTSEV
Postgraduate Student at the Department of Software Engineering,
Software Engineer, HAPPENING TECHNOLOGY LTD,
kolomytsev1996@gmail.com
***ORCID:** 0000-0002-3231-7593*

Yuliia KUZNETSOVA
Candidate of Technical Sciences,
Associate Professor at the Department of Software Engineering,
National Aerospace University "Kharkiv Aviation Institute",
y.kuznetsova@khai.edu
***ORCID:** 0000-0001-9416-6981*

THE ROLE OF MLOPS IN ADVANCING GREEN ENERGY: AN EVALUATION OF TECHNOLOGIES AND PRACTICES

Abstract. *Machine learning–driven decision making is essential for the efficient operation of cloud-hosted virtual power plants (VPPs) aggregating hundreds to thousands of distributed energy resources (DERs). However, manually deploying and maintaining ML models at scale introduces delays, inconsistency and high operational overhead. In this paper, we survey six widely adopted MLOps frameworks–Kubeflow, Apache Airflow, MLflow, Azure ML, AWS SageMaker and Google Vertex AI–against four criteria critical to VPP environments: industry adoption, feature-set completeness, interoperability with major ML frameworks and cloud platforms, and licensing or cost constraints. Drawing on public documentation, repository activity, case studies and market research, we identify trade-offs between open-source flexibility and managed-service convenience. Our analysis shows that Apache Airflow offers the most mature and extensible pipeline orchestration for on-premise and multi-cloud VPP deployments, while Kubeflow excels in Kubernetes-native contexts. Managed services like SageMaker and Azure ML deliver faster time-to-value for teams lacking dedicated infrastructure expertise but incur higher costs and vendor lock-in. Finally, we provide domain-tailored recommendations for integrating continuous training, evaluation and monitoring into VPP forecasting workflows, demonstrating how MLOps adoption can improve prediction latency and grid responsiveness.*

The goal of this article *is to evaluate and compare leading MLOps frameworks–open-source and managed cloud services–against key criteria (adoption, feature completeness, interoperability, and cost) and to recommend the most suitable solutions for cloud-hosted virtual power plants.*

Methodology. *We selected six MLOps frameworks based on adoption, features, interoperability and cost; extracted data from official docs, repositories and market reports; scored each tool against our criteria; and distilled domain-specific recommendations for cloud-hosted VPPs.*

Scientific Novelty. *This article explores the under-researched intersection of MLOps and virtual power plants (VPPs), addressing the specific challenges of applying automated ML workflows to large-scale, cloud-hosted VPP systems. It provides the first domain-specific comparison of MLOps tools tailored to the operational and forecasting needs of VPPs.*

Conclusion. *MLOps can significantly enhance the performance and scalability of virtual power plants. This study identifies the most suitable tools for VPP use cases, highlighting Apache Airflow and Kubeflow as strong open-source options, while managed services may suit teams with limited infrastructure expertise.*

Key words: *Machine learning, MLOps, Virtual power plant, Distributed energy, Cloud technology, Forecasting.*

Артем КОЛОМИЦЕВ, Юлія КУЗНЄЦОВА. РОЛЬ АВТОМАТИЗАЦІЇ РОЗГОРТАННЯ ІНФРАСТРУКТУРИ У ПРОСУВАННІ ЗЕЛЕНОЇ ЕНЕРГЕТИКИ: ОЦІНКА ТЕХНОЛОГІЙ ТА ПРАКТИК

Анотація. *Прийняття рішень на основі машинного навчання має важливе значення для ефективної роботи хмарних віртуальних електростанцій (ВЕС), що об'єднують сотні й тисячі розподілених енергоресурсів (РЕР). Однак ручне розгортання та підтримка моделей машинного навчання в масштабі призводить до затримок, неузгодженості та високих операційних витрат. У цій статті ми проаналізували шість широко розповсюджених фреймворків автоматизації розгортанні інфраструктури машинного навчання (MLOps) – Kubeflow, Apache Airflow, MLflow, Azure ML, AWS SageMaker і Google Vertex AI – за чотирма критеріями, критично важливими для середовищ ВЕС: впровадження в галузі, повнота набору функцій, сумісність з основними фреймворками машинного навчання і хмарними платформами, а також ліцензійні або фінансові обмеження. Спираючись на публічну документацію, активність репозиторіїв, тематичні дослідження та дослідження ринку, ми визначили компроміси між гнучкістю відкритого коду та зручністю керованого сервісу. Наш аналіз показує, що Apache Airflow пропонує найбільш зрілу та розширювану оркестровку конвеєрів для локальних та мультихмарних розгортань ВЕС, в той час як Kubeflow чудово працює в контекстах на базі Kubernetes. Керовані сервіси, такі як SageMaker та Azure ML, забезпечують швидшу окупність* *інвестицій для команд, яким не вистачає спеціалістів у галузі інфраструктури, але несуть більші витрати та прив'язку до певного постачальника. Нарешті, ми надаємо рекомендації щодо інтеграції безперервного навчання, оцінки та моніторингу в робочі процеси прогнозування ВЕС, демонструючи, як впровадження автоматизації розгортанні інфраструктури машинного навчання може покращити затримку прогнозування та швидкість реагування мережі.*

Мета статті: *оцінити та порівняти провідні фреймворки автоматизації розгортанні інфраструктури машинного навчання – відкриті та керовані хмарні сервіси – за ключовими критеріями (впровадження, повнота функцій, інтероперабельність та вартість) та рекомендувати найбільш підходящі рішення для віртуальних електростанцій, розміщених у хмарі.*

Методологія. *Ми обрали шість фреймворків автоматизації розгортанні інфраструктури машинного навчання на основі впровадження, функцій, сумісності та вартості; витягли дані з офіційних документів, репозиторіїв та ринкових звітів; оцінили кожен інструмент за нашими критеріями; і сформулювали рекомендації для хмарних ВЕС для конкретних доменів.*

Наукова новизна. *У цій статті досліджується малодосліджена сфера перетину автоматизації розгортанні інфраструктури машинного навчання і віртуальних електростанцій (ВЕС), розглядаються конкретні проблеми застосування автоматизованих робочих процесів машинного навчання до великомасштабних хмарних систем ВЕС. Це перше порівняння інструментів автоматизації розгортанні інфраструктури машинного навчання, пристосованих до операційних і прогнозних потреб ВЕС, з урахуванням специфіки конкретної галузі.*

Висновки. *У результаті дослідження доведено, методи автоматизації розгортанні інфраструктури машинного навчання можуть значно підвищити продуктивність і масштабованість віртуальних електростанцій. У цьому дослідженні визначено найбільш підходящі інструменти для використання ВЕС, зокрема Apache Airflow та Kubeflow як сильні варіанти з відкритим вихідним кодом, тоді як керовані сервіси можуть підійти командам з обмеженим досвідом роботи з інфраструктурою.*

Ключові слова: *машинне навчання, автоматизація розгортання інфраструктури, віртуальна електростанція, розподілена енергетика, хмарні технології, прогнозування.*

Introduction. The market for renewable energy is experiencing substantial expansion and is rapidly accelerating. Virtual power plants (VPP) are one of the next logical steps in electrical grid development [8; 13] and they are one of the best ways to integrate renewable energy into the grid, especially residential renewable energy production and storage (like rooftop solar panels and rechargeable lithium-ion wall-mount batteries). A single VPP can manage multiple (hundreds-thousands) distributed energy resources (usually connected to the same grid), which allows them to have a bigger impact on the energy market and better financial gain [13], if compared to those DERs accessing the market on their own. VPP collects telemetry from and sends control commands to those DERs.

VPP decisions can be based on various optimisation strategies, but often those strategies are implemented using machine learning [16; 17]. In many companies (especially early-stage startups) ML models are usually managed and deployed by data science engineers manually, which is a slow and error-prone process [3]. The software development industry has already solved a similar problem by automating operations using DevOps methodology, and an identical approach can be applied to ML [15]. ML Operations (usually shortened to MLOps) is the automation of ML model learning and deployment processes. According to [3] the principles of MLOps include CI/CD pipeline automation, workflow automation, reproducibility, versioning, collaboration, continuous ML training and evaluation, and continuous monitoring. A system that follows MLOps principles may consist of multiple components that allow it to achieve those principles.

ML application in VPP has its caveats: ML models should be able to make predictions fast [15] because VPP must be able to continuously send commands to thousands of DERs, and conditions are always changing: wholesale electricity market prices, weather conditions and forecasts, user energy consumption patterns etc. Some configurations of VPP involve running ML models on DER controllers as IoT edge software [10], but this article does not cover such case – it has additional pitfalls and issues: unreliable network connection [19] on the DER side leads to model not having recent market and weather forecast data, IoT edge hardware capabilities are usually limited, so ML model should be simplified [10] (which decreases predictions quality).

Research question: Evaluate existing MLOps technologies and choose those best fitting to be used in a VPP.

Method.

Related work. We analyzed existing publications on the subject and came to the conclusion that there is a scarcity of publications focusing on MLOps applications in the VPP field.

MLOps. Research regarding MLOps has been going on actively for the past few years. Kreuzberger, Kühl and Hirschl [3] have a very detailed definition of MLOps and a list of relevant technologies for each MLOps component but do not discuss domain-specific recommendations and do not actually compare/select the best technologies. Other relevant publications [9; 12] also lack domain-specific issues discussions and limitations.

ML in a renewable energy sector. Implementation of ML models for renewable energy forecasting and software-controlled distributed energy resources is a prominent topic and multiple publications discuss this issue from various angles [15; 16; 17] even considering applying MLOps approach [15], but they do not specifically cover MLOps application to a VPP.

ML in edge IoT. Multiple software and hardware solutions exist to run (and even retrain) ML models on the edge [4; 10], but this area has a list of very specific issues and a different set of issues if compared to a VPP, which is usually backend solution running on public cloud [10; 19] and talking to DERs via REST APIs of their respective vendors [5; 7]. This article does not cover ML on edge IoT use-case and problems, but the authors agree that this may be one of the ways to implement a VPP.

Selection Criteria. The comparison in the article mostly focused on workflow/pipeline automation tools for ML, excluding optional aspects like Feature Store and automatical triggers for training new models on performance degradation. These tools could be added to ML workflow later, as they are less important and should be built on the foundation of a solid automated ML training pipeline [9].

After going through the mentioned articles, online-accessible documentation [1; 11; 18] hosted by public cloud providers six tools were selected for comparison.

Inclusion criteria:

1. Adoption – The tool must be already adopted and used in the industry, successful case studies should exist for it. Integration of a new unproven solution into the platform may be difficult, because of a lack of documentation and features [14]. Good adoption also usually means it's easier to hire expertise.

2. Feature set – the more parts of the MLOps pipeline can be covered by the same toolset/framework, the fewer new dependencies the project would need to maintain.

3. Interoperability – what ML framework tool supports, what cloud provider integrations exist out of the box, etc.

4. Licensing and cost – is it an open-source solution that can be self-hosted or closed-source privately owned solution only available on vendor's cloud?

Exclusion criteria:

1. Obsolete tools – tools that are no longer maintained or replaced.

Evaluation Framework. Open source tools would be compared on initial configuration and maintenance cost, license limitations, public cloud feature integrations suite, and feature completeness.

Official documentation, public git repositories and market research tools (6sense) were used to estimate each metric's value.

Limitations. The Availability of existing case studies of MLOps applications in the VPP market is limited, so some metric values were approximate estimates.

This article does not evaluate the ease of implementation for an edge IoT MLOps solution, mostly focusing on VPP that works as a hosted backend solution on the public cloud.

Results.

Table 1 contains the final results for the comparison based on selected metrics.

Table 1

MLOps tools comparison

Name	Adoption	Feature set	Interoperability	Licensing & cost
Kubeflow	3	4	1	1
Airflow	1	1	1	1
MLFlow	2	1	1	1
Azure ML	4	2	1	3
AWS SageMaker	5	2	1	3
Google Vertex AI	7	2	1	2

Fig. 1 shows heatmap representing the scores.

Fig. 2 shows normalized comparison of selected MLOps frameworks.

Adoption. Each column contains a place in which a corresponding tool would fit in sorted based on the selected metric (lower is better).

For open-source tools adoption was estimated based on development activity on publicly hosted Git repositories [6]. Airflow is the most mature of the selected subset of MLOps automation tools and is still the most popular and actively supported.

For closed-source tools hosted on public cloud 6sense [2] estimates were used. Azure ML is almost twice as popular as AWS Sagemaker as as of this writing. Google Vertex AI was launched a few years later than all of the above tools and market adoption data for it is not yet available, so the authors assume the worst adoption for it in comparison to other tools from the list.

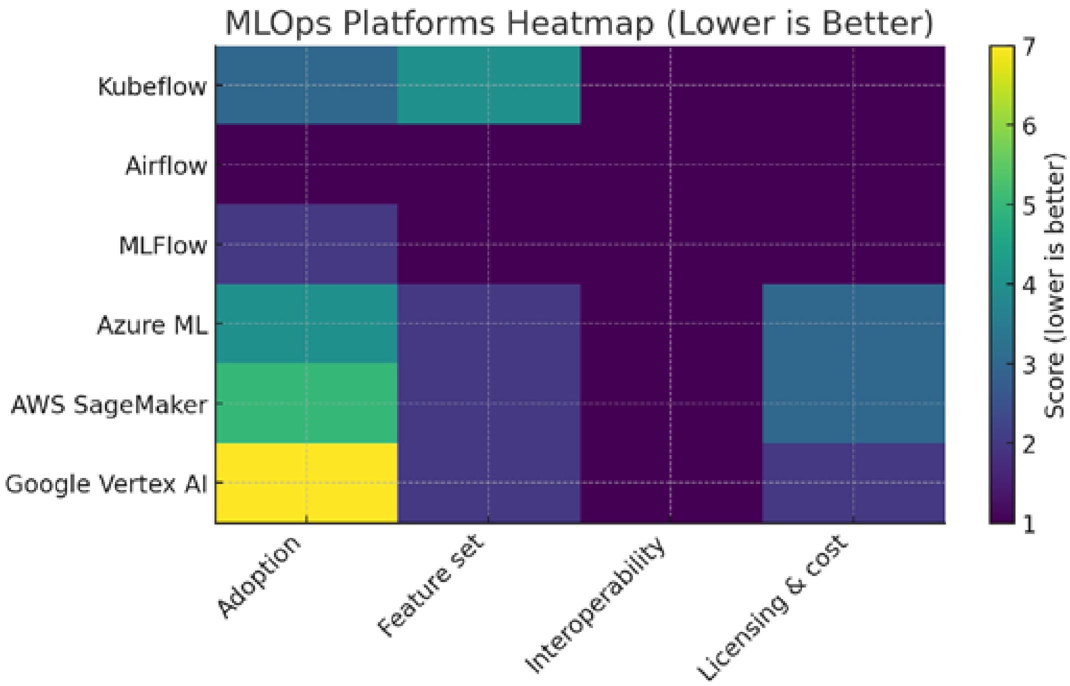

Fig. 1. Heatmap of scores

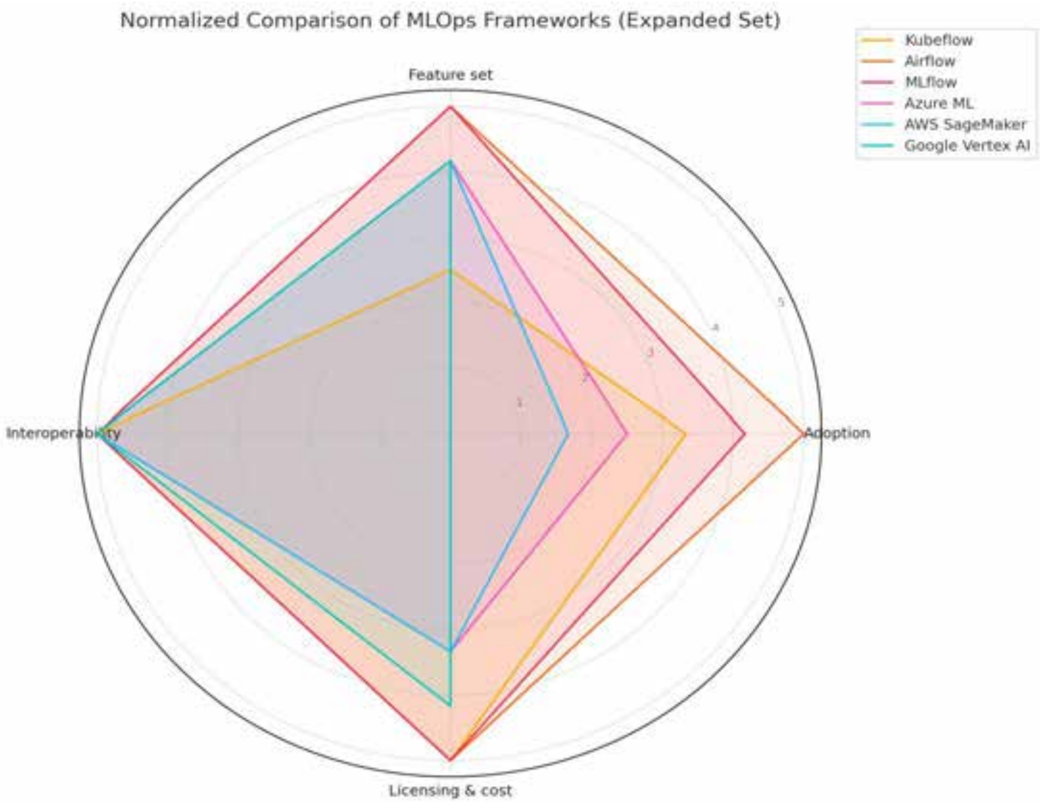

Fig. 2. Normalized comparison of MLOps frameworks

Feature set. Each column contains a place in which a corresponding tool would fit in sorted based on the selected metric (lower is better). Duplicate values represent equal/comparable feature sets.

Airflow and MLFlow are very mature open-source tools and any MLOps-related feature we might need (parametrised training, integration with orchestrators like Kubernetes etc) are implemented.

Azure ML, AWS SageMaker and Google Vertex AI have better integration with their respective cloud, but this leads to worse support for on-premise workloads or workloads on other cloud providers.

Kubeflow is very actively developed and mostly covers the use-case of training models in Kubernetes-managed environment. It's the easiest one to configure on an existing Kubernetes cluster, but if your environment is not running in Kubernetes orchestrator Kubeflow would be useless for you.

Interoperability. All listed MLOps pipeline automation tools support all major frameworks like TensorFlow and PyTorch.

Licensing & cost. Open-source tools allow you to use them free in commercial environments.

When comparing closed-source tools, Google Vertex AI is yet the cheapest to train models on.

Recommendations. Based on the results, the authors of the article recommend using Apache Airflow for most of the MLOps VPP use-cases.

If your environment is completely containerized and orchestrated by Kubernetes Kubeflow could be a good solution, but it's less mature than Airflow and Airflow can be run natively in the Kubernetes on its own.

Azure ML, AWS SageMaker and Google Vertex AI support less features and would be more expensive to host. They should be considered in organisations that do not have mature ops/infrastructure engineering teams that could maintain the MLOps pipeline on their own.

Limitations of Results. Market adoption was not measured based on real telemetry from production systems, but just an estimate.

All metrics were selected without any VPP specificity, albeit authors believe these metrics are the most important in most of the MLOps applications including VPP.

Adoption Barriers and Facilitators. MLOps adoption is mainly slowed by:

● organisational challenges (pure data-science driven team would not be able to achieve MLOps on their own, cross-functional team is required [3]);

● cost (ML is already an expensive endeavour for small or even medium-sized companies and pipeline automation might not be worth the cost for many of them if data-science team is small enough).

But, MLOps adoption can be accelerated by increased awareness and new easier to use tooling being built to abstract existing complexities of MLOps pipelines.

Conclusion. VPP would greatly benefit from MLOps integration, it would improve ML models' efficiency and decision-making processes. But, even considering these benefits MLOps is not yet widely used in VPP. It's mostly caused by the operational complexity of such tools and the cost of their introduction.

Multiple open and closed source tools exist to solve MLOps pipeline automation problems, and most of them cover all major features and support main frameworks.

Different MLOps tools have different limitations and their selection should be based on market adoption, feature set and interoperability with ML frameworks your organisation uses.

Open-source tools like Airflow or MLFlow would be best for most use cases, but their usage requires having a cross-functional data science/operations team. In some organisations it might not be a great fit and cloud-provider tools like AWS SageMaker or Azure ML should be used.

As VPPs continue to advance and assume a more critical role in renewable energy integration, the strategic application of MLOps will be vital.

Bibliography:

1. "Azure Machine Learning – ML as a Service | Microsoft Azure." Accessed: Apr. 13, 2024. URL: https://azure.microsoft.com/en-gb/products/machine-learning

2. "Azure Machine Learning vs Amazon SageMaker: Data Science And Machine Learning Comparison," 6sense. Accessed: Apr. 13, 2024. URL: https://www.6sense.com/tech/data-science-and-machine-learning/azuremachinelearning-vs-amazonsagemaker

3. D. Kreuzberger, N. Kühl, and S. Hirschl, "Machine Learning Operations (MLOps): Overview, Definition, and Architecture," *IEEE Access*, 2023. vol. 11, pp. 31866–31879, doi: 10.1109/ACCESS.2023.3262138.

4. E. Peltonen and S. Dias, "LinkEdge: Open-sourced MLOps Integration with IoT Edge," in Proceedings of the 3rd Eclipse Security, AI, Architecture and Modelling Conference on Cloud to Edge Continuum, Ludwigsburg Germany: ACM, Oct. 2023, pp. 67–76. doi: 10.1145/3624486.3624496.

5. F. Sattar, A. Husnain, and T. Ghaoud, "Integration of Distributed Energy Resources into a Virtual Power Plant-A Pilot Project in Dubai," 2023 IEEE/IAS Industrial and Commercial Power System Asia (I&CPS Asia), pp. 526–531, Jul. 2023, doi: 10.1109/ICPSAsia58343.2023.10294691.

6. G. Robles, A. Capiluppi, J. M. Gonzalez-Barahona, B. Lundell, and J. Gamalielsson, "Development effort estimation in free/open source software from activity in version control systems," *Empir Software Eng*, 2022. vol. 27, no. 6, p. 135, doi: 10.1007/s10664-022-10166-x.

7. H.-M. Chung, S. Maharjan, Y. Zhang, F. Eliassen, and K. Strunz, "Optimal Energy Trading With Demand Responses in Cloud Computing Enabled Virtual Power Plant in Smart Grids," *IEEE Trans. Cloud Comput.*, 2022. vol. 10, no. 1, pp. 17–30, doi: 10.1109/TCC.2021.3118563.

8. J. Rodríguez-García, D. Ribó-Pérez, C. Álvarez-Bel, and E. Peñalvo-López, "Novel Conceptual Architecture for the Next-Generation Electricity Markets to Enhance a Large Penetration of Renewable Energy," *Energies*, 2019. vol. 12, no. 13, Art. no. 13, doi: 10.3390/en12132605.

9. K. Salama, J. Kazmierczak, and D. Schut, "Practitioners guide to MLOps: A framework for continuous delivery and automation of machine learning.".

10. M. Merenda, C. Porcaro, and D. Iero, "Edge Machine Learning for AI-Enabled IoT Devices: A Review," *Sensors*, 2020. vol. 20, no. 9, p. 2533, doi: 10.3390/s20092533.

11. "Machine Learning Service – Amazon SageMaker – AWS," Amazon Web Services, Inc. Accessed: Apr. 13, 2024. URL: https://aws.amazon.com/sagemaker/

12. "MLOps: Continuous delivery and automation pipelines in machine learning. Cloud Architecture Center," Google Cloud. Accessed: Mar. 16, 2024. URL: https://cloud.google.com/architecture/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning

13. N. Naval and J. M. Yusta, "Virtual power plant models and electricity markets – A review," *Renewable and Sustainable Energy Reviews*, 2021. vol. 149, p. 111393, doi: 10.1016/j.rser.2021.111393.

14. R. S. Kenett, X. Franch, A. Susi, and N. Galanis, "Adoption of Free Libre Open Source Software (FLOSS): A Risk Management Perspective," 2014 IEEE 38th Annual Computer Software and Applications Conference, pp. 171–180, Jul. 2014, doi: 10.1109/COMPSAC.2014.25.

15. R. Subramanya, S. Sierla, and V. Vyatkin, "From DevOps to MLOps: Overview and Application to Electricity Market Forecasting," *Applied Sciences*, 2022. vol. 12, no. 19, Art. no. 19, doi: 10.3390/app12199851.

16. T.-Y. Kim and S.-B. Cho, "Predicting residential energy consumption using CNN-LSTM neural networks," *Energy*, 2019. vol. 182, pp. 72–81, Sep. doi: 10.1016/j.energy.2019.05.230.

17. V. Omelčenko and V. Manokhin, "Optimal Balancing of Wind Parks with Virtual Power Plants," in Frontiers in Energy Research, Nov. 2021, p. 665295. doi: 10.3389/fenrg.2021.665295.

18. "Vertex AI," Google Cloud. Accessed: Apr. 13, 2024. URL: https://cloud.google.com/vertex-ai

19. W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge Computing: Vision and Challenges," *IEEE Internet Things J.*, vol. 3, no. 5, pp. 637–646, Oct. 2016, doi: 10.1109/JIOT.2016.2579198.

Дата надходження статті: 24.06.2025
Дата прийняття статті: 30.06.2025
Опубліковано: 23.09.2025