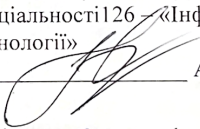



Вінницький національний технічний університет
Факультет інтелектуальних інформаційних технологій та автоматизації
Кафедра системного аналізу та інформаційних технологій

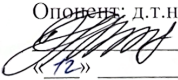
Комплексна магістерська кваліфікаційна робота на тему:

«АНАЛІЗ ТА ПЕРЕДБАЧЕННЯ ДОБОВОГО ГЕНЕРУВАННЯ ЕЛЕКТРОЕНЕРГІЇ ФОТОЕЛЕКТРОСТАНЦІЄЮ. ЧАСТИНА 1. АНАЛІЗ ТА ПРОГНОЗУВАННЯ ДОБОВОГО ГЕНЕРУВАННЯ ЕЛЕКТРОЕНЕРГІЇ»

Виконав: студент 2 курсу, групи 2ІСТ–22м спеціальності 126 «Інформаційні системи та технології»

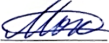

_____ Андрій БАРТЕЦЬКИЙ

Керівник: д.т.н., проф. каф. САІТ

_____ Олександр МОКІН
« 08 » _____ 12 _____ 2023 р.

Оponent: д.т.н., проф. каф. КН

_____ Ярослав ІВАНЧУК
« 12 » _____ 12 _____ 2023 р.

Допущено до захисту

Завідувач кафедри САІТ


_____ д.т.н., проф. Віталій МОКІН


« 08 » _____ 12 _____ 2023 р.

Вінниця ВНТУ – 2023 рік

Вінницький національний технічний університет
Факультет інтелектуальних інформаційних технологій та автоматизації
Кафедра системного аналізу та інформаційних технологій
Рівень вищої освіти – II-й (магістерський)
Галузь знань – 12 Інформаційні технології
Спеціальність – 126 Інформаційні системи та технології
Освітньо-професійна програма – Інформаційні технології аналізу даних та зображень

ЗАТВЕРДЖУЮ

Завідувач кафедри САІТ



 д.т.н., проф. Віталій МОКІН
« 04 » _____ 09 _____ 2023 р.

ЗАВДАННЯ НА КОМПЛЕКСНУ МАГІСТЕРСЬКУ КВАЛІФІКАЦІЙНУ РОБОТУ СТУДЕНТУ

Бартецькому Андрію Анатолійовичу

1. Тема роботи: “ Аналіз та передбачення добового генерування електроенергії фотоелектростанцією. Частина 1. Аналіз та прогнозування добового генерування електроенергії”;
керівник роботи: Олександр МОКІН, д.т.н., проф. каф. САІТ,
затверджені наказом ВНТУ « 18 » _____ 01 _____ 2023 року № 247
2. Строк подання студентом роботи « 30 » _____ 11 _____ 2023 року
3. Вихідні дані для проведення робіт: Kaggle Dataset “ energy
”<https://www.kaggle.com/datasets/andriibartetskiyi/energy>” наданий, з дозволу адміністрації фотоелектричної станції, розташованої у Вінницькій області.
4. Зміст текстової частини:
 - 1) Загальна характеристика об’єкту досліджень;
 - 2) Вибір оптимальних налаштувань для розв’язання поставленої задачі;
 - 3) Реалізація інформаційної технології прогнозування генерації електроенергії;
 - 4) Економічна частина;
5. Перелік ілюстративного матеріалу:
 - 1) Вхідний датасет генерації електроенергії на ФЕС;
 - 2) Графік генерації електроенергії ФЕС за увесь період спостережень;
 - 3) Графіки генерації електроенергії ФЕС в межах доби;
 - 4) Графік автокореляції датасету генерації електроенергії ФЕС;
 - 5) Voxplot датасету генерації електроенергії ФЕС;
 - 6) Діаграма розсіювання датасету генерації електроенергії ФЕС
 - 7) Графіки річної, тижневої, сезонної, а також денної циклічності.
 - 8) Графік прогнозування генерації електроенергії на ФЕС.

6. Консультанти розділів МКР

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
4	Наталія БУРЕННІКОВА, д. е. н., проф. каф. ЕПВМ	29.10 	25.11. 

7. Дата видачі завдання «04» 09 2023 року

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва та зміст етапу	Термін виконання		Примітка
		початок	закінчення	
1	Загальна характеристика об'єкту досліджень	04.09	30.09	вкл
2	Вибір оптимальних налаштувань для розв'язання поставленої задачі	01.10	09.10	вкл
3	Реалізація інформаційної технології прогнозування генерації електроенергії на ФЕС	10.10	28.10	вкл
4	Економічна частина	29.10	05.11	вкл
5	Оформлення матеріалів до захисту МКР	04.11	25.11	вкл

Студент

Андрій БАРТЕЦЬКИЙ

Керівник роботи

Олександр МОКІН

АНОТАЦІЯ

УДК 004.09

Бартецький А.А. Аналіз та передбачення добового генерування електроенергії фотоелектростанцією. Частина 1. Аналіз та прогнозування добового генерування електроенергії. Комплексна магістерська кваліфікаційна робота зі спеціальності 126 – інформаційні системи та технології, освітньо–професійна програма – інформаційні технології аналізу даних та зображень. Вінниця: ВНТУ, 2023. 126 с.

Укр. мовою. Бібліогр.: 115 назв; рис.: 42; табл.: 8.

В магістерській кваліфікаційній роботі проведено аналіз предметної області, обґрунтовано доцільність та актуальність даного дослідження. Проаналізовано фактори, які впливають на генерацію електроенергії фотоелектростанціями.

Досліджено існуючі методи прогнозування генерації електроенергії на ФЕС та розроблено інформаційну систему передбачення прогнозування генерації електроенергії на ФЕС за допомогою методів машинного навчання.

Ілюстративна частина складається з 8 плакатів.

У розділі економічної частини розглянуто питання про доцільність розробки та впровадження Інформаційної системи прогнозування добового генерування електроенергії.

Ключові слова: фотоелектростанція, методи машинного навчання, часові ряди, Prophet.

ABSTRACT

Bartetskyi A. Analysis and prediction of daily electricity production by a photovoltaic plant. Part 1(complex MCR). Analysis and forecasting of daily electricity production. Master's qualification work in the specialty 126 - Information Systems and Technologies, Educational and Professional Program - Information Technologies for Data and Image Analysis. Vinnytsia: VNTU, 2023. 126 c.

In Ukrainian. Bibliography: 115 titles; Figures: 42; tables: 8.

In the master's qualification work, an analysis of the subject area was carried out, the expediency and relevance of this research was substantiated. The factors that affect the generation of electricity by photovoltaic power plants are analyzed.

The existing methods of forecasting electricity generation at the FES were studied and an information system for forecasting the forecasting of electricity generation at the FES was developed using machine learning methods.

The illustrative part consists of 8 posters.

In the section of the economic part, the issue of the feasibility of developing and implementing the Information System for forecasting daily electricity generation is considered.

Keywords: photoelectric plant, machine learning methods, time series, Prophet

ЗМІСТ

ВСТУП.....	4
1 ЗАГАЛЬНА ХАРАКТЕРИСТИКА ОБ’ЄКТУ ДОСЛІДЖЕНЬ.....	8
1.1 Проблеми прогнозування генерації електроенергії на ФЕС.....	8
1.2 Огляд існуючих методів вирішення технічної проблеми.....	10
1.3 Висновки.....	24
2 ВИБІР ОПТИМАЛЬНИХ НАЛАШТУВАНЬ ДЛЯ РОЗВ’ЯЗАННЯ ПОСТАВЛЕНОЇ ЗАДАЧІ.....	25
2.1 Розвідувальний аналіз даних.....	25
2.2 Огляд методів вирішення поставленої задачі.....	39
2.3 Висновки.....	61
3 РЕАЛІЗАЦІЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ ПРОГНОЗУВАННЯ ГЕНЕРАЦІЇ ЕЛЕКТРОЕНЕРГІЇ НА ФЕС.....	63
3.1 Вибір платформи для побудови моделі прогнозування.....	63
3.2 Прогнозування генерації електроенергії на ФЕС.....	65
3.3 Висновки.....	80
4 ЕКОНОМІЧНА ЧАСТИНА.....	82
4.1 Проведення комерційного та технологічного аудиту науково-технічної розробки.....	82
4.2 Розрахунок витрат на проведення науково-технічної роботи.....	86
4.2.1 Витрати на оплату праці.....	86
4.2.2 Відрахування на соціальні заходи.....	89
4.2.3 Сировина та матеріали.....	89
4.2.4 Розрахунок витрат на комплектуючі.....	90

4.2.5 Спецустаткування для наукових (експериментальних) робіт	91
4.2.6 Програмне забезпечення на наукових (експериментальних) робіт	91
4.2.7 Амортизація обладнання, програмних засобів та приміщень	91
4.2.8 Паливо та енергія для науково-виробничих цілей	92
4.2.9 Службові відрядження	93
4.2.10 Витрати на роботи, які виконують сторонні підприємства, установи і організації	93
4.2.11 Інші витрати	93
4.2.12 Накладні (загальновиробничі) витрати	94
4.3 Розрахунок економічної ефективності науково-технічної розробки за її можливої комерціалізації потенційним інвестором	95
4.4 Висновки	101
ВИСНОВКИ	103
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	106
Додаток А (обов'язковий). Технічне завдання	127
Додаток Б (обов'язковий). Протокол перевірки магістерської кваліфікаційної роботи на наявність текстових запозичень	130
Додаток В (довідковий). Лістинг коду	131
Додаток Г (обов'язковий). Ілюстративна частина	134

ВСТУП

Актуальність теми. Останнім часом фотоелектричні станції визначають нові стандарти у виробництві електроенергії, прискорюючи перехід до екологічно чистих та стійких енергетичних систем. Політика підтримки відновлюваних джерел енергії, зокрема застосування пільгових тарифів, є важливим інструментом у стимулюванні розвитку фотоелектричних станцій. За даними Державного агентства з енергоефективності та енергозбереження України, ця стратегія призвела до істотного зростання виробленої сонячною енергією в країні. Однак важливо визначити, що успіх фотоелектричних станцій не обмежується лише пільговими тарифами. Зростаюча екологічна свідомість суспільства та підприємств сприяє популяризації цих станцій як чистого та невичерпного джерела енергії. Їх використання стає не лише стратегічним вибором для генерації електроенергії, але й актом відповідальності перед природою та майбутніми поколіннями. Нові технології та інноваційні підходи в розробці фотоелектричних систем дозволяють постійно підвищувати їхню продуктивність та знижувати вартість генерованої електроенергії. Це робить їх більш доступними та конкурентоспроможними на ринку енергетики, забезпечуючи зростаючий внесок у енергосистему. Розширення частки фотоелектричних станцій у енергосистемі України вказує на перспективність та важливість цього напрямку. Використання сонячної енергії стає ключовим елементом стратегії зменшення викидів та розвитку сталої та екологічно чистої енергетики в Україні. За даними Державного агентства з енергоефективності та енергозбереження України [1] очікуваний результат впровадження сонячної енергетики на період до 2025 року 2420 ГВт*год. Проте збільшення частки стохастичних джерел енергії в енергосистемі супроводжується виникненням додаткових ризиків, пов'язаних з їх імовірною природою і менш стабільними характеристиками, що може призвести до порушення балансової стійкості енергосистеми [2 - 4] та якості електроенергії [5]. Відповідно до закону "Про ринок електричної енергії України" [6] для виробників

електроенергії, що використовують енергію вітру і сонячного випромінювання, для яких встановлено "зелений тариф", встановлюється плата за небаланс з відповідним відсотком щороку, причому, з 1 січня 2025 року складатиме 100 %, а небаланс повинен знаходитися в межах 5%. У зв'язку із необхідністю забезпечення балансової стійкості енергосистеми необхідно створити ряд підходів та рекомендацій, використання яких дасть можливість реалізувати регульовані відновлювані джерела енергії, як складові розподілених віртуальних електростанцій, що в свою чергу є стабільними джерелами електроенергії.

Відповідно до правил ринку електроенергії [7] учасники ринку складають заявку про продаж електричної енергії на добу наперед, із вказанням обсягів які не повинні мати відхилення в межах 10%. Однією з найбільших проблем є складність прогнозування генерування електроенергії фотоелектростанціями, оскільки на сам процес впливає багато факторів, враховуючи вище згадане, спроби побудови адекватних моделей є дуже актуальними.

Мета і завдання роботи.

Метою дослідження є підвищення точності прогнозування генерування електроенергії фотоелектростанцій за рахунок використання методів машинного навчання.

Відповідно до вказаної мети в роботі необхідно розв'язати такі **основні завдання**:

- провести огляд існуючих методів і засобів для прогнозування генерації електроенергії нестабільними джерелами енергії. Визначити чинники які впливають на генерацію електроенергії фотоелектростанціями;
- провести аналіз закономірностей зміни генерації електроенергії на ФЕС;
- здійснити вибір оптимальних інформаційних технологій та моделей машинного навчання;
- побудувати модель генерування електроенергії на ФЕС;
- здійснити прогнозування генерації електроенергії на ФЕС на тестових даних та оцінити ефективність роботи моделі.

Об'єктом дослідження є процес прогнозування генерування електроенергії на ФЕС.

Предметом дослідження є методи інформаційних технологій, які використовувались для аналізу та прогнозування генерації електроенергії на фотоелектростанції.

У центрі завдань цього дослідження стоїть мета побудувати ефективну модель для прогнозування генерації електроенергії на фотоелектростанції. Результати розвідувального аналізу даних створять основу для реалізації моделей, спрямованих на точне прогнозування генерації електроенергії на фотоелектростанції протягом певного періоду часу. Важливим етапом роботи буде відбір оптимальної моделі, яка виявиться найбільш точною та надійною у прогнозуванні. Це дозволить потенційним користувачам, таким як операторам енергосистеми чи іншим зацікавленим сторонам, здійснювати прогнози генерації електроенергії на фотоелектростанції з високою точністю. Отримані результати будуть мати важливе значення для підвищення надійності та стійкості енергосистеми, а також для підвищення рівня енергетичної безпеки країни. Цей підхід сприятиме ефективнішому управлінню енергетичними ресурсами та забезпечить стабільність генерації електроенергії, що важливо для сучасного енергетичного сектора.

Новизна одержаних результатів. Подальшого розвитку набула модель генерування електроенергії, яка, на відміну від існуючих, для ідентифікації параметрів використовує методи машинного навчання, що дозволяє підвищити точність передбачення.

Практична цінність одержаних результатів. Результати магістерської кваліфікаційної роботи мають допомогти операторам фотоелектростанцій більш точно скласти графік генерування електроенергії, що в умовах об'єднаної енергосистеми дозволить краще здійснювати балансування енергосистеми, відповідно це підвищує надійність роботи енергосистеми, а також покращує якість енергетичних показників. Результати роботи можуть бути універсалізовані

та використані для прогнозування генерації на будь-яких нестабільних відновлюваних джерелах генерації електроенергії.

Апробація результатів магістерської кваліфікаційної роботи.

Результати роботи доповідались на ЛІІ Науково-технічній конференції факультету інтелектуальних інформаційних технологій та автоматизації з публікацією тез ВНТУ (м. Вінниця, 2023р.)

Публікації результатів магістерської кваліфікаційної роботи.

Опубліковано тези на ЛІІ Науково-технічній конференції факультету інтелектуальних інформаційних технологій та автоматизації ВНТУ (м. Вінниця, 2023 р.) [8].

1 ЗАГАЛЬНА ХАРАКТЕРИСТИКА ОБ'ЄКТУ ДОСЛІДЖЕНЬ

1.1 Проблеми прогнозування генерації електроенергії на ФЕС

Однією з основних труднощів, що виникають у роботі нестабільних відновлюваних джерел енергії (ВДЕ), зокрема сонячних електростанцій, є забезпечення балансової стійкості енергосистеми. Підвищення частки нерегульованих та нестабільних джерел енергії у енергосистемі може викликати розбалансування та призвести до серйозних проблем. Однією з головних аспектів є вирішення питання балансу між виробництвом та споживанням електроенергії, особливо коли виробництво залежить від випадкових факторів, наприклад інтенсивність сонячної радіації. Великі коливання у виробництві можуть порушити баланс і викликати нестабільність в енергосистемі. З іншого боку, для власників сонячних електростанцій велике відхилення фактичного графіка генерування від передбаченого може призвести до серйозних фінансових збитків. Точність прогнозування та управління виробництвом є ключовими елементами для забезпечення ефективності та економічної доцільності сонячних електростанцій. Такі виклики стають особливо важливими у контексті розвитку великих мережевих систем, де інтеграція великої кількості сонячних електростанцій вимагає вдосконалених технологій та стратегій для збалансування та забезпечення стабільності енергетичної інфраструктури.

Оскільки фотоелектричні станції розраховані на використання сонячної радіації для генерації електроенергії, їх ефективність надто чутлива до погодних умов. Швидкозмінні явища, такі як раптова локальна захмареність неба, можуть спричинити раптові зміни в інтенсивності сонячної радіації, що, в свою чергу, призведе до значних коливань у виробництві електроенергії. Отже, врахування швидкозмінних погодних умов є ключовим аспектом в управлінні та прогнозуванні ефективності сонячних електростанцій, спрямованим на зменшення можливих втрат і оптимізацію виробництва з урахуванням змін у погодних умовах. У статті [9], автори представляють рішення проблеми впливу

метеопараметрів на точність прогнозу генерації фотоелектричними станціями. Одним із ключових елементів цього підходу є використання прогнозу погоди на наступний день та інформації із метеопостів, які встановлені безпосередньо на електростанції. Такий комбінований підхід спрямований на зменшення збитків власників електростанцій, що виникають внаслідок неточностей у прогнозуванні.

Важливо відзначити, що автори не розглядали методів оперативного керування фотоелектростанцією для компенсації похибок у прогнозуванні. Це відкриває можливості для подальших досліджень та впровадження інтегрованих стратегій управління, які враховують оперативні зміни у прогнозі та дозволяють забезпечити ефективніший контроль над генерацією на фотоелектричних станціях. В роботі [10] автори пропонують розраховувати оптимальну потужність резерву електростанції, яка враховує нестабільність процесу генерування ВДЕ. Цей підхід дозволяє враховувати імовірні значення невідпущеної електроенергії, що може виникнути внаслідок нестабільності в генерації електроенергії від ВДЕ. Розрахунок оптимальної потужності резерву є важливим етапом в управлінні електростанцією, оскільки це дозволяє ефективно використовувати потенційні резерви та мінімізувати втрати внаслідок нестабільності генерації. Такий підхід може стати важливим кроком у розв'язанні проблеми балансової стійкості в системах з великою часткою електростанцій на основі ВДЕ. У статті [11] автори пропонують розробку віртуальних електростанцій як функціональної одиниці Об'єднаної енергетичної системи. У роботі проведено аналіз взаємодії гідро- та вітрогенеруючих установок на локальну енергетичну мережу. Важливим аспектом є те, що отримані результати виявилися неефективними в умовах нового законодавства, зокрема закону "Про ринок електричної енергії України".

Ця розробка вказує на потребу в постійному вдосконаленні та адаптації енергетичних систем до сучасних законодавчих та ринкових умов. Урахування змін у законодавстві є критичним аспектом для ефективного функціонування віртуальних електростанцій та інших елементів енергетичних систем.

1.2 Огляд існуючих методів вирішення технічної проблеми

Неточність відомих методів прогнозування пропонується вирішувати на основі критерію мінімального відхилення від прогнозованого графіка генерування [12]. Відповідно до цього критерію генерація електроенергії на фотоелектростанції (ФЕС) розглядається в контексті наступних трьох випадків:

1. В сценарії використання фотоелектричної станції, яка підключена до локальної електричної мережі без застосування електроакумуляторів та без взаємодії із стабільними ВДЕ, виникає обмеження, що полягає в неможливості покриття миттєвих провалів генерування. Це обумовлено відсутністю системи накопичування електроенергії для компенсації флуктуацій в генерації.

В такому випадку модель оптимізації стикається з обмеженням щодо неможливості компенсації миттєвих піків перегенерування. Це означає, що електростанція не може ефективно вирішувати ситуації, коли виробництво електроенергії не відповідає споживанню в конкретний момент часу.

Недоліком такого підходу є обмежена можливість забезпечення повної керованості генерації, що, в умовах нового закону "Про ринок електричної енергії України", може призвести до фінансових збитків підприємства у вигляді штрафів за недовипуск електричної енергії [13]. Застосування системи оптимізації фотоелектричних станцій у вказаному випадку може ефективно запобігти виникненню миттєвих локальних перенапруг. Це особливо важливо для уникнення можливих аварійних ситуацій, які можуть виникнути внаслідок пробією ізоляції в електротехнічних установках. Система оптимізації може допомагати вирішувати проблеми, пов'язані із змінами у виробництві та споживанні електроенергії, попереджаючи перенапруги та забезпечуючи баланс між виробництвом та споживанням. Це сприяє збереженню стабільності електричної системи та уникненню можливих технічних неполадок. Такий підхід до оптимізації роботи фотоелектричних станцій може покращити їхню

ефективність та надійність, зменшуючи ризики аварій та підвищуючи загальну безпеку електроенергетичної системи [14].

2. Використання фотоелектричної станції із накопичувачами електричної енергії надає значні переваги у досягненні оптимальних енергетичних показників та компенсації провалів генерування та регенерування електроенергії. Завдяки наявності накопичувачів можливе повне використання виробленої енергії, навіть при тимчасовій відсутності сонячної активності. Обмеження, такі як ємність накопичувача та залишковий заряд, можуть визначати максимальну кількість енергії, яку можна зберігати та використовувати. Проте, гнучкість використання та висока швидкодія накопичувачів дозволяють ефективно компенсувати миттєві відхилення у генеруванні, що робить фотоелектричну станцію стабільним та надійним джерелом енергії. Цей підхід сприяє оптимізації використання відновлювальних джерел енергії, підвищує ефективність їхньої експлуатації та забезпечує стійкість енергетичної системи [15]. Незважаючи на переваги фотоелектричних станцій із накопичувачами електричної енергії, існують і деякі недоліки, які потрібно враховувати при їхньому впровадженні. Один із головних недоліків – це висока вартість системи накопичувачів, яка може виявитися значною і навіть порівнянною з вартістю самої фотоелектричної станції. Це може ускладнити економічну вигоду від впровадження таких систем. Крім того, важливим аспектом є екологічний аспект використання накопичувачів електроенергії. Виробництво та утилізація елементів накопичувачів може призводити до екологічного впливу та виникнення проблем з управління відходами. Окрім цього, невизначеність у сфері екологічних аспектів може породжувати питання щодо сталості та екологічної придатності фотоелектричних станцій із накопичувачами. Це питання стає актуальним у контексті загального підвищення уваги до екологічних стандартів та відповідальності в сфері виробництва та використання енергетичних технологій.

3. У випадку сумісної роботи ФЕС із стабільними ВДЕ, вирішення задачі оптимізації передбачає оперативне керування їхньою спільною роботою. Основна

мета - забезпечити мінімальне відхилення від сумарного графіка генерування ФЕС та стабільної ВДЕ. У цьому контексті, стабільні ВДЕ можуть використовуватися з резервом потужності для компенсації провалів генерування електричної енергії ФЕС. Серед стабільних ВДЕ, які можуть працювати паралельно із ФЕС в Україні, варто визначити малі гідроелектростанції (МГЕС), біогазові електроустановки (БГЕУ) та повітряокумулюючі електростанції (ПАЕС) [16]. Цей підхід дозволяє використовувати стабільні ВДЕ як резерв для забезпечення стійкості та ефективності роботи ФЕС, зокрема, покриваючи недоліки генерації, що можуть виникнути через зміни у погодних умовах або інші фактори

Розглянемо випадок, коли мала ГЕС розташована в одній локальній ЕС разом із ФЕС (рисунок 1.1).

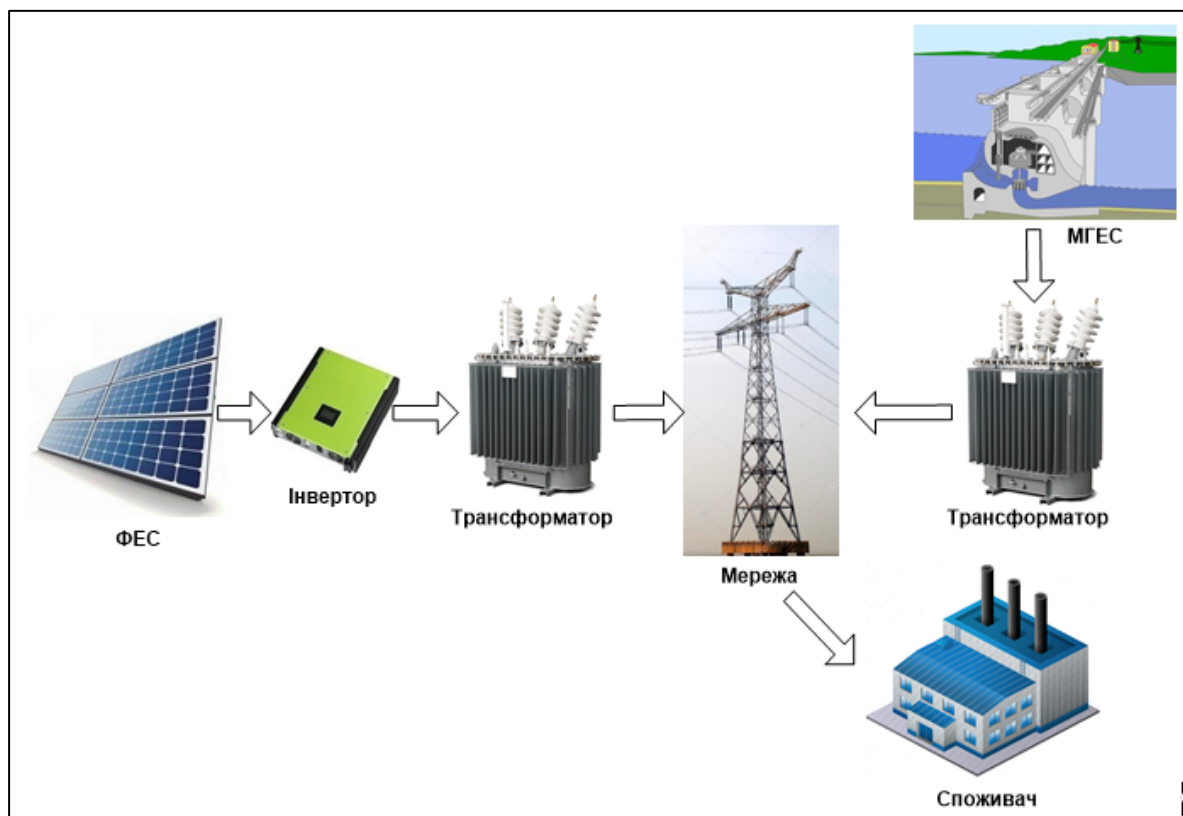


Рисунок 1.1 – Структурна схема сумісної роботи ФЕС та МГЕС в локальній ЕС

У цьому випадку, провали в генерації ФЕС можуть бути компенсовані за рахунок резерву потужності, наданого малими гідроелектростанціями (МГЕС) з

інвертором, як зображено на рисунку 1.2. Використання інвертора в МГЕС дозволяє забезпечити високу швидкість регулювання, що призводить до покращення енергетичних показників та ефективного використання електроенергії.

Цей підхід до оптимізації роботи ФЕС за участю МГЕС із застосуванням інвертора дозволяє більш ефективно вирішувати проблеми провалів в генерації, забезпечуючи стабільність системи та оптимальне використання резерву потужності.

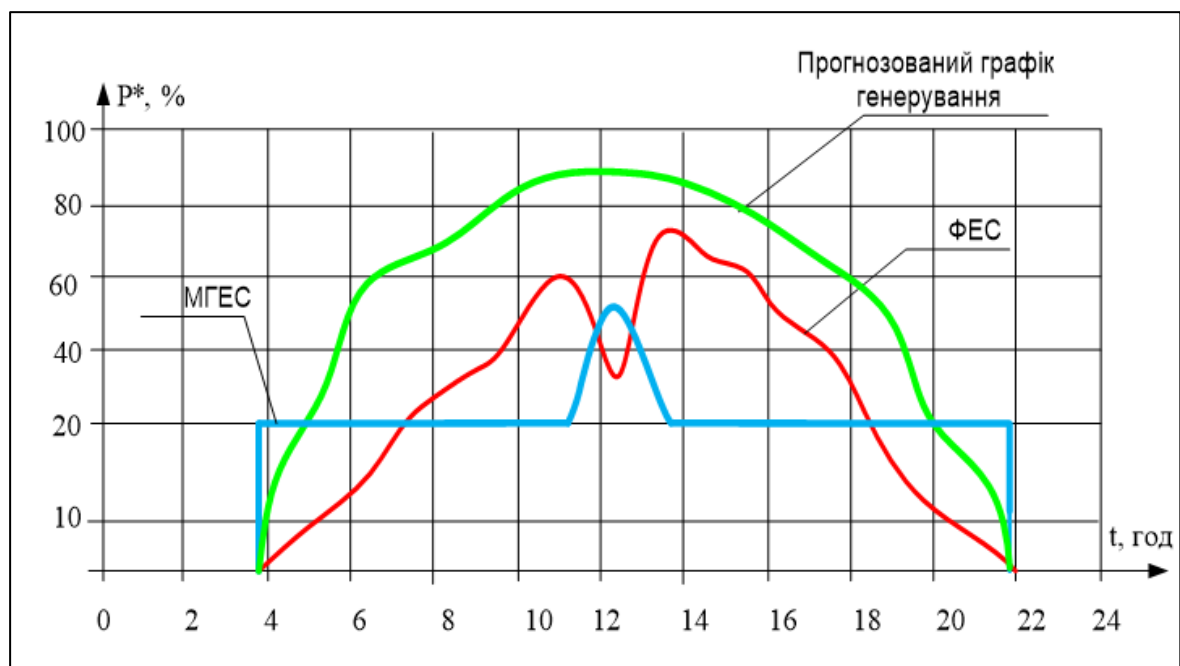


Рисунок 1.2– Графік сумісної роботи ФЕС та МГЕС

У випадку, коли ФЕС та мала гідроелектростанція (МГЕС) знаходяться на невеликій відстані одна від одної, це створює зручні умови для оптимізації просторового розташування. Такий підхід дозволяє економити місце, особливо враховуючи, що при будівництві МГЕС залишаються невикористані земельні ділянки, які можуть стати ідеальним варіантом для розміщення ФЕС.

Використання цих невикористаних земельних ділянок для встановлення сонячних батарей ФЕС дозволяє оптимально використовувати обмежений простір і максимізувати виробництво зеленої енергії на даній території. Такий

синергетичний підхід може бути ефективним з точки зору ефективності використання ресурсів та раціонального використання доступного простору.

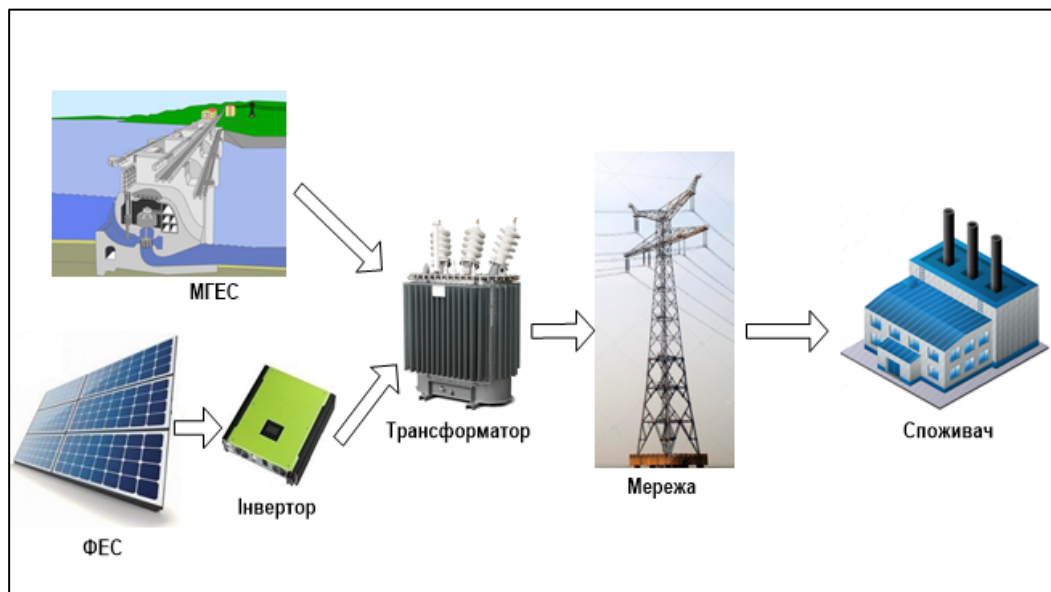


Рисунок 1.3– Структурна схема сумісної роботи ФЕС та МГЕС в локальній ЕС

ФЕС співпрацюючи із міні-ГЕС, може використовуватися в схожому режимі роботи. У випадку такого партнерства міні-ГЕС спрямована на покриття можливих провалів у генерації електроенергії.

Це означає, що в моменти, коли ФЕС не може забезпечити достатню кількість електроенергії, міні-ГЕС може взяти на себе функцію додаткового джерела електропостачання. Такий сценарій роботи дозволяє компенсувати можливі тимчасові збої у виробництві електроенергії та забезпечує більш стабільний вихід електроенергії у систему. При цьому графік генерування матиме вигляд (рисунок 1.4).

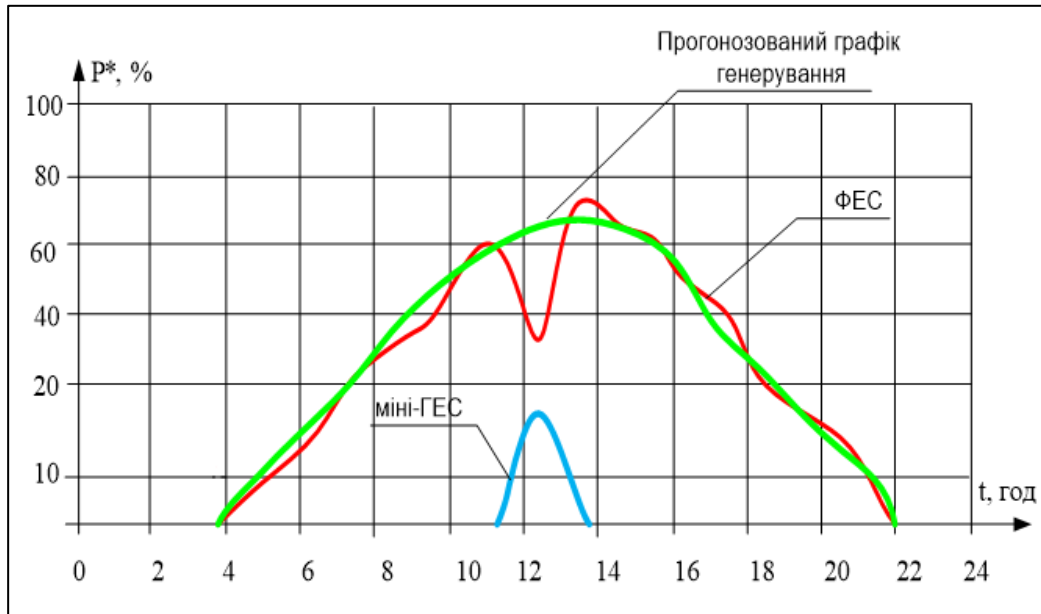


Рисунок 1.4 – Графік сумісної роботи ФЕС та міні-ГЕС

Використання міні-гідроелектростанцій (міні-ГЕС) як допоміжного джерела енергії для компенсації можливих провалів у генерації фотоелектричних станцій (ФЕС) є раціональним підходом. Цей метод дозволяє швидко реагувати на випадкові тимчасові відключення ФЕС та забезпечує високий рівень керованості системи. Проте, варто зауважити, що жоден з розглянутих методів не передбачає використання ФЕС в комплексі із стабільним джерелом енергії для компенсації провалів у генерації. Це може вплинути на фінансові та фізичні обмеження проєкту, включаючи бюджет на будівництво та експлуатацію самої ФЕС. Такий комплексний підхід потребує детального аналізу та урахування різних факторів для забезпечення ефективності та стабільності енергетичної системи.

В роботі [17] автором проводилося дослідження проблеми передбачення генерації електроенергії на ФЕС. Автор намагається врахувати ряд кліматичних параметрів, зокрема температуру повітря а також інсоляцію. Проте використання бібліотеки RandomForestRegressor не забезпечує відповідної точності передбачення для поставленої задачі, а також отримати достовірні кліматичні дані отримати безпосередньо над місцем розташування ФЕС досить складно.

Загалом, прогнозування виконується на основі різних періодів часу, відомих як горизонти прогнозування. Дуже короткострокове прогнозування ($1 \text{ с} \leq 1 \text{ год}$) допомагає розподіляти електроенергію в реальному часі, оптимізувати ресурси та вирівнювати потужність [18]. Короткострокове прогнозування (1 год–24 год) допомагає підвищити надійність мережі та імпровізувати роботу енергосистеми [19]. Середньострокове прогнозування (тиждень – місяць) встановлює графік планування та обслуговування енергосистем шляхом прогнозування наявної електроенергії. Довгострокове прогнозування (місяць - рік) включає органи управління розподілом і передачею електроенергії, планування виробництва електроенергії, а також заходи щодо забезпечення енергоспоживання та безпеки [20, 21].

Для прогнозування фотоелектричної генерації використовувалися різні методи, такі як ARMA, ARIMA, ARMAX, пов'язані авторегресійні та динамічні системи (CARDS), регресія та дерева регресії. Точність прогнозу цих методів краща для короткострокових горизонтів. Однак точність знижується зі збільшенням горизонту прогнозування та розмірів випуску [22–23]. Нелінійні дані також є обмеженням цих методів. Ґрунтуючись на відстеженні та прогнозуванні хмар, фотографії неба та супутники були використані для прогнозування сонячного опромінення на ультракороткостроковій основі [24–25]. Точність прогнозування методів на основі зображень безпосередньо залежить від алгоритмів обробки зображень. Однак, базуючись на супутникових даних із низькою роздільною здатністю та обмеженому охопленні зображень неба з землі, точність прогнозування цих методів потребує подальшого вдосконалення. Числовий прогноз погоди (ЧПП) використовується для середньострокового (на 15 днів вперед) прогнозу сонячної радіації. Однак його застосування обмежене через обмеження щодо пошуку даних, які накладають вітчизняні метеорологічні відділи [26–32].

Штучна нейронна мережа (ANN) у [33–36] та адаптивна нейронечітка система логічного висновку (ANFIS) у [37] є одними з методів машинного

навчання, які застосовуються для прогнозування сонячної енергії. Вони краще справляються з нелінійними системами та справляються зі змінною поведінкою сонячної енергії. Однак проблеми випадкових початкових даних, локальних мінімумів, переобладнання та підвищеної складності через багат шарову структуру впливають на надійність енергосистем [38, 39]. Тим часом опорні векторні машини (SVM) показують кращу точність прогнозу для прогнозування сонячної енергії в [40–42]. Однак вони надзвичайно чутливі до таких параметрів, як коефіцієнт штрафу (C), функція ядра та радіус труби (ϵ). Тому зробити відповідний вибір для цих параметрів є складним завданням. Ваги та зміщення прихованих вузлів у машині екстремального навчання (ELM) вибираються випадковим чином [43–45].

Нові досягнення в технології штучного інтелекту (ШІ) призвели до використання різних методів глибокого навчання в певних сферах для подолання недоліків традиційних нейронних мереж та інших методів машинного навчання [46–47]. Глибоке навчання має наступні основні риси на відміну від фізичних, наполегливих і статистичних методів [48]: неконтрольоване вилучення ознак, домінуючу здатність до узагальнення та навчання на великих даних.

У багатьох публікаціях науковці та інженери проводили дослідження щодо прогнозування виробництва фотоелектричної енергії, яке в основному поділяється на дві категорії: фізичні моделі та статистичні моделі. У фізичних моделях прогнозне значення сонячного опромінення та інформація про географічне розташування в поєднанні з режимом роботи фотоелектричних модулів використовуються для виконання математичного моделювання [49], а система накопичення енергії використовується для вирішення негативних наслідків нестабільної генерації та низької надійності електропостачання. У практичних застосуваннях під час використання фотоелектричної генерації неминуче виникають помилки через втрату електроенергії та інші проблеми. Поліпшення властивостей матеріалу є найбільш прямим шляхом до підвищення ефективності фотоелектричного перетворення [50, 51]. У статистичній моделі в

основному покладаються на історичні дані фотоелектричних електростанцій. Тому дослідники віддають перевагу алгоритмам штучного інтелекту. До них відносяться алгоритми машинного навчання, такі як штучні нейронні мережі [52, 53] (ANN) і опорні векторні машини [54, 55] (SVM). Ці алгоритми широко використовуються в галузі прогнозування виробництва фотоелектричної енергії. Наприклад, у статті [56] запропоновано ефективну модель прогнозування SVM для вивчення зв'язку між метеорологічними даними та виробництвом електроенергії фотоелектростанціями. Автори [57] запропонували розширену модель на основі SVM для отримання більш точного набору даних. Точність прогнозування моделей машинного навчання часто залежить від якості заданого набору даних і налаштувань внутрішніх параметрів. Подібним чином невеликі відмінності в наборах даних можуть призвести до значних змін у результатах прогнозування [58]. Тому одна за одною з'являються гібридні моделі прогнозування. Оптимізуючи використаний набір даних і обчислюючи найкращі параметри, модель прогнозування може отримати найкращий прогноз. В роботі [59] запропоновано систему оцінки для короткострокового прогнозування фотоелектричної потужності та порівняльний аналіз між різними моделями машинного навчання та методами вибору функцій, і результати показали, що метод екстремального градієнтного посилення (XGBoost) перевершив окремі методи машинного навчання. За словами авторів [60], завдяки поєднанню XGBoost із технологією розробки функцій, важлива інформація була вилучена з прогнозів погоди для досягнення покращеної точності прогнозування.

У порівнянні з традиційними методами машинного навчання, моделі глибокого навчання мають кращу продуктивність і здатні виявляти внутрішні зв'язки у багатовимірних даних [61]. Моделі глибокого навчання включають згорткові нейронні мережі (CNN) [62, 63], рекурентні нейронні мережі (RNN) [64] та інші класичні моделі, а також їх варіанти і комбіновані моделі. Як варіант моделі RNN, мережа довгострокової короткочасної пам'яті (LSTM) може ефективно фіксувати довгострокові залежності часових рядів і стала дуже

популярною в галузі короткострокового прогнозування вихідної потужності. Наприклад, експериментальні результати в [65] показали, що ефективність методу прогнозування виробництва електроенергії на основі LSTM краща, ніж у багатошарових перцептронів (MLP) і глибоких згорткових мереж. Автори [66] використовували мережу LSTM для прогнозування сонячного опромінення за попередній день, і її результат був кращим, ніж результати нейронної мережі зворотного поширення (BP) і лінійної регресії найменших квадратів. Автори в [67] запропонували гібридну модель глибокого навчання CNN-LSTM, яка використовує багатошарову CNN для виділення ознак і рівень LSTM для прогнозування, таким чином ефективно покращуючи ефект прогнозування LSTM.

Незалежно від обраного алгоритму прогнозування, етап обробки даних є проблемою, яку не можна ігнорувати. Вихідна потужність має нелінійні характер. Розкладання такого часового ряду на кілька підпоследовностей може ефективно зменшити складність даних і є ефективним засобом для підвищення точності прогнозування використовуваної моделі [68]. Загальні методи декомпозиції послідовності включають емпіричну декомпозицію (EMD), ансамблеву EMD (EEMD) і вейвлет-декомпозицію (WD) [69, 70]. Однак результати наведених вище методів декомпозиції послідовності спричиняють модальне накладення псевдонімів, що ускладнює прогнозування. Будучи методом, який виконує декомпозицію та реконструкцію послідовності [71], сингулярний спектральний аналіз (SSA) може ефективно розкласти послідовність на послідовність тенденцій, періодичну послідовність і шумову послідовність без вибору апріорної базисної функції або складного операційного процесу, і ця методика досягає кращої об'єктивності та адаптивності [72]. Він підходить для різних інженерних задач і широко використовується в прогнозуванні енергії вітру та прогнозуванні енергетичного навантаження [73,74]. Наприклад, у [75] автор розклав ряд потужності вітру на дві підпоследовності (ряд тренду та ряд шуму) через SSA та використав гібридну нейронну мережу Лагерра для прогнозування розкладених сигналів. У [76] була запропонована багатоступенева модель прогнозування

швидкості вітру, яка попереджає, шляхом поєднання варіаційної декомпозиції моди (VMD) і SSA з моделлю LSTM.

Обробка характеристик погоди також є важливою ланкою в прогнозуванні фотоелектричної потужності. Хоча вихідна потужність коливається, діапазон коливань вихідної потужності є подібним за одного типу погоди. Таким чином, при створенні набору даних для прогнозування, кластеризація даних у подібні дні відповідно до пов'язаних типів погоди може зменшити надмірність даних і помилки прогнозування [77]. Зазвичай використовувані методи кластеризації включають К-найближчих сусідів (KNN) [78] і кластеризацію К-середніх (K-means) [79]. Автори [80] використовували алгоритм кластеризації нечітких С-середніх (FCM) для кластеризації та аналізу історичних метеорологічних даних і інформації про прогноз погоди, а також використовували алгоритм найменших квадратів SVM (LSSVM) для прогнозування. У [81] кластеризація К-середніх була використана для вибору подібних історичних даних із днів прогнозування як навчальних зразків, а потім повний EEMD з адаптивним шумом (CEEMDAN) і стробований рекурентний блок (GRU) були використані для прогнозування фотоелектричної потужності. Результати моделювання показали, що запропонована модель перевершила інші моделі. Можна побачити, що під час обробки наборів даних про виробництво фотоелектричної енергії, незалежно від того, кластеризує типи погоди чи шукає подібні дні, встановлення відповідних моделей для різних типів даних може підвищити точність отриманого прогнозу. Наведені вище методи розбивають весь набір даних на безліч менших наборів даних для навчання моделі прогнозування. Коли обсяг даних недостатній, декомпозиційний набір даних може бути дуже малим, що може легко призвести до недостатньої кількості навчальних вибірок для алгоритму та переобладнання результатів прогнозу [82]. Підводячи підсумок, у цьому документі пропонується гібридна модель прогнозування на основі SSA-LSTM. Розкладання SSA виконується на заданій послідовності вихідної потужності із сильною мінливістю; виділяється послідовність тренду, періодична послідовність і шумова

послідовність вихідної потужності і аналіз головних компонент виконується на послідовності. Важливі компоненти виділяються для реконструкції послідовності, а моделі прогнозування LSTM окремо створюються для реконструйованих послідовностей. Мета цього полягає в тому, щоб дозволити LSTM безпосередньо вивчати регулярні дані послідовності, зменшити складність моделі та підвищити точність передбачення.

Короткострокове прогнозування генерації, погодинне, денне або пікове навантаження, є предметом інтенсивних досліджень. Авторами в [83], виділено різницю між методами та методологіями прогнозування; перша відноситься до групи моделей, що належать до подібного сімейства, наприклад авторегресійного інтегрованого ковзного середнього (ARIMA) [84], тоді як остання відноситься до загальної основи рішення, застосовної до різних сімейств моделей. Такі моделі можуть включати, серед іншого, вибір функцій, підхід подібного дня та ієрархічне прогнозування.

З точки зору моделювання, статистичні моделі часових рядів, такі як ARIMA [85], і моделі машинного навчання, такі як штучні нейронні мережі (ANN) [86, 88] і опорні векторні машини (SVM) [89], є одними з найпопулярніших програм для короткострокового прогнозування генерації. Модель експоненційного згладжування та її різні варіації, такі як моделі подвійного сезонного експоненціального згладжування, показують хороші результати [90], але менш популярні в реальних програмах через їх нездатність включати природні факторів, наприклад, прогноз погоди. Популярним підходом серед дослідників і практиків є поєднання різних моделей і методологій, що призводить до так званих гібридних підходів. Наприклад, автори [91] розглянули аналіз основних компонентів (PCA) і множинну лінійну регресію (MLR) для прогнозування щоденного навантаження.

В останні роки моделі глибокого навчання, дуже успішні для програм розпізнавання образів зображень і обробки природної мови (NLP), з'явилися як чудові кандидати для програм короткострокового прогнозування генерації

[92, 93]. Зокрема, рекурентні нейронні мережі (RNN), такі як довготривала короткочасна пам'ять (LSTM) [94], є популярним вибором для прогнозування часових рядів, оскільки вони спеціально розроблені для обробки послідовних даних. Застосування різних RNN для прогнозування кількох синтетичних і реальних часових рядів можна знайти в [95], причому результати непереконливі щодо того, яка архітектура є кращою для прогнозування попиту, але радше вказують на те, що слід дотримуватися підходу до конкретного випадку. У [96] для прогнозування житлового навантаження запропоновано структуру прогнозування LSTM, яка перевершує кілька еталонних моделей. Автори [97] запропонували вдосконалену модель LSTM для прогнозу споживання електроенергії енергосистемою в столичній зоні у Франції. Гібридний підхід для короткострокового прогнозування навантаження представлено в [98], де сигнал навантаження обробляється за допомогою використання як LSTM, так і згорткових нейронних мереж (CNN). У [99] була запропонована спеціальна модель глибокого навчання, яка поєднує кілька рівнів CNN для виділення ознак, шари LSTM для прогнозування та паралельні щільні шари для перетворення екзогенних змінних.

Незалежно від обраної основної методики, важливою проблемою з методологічної точки зору є робота з притаманними сезонними моделями. Стандартні підходи для роботи з окремою сезонністю включають сезонне експоненціальне згладжування та сезонні моделі ARIMA. Погодинне навантаження, однак, є високочастотним часовим рядом, що регулюється кількома сезонними моделями, а саме щоденними, тижневими та річними. У зв'язку з цим традиційні підходи були розширені для врахування кількох сезонних моделей, таких як подвійне сезонне експоненціальне згладжування [100] або динамічна гармонічна регресія з термінами Фур'є [101]. Зокрема, для ШНМ загальний консенсус полягає в тому, що десезонізація перед навчанням призводить до покращення продуктивності [102]. Таким чином, стандартний підхід полягає в тому, щоб спочатку видалити сезонність, змодельовати сезонно

скоригований ряд, випустити прогноз і, нарешті, знову ввести сезонність. Іншим популярним підходом до ШНМ, спрямованим на зменшення складності моделі, є сегментація часового ряду на підсерії для кожної години доби. Потім розгортається стратегія прямого прогнозування, а окремі прогнози об'єднуються в кінці [103].

Декомпозиція часових рядів включає альтернативний підхід роботи зі складною сезонністю. По-перше, вихідний часовий ряд розкладається на кілька компонентів, або адитивну, або мультиплікативну, що зазвичай відповідає тренду, сезонність (коливальні або періодичні компоненти) і випадковий шум. У свою чергу, ці компоненти моделюються та прогнозуються незалежно перед фазою повторної сезонізації, де вони об'єднуються в остаточний прогноз. Відомі методи включають сезонну та трендову декомпозицію з використанням Лесса (STL) [104], яка може бути розширена для кількох сезонів, тригонометричну декомпозицію (TBATS) [105] та сингулярний спектральний аналіз (SSA) [106]. У [107] була запропонована модель LSTM на основі декомпозиції для роботи з кількома сезонними моделями. У цій моделі часовий ряд розкладається на різні сезонні компоненти, які використовуються як природні фактори в LSTM. Результати показали, що цей підхід, позначений як сезонний екзогенний підхід, показує покращену продуктивність у прогнозуванні однорідних часових рядів, таких як споживання навантаження. SSA застосовувався в енергетичній сфері головним чином для цілей прогнозування швидкості вітру [108, 110]. У [111] SSA було об'єднано з LSTM разом із варіаційною декомпозицією режиму (VMD) і Extreme Learning Machine (ELM) для того, щоб проводити багатоетапне прогнозування швидкості вітру. Подібним чином [112] автор представив модель для багатокрокового прогнозування швидкості вітру на основі VMD і SSA в поєднанні з моделлю LSTM.

1.3 Висновки

У даному розділі проведено аналіз проблематики генерацій електроенергії фотоелектричними електростанціями. Було окреслено важливість прогнозування генерації електроенергії фотоелектричними електростанціями при роботі на об'єднану енергосистему. Розглянуто чинники які впливають на генерацію.

Також був проведений огляд існуючих підходів та моделей машинного навчання в питанні прогнозування генерації електроенергії нестабільними джерелами енергії зокрема фотоелектростанціями.

Отже, враховуючи технологічну необхідність в балансуванні енергосистеми та імовірнісний характер генерації електроенергії фотоелектростанціями, які працюють в об'єднаній енергосистемі, виникає необхідність у точному короткостроковому прогнозі генерації електроенергії.

Відомі методи прогнозування варіюються в залежності від задачі прогнозування, а саме довгострокове прогнозування та короткострокове прогнозування. Переважна більшість існуючих моделей прогнозування базуються зокрема на моделях: ARMA, ARIMA, ARMAX, штучна нейронна мережа ANN та адаптивна нейронечітка система логічного висновку ANFIS, рекурентні нейронні мережі (RNN), також в існуючих моделях прогнозування використовують розпізнавання образів зображень і обробки природної мови (NLP). Недоліком існуючих моделей є те, що для навчання автори використовують дуже велику кількість даних, в іншому випадку моделі не забезпечують необхідну точність прогнозу, а саме в межах 10%. У випадку зі збільшенням кількості даних, модель стає менш гнучкою, оскільки переважна більшість фотоелектричних станій є новими або за період у 5-10 років неодноразово були модернізовані.

2 ВИБІР ОПТИМАЛЬНИХ НАЛАШТУВАНЬ ДЛЯ РОЗВ'ЯЗАННЯ ПОСТАВЛЕНОЇ ЗАДАЧІ

2.1 Розвідувальний аналіз даних

Розвідувальний аналіз даних (EDA) - це метод аналізу, який виявляє загальні закономірності в даних. До цих закономірностей належать викиди та особливості даних, які можуть бути неочікуваними.

EDA є важливим першим кроком у будь-якому аналізі даних. Розуміння того, де відбуваються викиди, і як пов'язані змінні, може допомогти створити статистичні аналізи, які дають значущі результати. У випадку даних біологічного моніторингу, місця можуть бути під впливом кількох стресорів. Таким чином, початкові дослідження кореляцій стресорів є критичними, перш ніж намагатися пов'язати стресорні змінні з біологічними відгуковими змінними. EDA може надати висновки щодо можливих причин, які слід включити в оцінку причинності. про можливі причини, які слід включити в причинно-наслідкову оцінку.

У цьому розділі здійснюється глибокий аналіз та розвідка даних, які використовуються для побудови моделі прогнозування генерації електроенергії на приватній фотоелектростанції, розташованій у Вінницькій області (рисунок 2.1). Цей етап є ключовим у розумінні особливостей та характеристик даних, а також визначенні їхньої придатності для моделювання. Дані для аналізу складаються з результатів моніторингу генерації електроенергії на фотоелектростанції протягом двох років. Ці дані включають в себе різноманітні показники, зокрема:

- «Date» – Дата фіксації генерації ;
- «P» – Сумарна енергія за добу[кВт*год.];
- «h1-h48» – Енергія в точках фіксації генерації[кВт*год.].

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	
	id	DATE	P	h13	h14	h15	h16	h17	h18	h19	h20	h21	h22	h23	h24	h25	h26	h27	h28
2	1080113	31.12.2022	812	0	0	0	0	0	5	38	37	67	75	98	90	75	83	69	
3	1661195	07.11.2022	3170	0	0	3	18	14	22	91	187	219	155	33	719	490	247	558	
4	1659626	06.11.2022	1803	0	0	0	3	9	23	54	77	70	94	153	232	202	273	185	
5	1658057	05.11.2022	938	0	0	0	4	30	59	60	52	69	71	73	87	74	50	34	
6	1656450	04.11.2022	982	0	0	0	38	93	139	175	7	0	0	0	0	0	0	0	
7	1654843	03.11.2022	1739	0	0	2	18	28	54	111	293	412	165	523	88	42	0	0	
8	1652863	02.11.2022	582	0	0	0	6	32	36	36	35	69	49	65	77	42	0	21	
9	1651294	01.11.2022	5149	0	0	22	83	301	472	577	659	720	762	785	460	0	0	0	
0	1649687	31.10.2022	605	0	0	0	21	133	0	0	0	0	0	0	0	0	0	0	
1	1648080	30.10.2022	4674	0	0	0	33	108	83	83	121	157	180	262	198	534	754	328	
2	1646435	29.10.2022	5867	0	0	0	0	30	116	278	261	508	536	677	690	717	505	240	
3	1644828	28.10.2022	7693	0	0	0	1	37	72	221	234	326	515	527	673	628	548	694	
4	1643221	27.10.2022	4677	0	0	0	0	1	15	30	45	64	96	126	584	771	662	256	
5	1641614	26.10.2022	3237	0	0	0	0	0	19	53	104	82	120	100	125	211	196	235	
6	1640007	25.10.2022	562	0	0	0	0	6	48	59	49	63	83	32	28	21	16	6	
7	1638400	24.10.2022	5756	0	0	0	0	47	125	139	246	453	487	426	638	411	394	636	

Рисунок 2.1– Базові ознаки датасету по генерації електроенергії ФЕС

Датасет також містить ще «id», але, як правило його видаляють, оскільки даний параметр не несе корисної інформації.

Саме ознака «h1-h48» і вважається ознакою, яка відображає поточне генерування.

Оскільки відповідно до завдання поставлена задача прогнозування, для дослідження доцільно використовувати моделі часових рядів.

Однією із особливостей моделей часових рядів є те, що і тестовий, і валідаційний датасети вибираються в кінці даних спостережень. Основна задача валідаційного датасету – допомогти відібрати оптимальну модель, яка найкраще буде прогнозувати майбутні значення, а не значення в минулому.

Оскільки датасет містить велику кількість даних, що уповільнює роботу інтерпретатора, як наслідок виконання програми триває досить довго, тому видалення та переформатування даних здійснювалося програмно локально.

Для первинної обробки даних використовувалася студія розробки IntelliJ IDEA, програма написана мовою програмування java.

Основна програма конвертації даних наведена на рисунку 2.2.

```
package com.company;

import com.company.converter.DsConverter;

import java.io.*;

public class Main {

    public static void main(String[] args) throws IOException {

        if (args.length < 2) {
            System.out.println("Command line parameters are incorrect! It must be passed: [input_file] [output_file]");
            return;
        }

        InputStream fis = new FileInputStream(args[0]);
        OutputStream fos = new FileOutputStream(args[1]);

        System.out.println("Conversion started..");
        long startTs = System.currentTimeMillis();
        DsConverter.convert(fis, fos);
        System.out.printf("Conversion completed, it took %d ms\r\n", System.currentTimeMillis() - startTs);

        fis.close();
        fos.close();
    }
}
```

Рисунок 2.2– Основна програма конвертації даних

Програма перевіряє, чи були передані необхідні параметри командного рядка. Якщо кількість параметрів менше двох, виводиться повідомлення про неправильність та програма завершується.

Потоки використовуються для читання даних з вхідного файлу (fis) та запису даних у вихідний файл (fos).

Викликається метод convert класу DsConverter, який, має реалізацію для конвертації даних. Тут fis і fos використовуються для передачі потоків даних у цей метод.

Після завершення конвертації закриваються відкриті потоки для правильного вивільнення ресурсів.

Підпрограма DsConverter для конвертації файлів наведена на рисунку 2.3

```

2 usages
public class DsConverter {

    2 usages
    private final static String CSV_DELIMITER = ";";

    1 usage
    public static void convert(InputStream fis, OutputStream fos) {
        Scanner scanner = new Scanner(fis);
        PrintWriter writer = new PrintWriter(fos);

        scanner.nextLine(); // skip header
        writer.println(getHeader());

        while (scanner.hasNextLine()) {
            String inputLine = scanner.nextLine();
            List<String> outputLines = DsLineConverter.convert(inputLine, columnNumbers: 49, CSV_DELIMITER);
            outputLines.forEach(writer::println);
        }

        scanner.close();
        writer.flush();
        writer.close();
    }

    1 usage
    private static String getHeader() { return "DATE" + CSV_DELIMITER + "VALUE"; }
}

```

Рисунок 2.3 – Підпрограма DsConverter

DsConverter відповідає за конвертацію вхідного потоку даних в інший формат та запис його у вихідний потік даних (наприклад, у файл). Давайте розглянемо його деталі:

Метод (convert) отримує вхідний потік (InputStream) та вихідний потік (OutputStream) і використовує їх для зчитування та запису даних відповідно. Створюється об'єкт Scanner для зчитування даних з вхідного потоку. Створюється об'єкт PrintWriter для запису даних у вихідний потік.

Метод (Scanner.nextLine()) пропускає перший рядок (заголовок) у вхідних даних та записує заголовок у вихідні дані за допомогою методу getHeader().

Пропускає перший рядок (заголовок) у вхідних даних та записує заголовок у вихідні дані за допомогою методу `getHeader()`.

Далі в циклі обробляється кожен рядок вхідних даних, викликається метод `DsLineConverter.convert` для конвертації рядка та записує результат у вихідний потік. Закривається вхідний та вихідний потоки. Важливо викликати `flush()` перед закриттям `PrintWriter`, щоб гарантувати, що всі дані були записані. Метод (`getHeader`) генерує та повертає рядок заголовка для вихідного файлу.

В результаті для подальших досліджень вже використовувався відформатований датасет, фрагмент якого наведений на рисунку 2.4.

	A	B	C
1	ds	y	
2	31.12.2022 0:00	0	
3	31.12.2022 0:30	0	
4	31.12.2022 1:00	0	
5	31.12.2022 1:30	0	
6	31.12.2022 2:00	0	
7	31.12.2022 2:30	0	
8	31.12.2022 3:00	0	
9	31.12.2022 3:30	0	
10	31.12.2022 4:00	0	
11	31.12.2022 4:30	0	
12	31.12.2022 5:00	0	
13	31.12.2022 5:30	0	
14	31.12.2022 6:00	0	
15	31.12.2022 6:30	0	
16	31.12.2022 7:00	0	
17	31.12.2022 7:30	0	
18	31.12.2022 8:00	0	
19	31.12.2022 8:30	5	
20	31.12.2022 9:00	38	
21	31.12.2022 9:30	37	
22	31.12.2022 10:00	67	
23	31.12.2022 10:30	75	
24	31.12.2022 11:00	98	

Рисунок 2.4 – Датасет по генерації електроенергії ФЕС

В результаті отримано датасет в зручному форматі для подальшої обробки, де:

- «ds» – Дата та час фіксації генерації ;
- «y» – Енергія в точках фіксації генерації[кВт*год].

Візуалізація отриманого датасету наведена на рисунку 2.5.

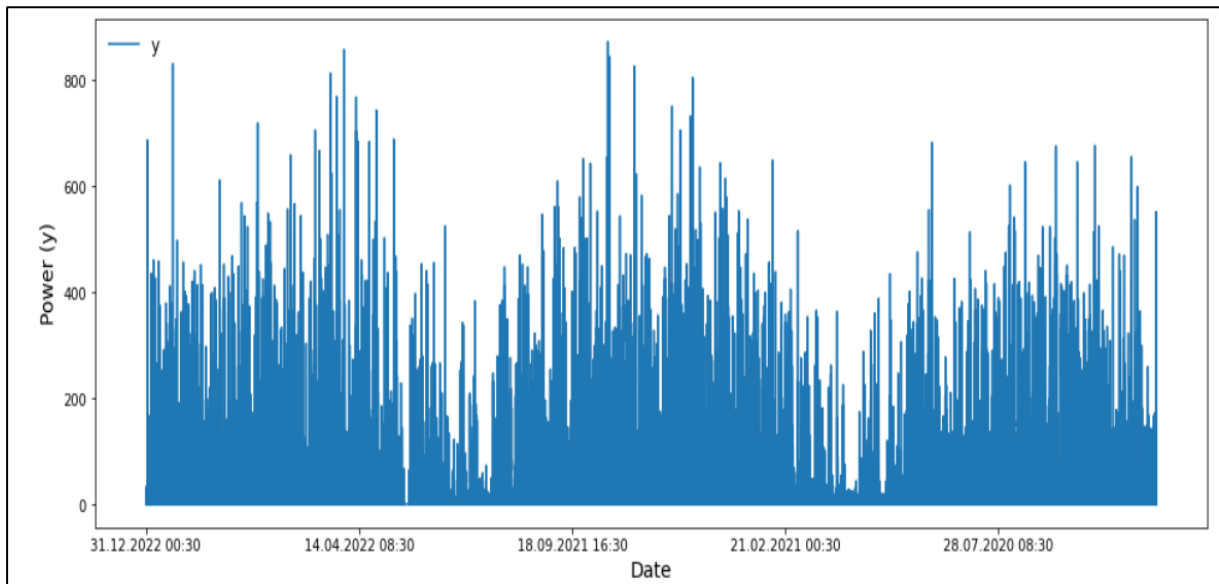


Рисунок 2.5 – Графік генерації електроенергії ФЕС за увесь період спостережень

Щоб отримати краще розуміння у сезонні зміни в генерації електроенергії, дослідимо графіки генерації в межах доби. Це дозволить нам виявити, як змінюється виробництво електроенергії протягом різних годин доби і визначити можливі сезонні патерни (рис. 2.6).

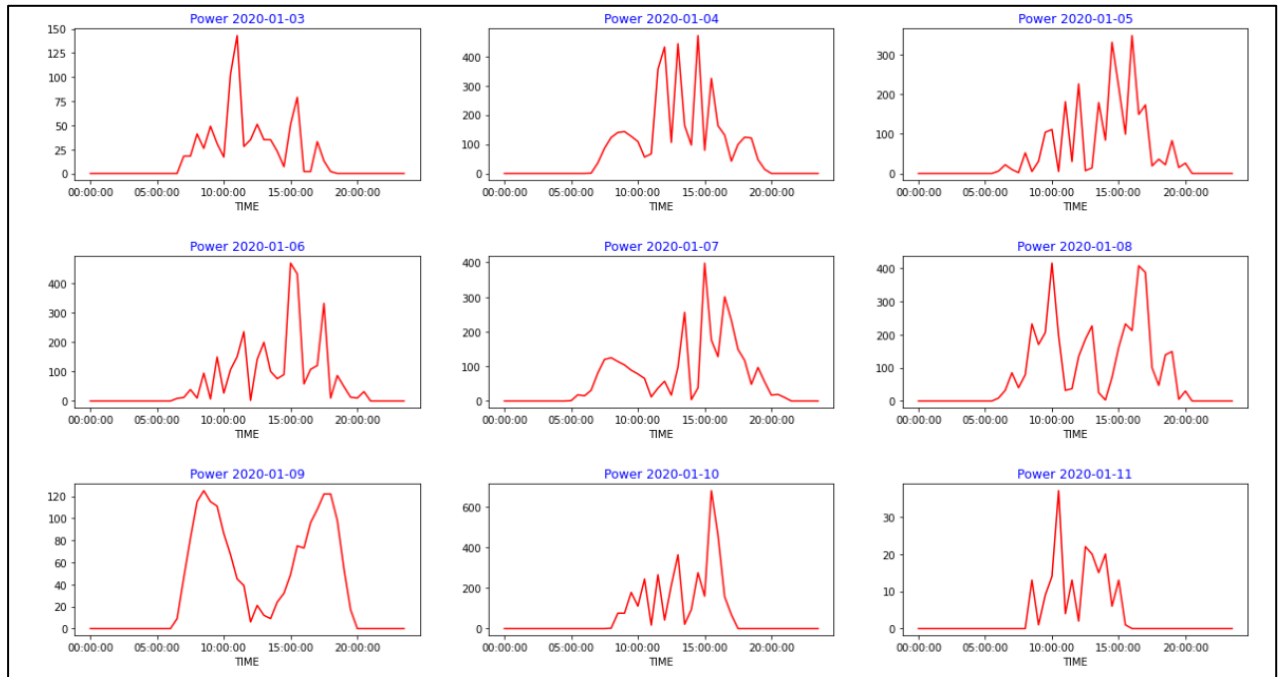


Рисунок 2.6 – Графіки генерації електроенергії ФЕС в межах доби

Як видно з графіків за кілька днів, генерація дійсно має досить нестабільний характер, зокрема суттєво спотворюють графіки провали в генерації. Це закономірні провали пов'язані з випадковими погодними процесами, проте яскравих аномалій на графіках не спостерігається.

Для оцінки отриманого часового ряду оцінимо його статистичні метрики. Статистичні метрики є числовими показниками, які використовуються для характеристики різних аспектів часового ряду. В контексті часових рядів електроенергії такі метрики можуть допомогти оцінити основні характеристики ряду та його змін.

Основні статистичні показники включають:

Середнє (Mean): Визначається як сума всіх значень ряду, поділена на кількість спостережень. Математично:

$$Mean = \frac{\sum_{i=1}^n y_i}{n} \quad (2.1)$$

де y_i - значення у момент часу;

i , n - кількість спостережень.

Стандартне відхилення (Standard Deviation): Вказує на ступінь розподілу значень від середнього. Чим вище стандартне відхилення, тим більше варіації в даних.

Медіана (Median): Значення, яке розташоване в середині впорядкованого ряду. Менше чутлива до викидів порівняно з середнім.

Мінімум та максимум: Найменше і найбільше значення в ряді.

Коефіцієнт варіації (Coefficient of Variation): Відношення стандартного відхилення до середнього значення, виражене у відсотках.

Для отримання статичних метрик нашого часового ряду потрібно виконати код, який наведено на рисунку 2.7.

```
[188]: # Output of statistical metrics
print("Середнє:", df_data['y'].mean())
print("Стандартне відхилення:", df_data['y'].std())
print("Медіана:", df_data['y'].median())

Середнє: 151.23970985155196
Стандартне відхилення: 260.6145362544969
Медіана: 0.0
```

Рисунок 2.7 – Статистичні показники датасету генерації електроенергії ФЕС

Середнє (Mean): Середнє значення ряду дорівнює 151.24. Це представляє узагальнене значення генерації електроенергії за розглянутий період.

Стандартне відхилення (Standard Deviation): Стандартне відхилення досить велике (260.61), що може вказувати на значну варіацію та коливання в генерації електроенергії. Це може бути пов'язано з різними факторами, такими як коливання в погоді, технічні аспекти сонячної електростанції та інші.

Медіана (Median): Медіана дорівнює 0.0. Це вказує на те, що половина значень менше або рівна 0.0, а інша половина - більше. Враховуючи велике

стандартне відхилення, це може бути пов'язано з великою кількістю нульових або дуже малих значень у ряді.

Автокореляція є важливим інструментом у часовому аналізі і використовується для визначення кореляції між часовим рядом та його затримками (лагами). Автокореляція вимірює ступінь схожості між значеннями часового ряду в різних моментах часу.

Головні цілі графіка автокореляції включають:

Виявлення патернів в часових рядах: Графік автокореляції може допомогти виявити повторювані патерни або циклічні залежності в часовому ряді. Наприклад, якщо є циклічна компонента з періодом, автокореляційна функція може показати кореляцію на кожному лозі цього циклу.

Визначення затримок (лагів): Графік автокореляції допомагає визначити, на якому лозі відбуваються значущі кореляції. Це може бути корисним при визначенні оптимального значення параметра для моделей прогнозування.

Виявлення структури часового ряду: За допомогою графіка автокореляції можна виявити наявність сезонності, циклів або інших регулярних патернів у часовому ряді.

Графік автокореляції зазвичай містить лінії або стовпчики, які представляють кореляції на різних лагах. Це дає змогу аналізувати структуру та патерни у часовому ряді для покращення його розуміння та використання у прогнозуванні.

Для побудови графіка автокореляції потрібно виконати код наведений на рисунку 2.8.

```
[189]: # Autocorrelation plot
plot_acf(df_data['y'])
plt.show()
```

Рисунок 2.8 – Побудова графіка автокореляції датасету генерації електроенергії ФЕС

Графік автокореляції датасету генерації електроенергії ФЕС наведений на рисунку 2.9

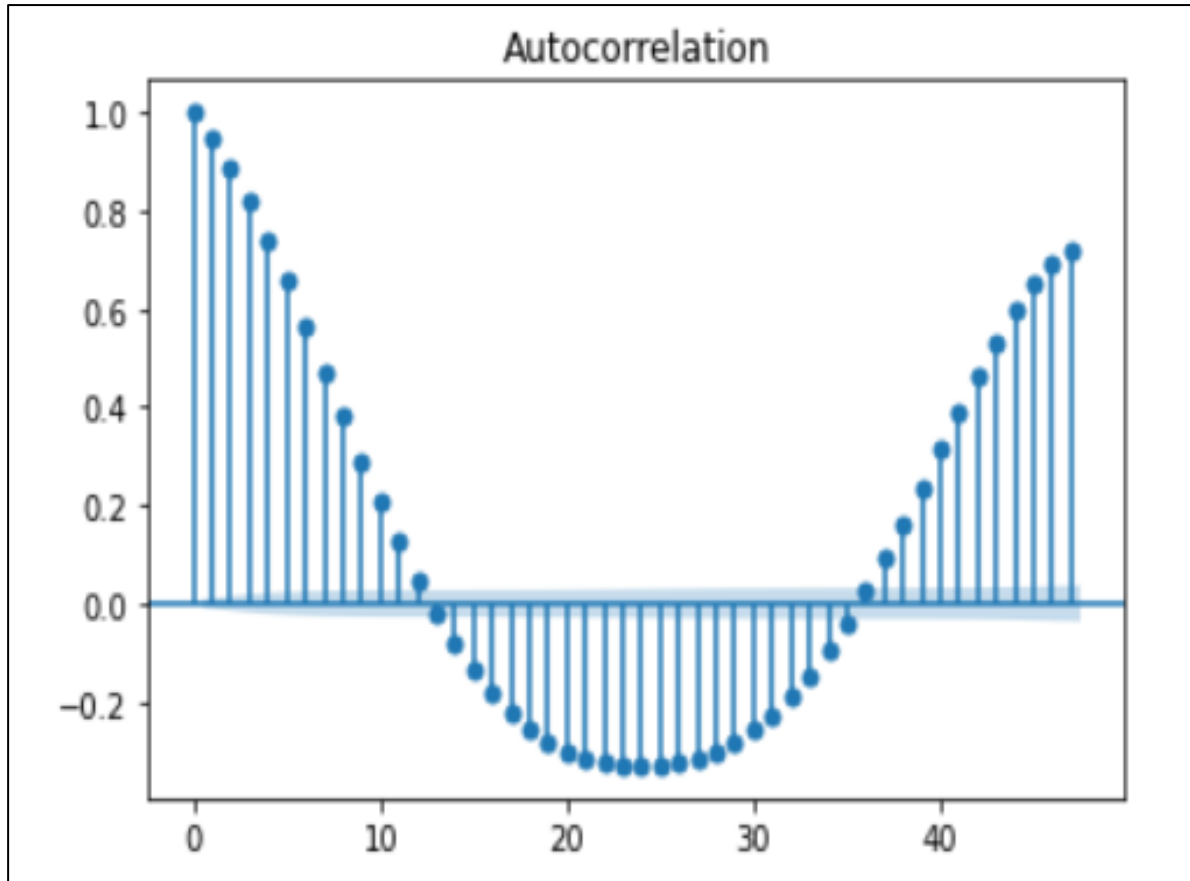


Рисунок 2.9 – Графік автокореляції датасету генерації електроенергії ФЕС

Такий характер автокореляції вказує на яскраво виражену сезонність, що характерно для генерації електроенергії сонячними електростанціями і обумовлено зміною світлового дня в залежності від пори року.

Гістограма підсумовує розподіл даних шляхом розміщення спостережень за інтервалами (які також називаються класами або бінами) і підрахунку кількості спостережень у кожному інтервалі. На осі Y може бути кількість спостережень, відсоток від загальної кількості, частка від загальної кількості (або ймовірність) або щільність (у якій висота стовпчика, помножена на ширину інтервалу, відповідає відносній частоті інтервалу). Вигляд гістограми може залежати від того, як визначено інтервали.

Побудуємо гістограму для досліджуваного датасету (рис. 2.10)

```
[41]:  
sns.set(color_codes=True)  
sns.set_palette(sns.color_palette("muted"))  
sns.distplot(df_data["y"].dropna())  
  
plt.show()
```

Рисунок 2.10 – Побудова гістограми розподілу даних датасету генерації електроенергії ФЕС

Гістограма розподілу даних датасету генерації електроенергії ФЕС наведена на рисунку 2.11.

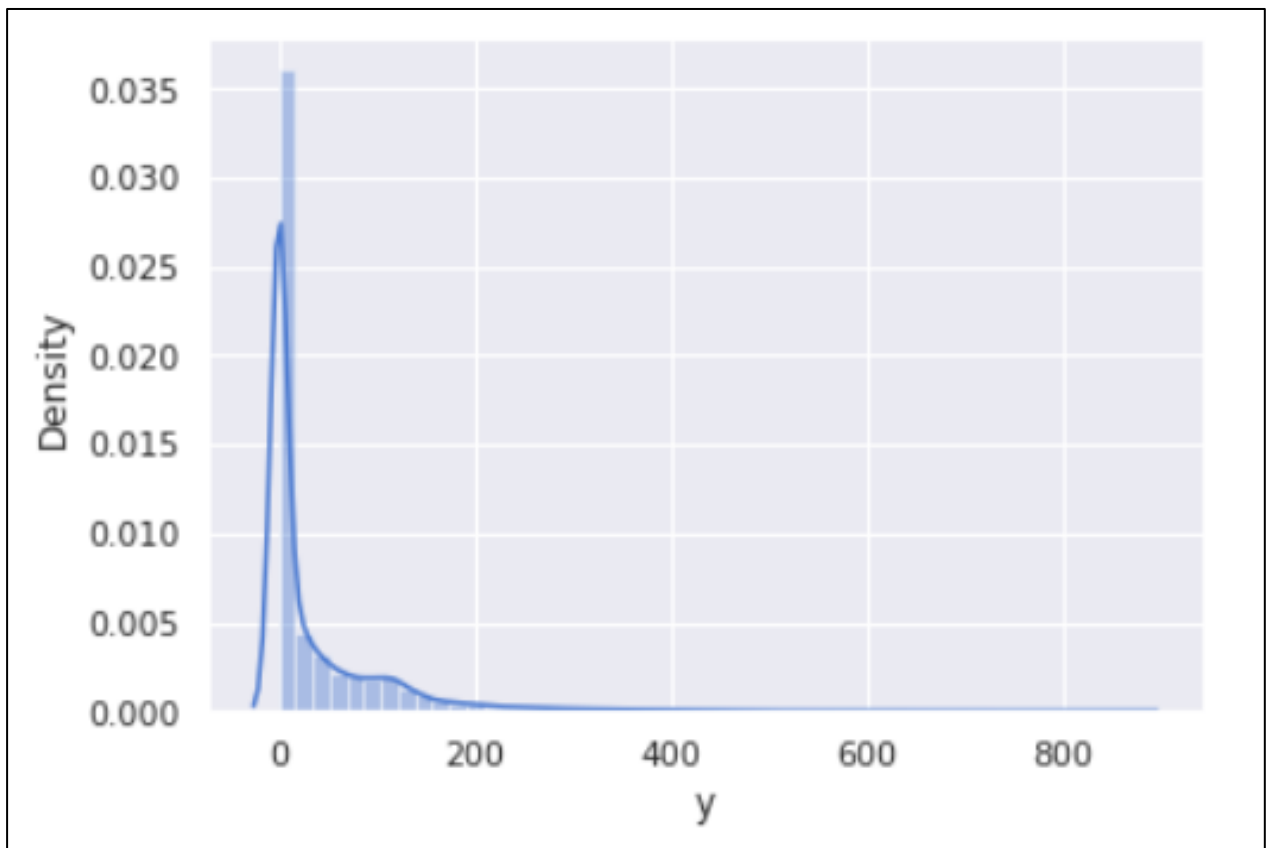


Рисунок 2.11 – Гістограма розподілу даних датасету генерації електроенергії ФЕС

Коробчатий графік (також відомий як boxplot) надає компактний підсумок розподілу змінної. Стандартна коробкова діаграма складається з прямокутника, визначеного 25-м і 75-м процентами, горизонтальної лінії або точки на прямокутнику на медіані та вертикальних ліній (вусів), проведених від кожного шарніра (квартиля) до екстремального значення. У невеликій варіації стандартної прямокутної діаграми вуси простягаються на відстань прольоту від шарніра. Коробкові графіки особливо корисні для порівняння розподілу різних підмножин однієї змінної.

От же побудуємо boxplot графік для датасету генерації електроенергії ФЕС (рис. 2.12).

```
df_data['Year'] = df_data['ds'].dt.year

plt.figure(figsize=(12, 6))
sns.boxplot(x='Year', y='y', data=df_data)
plt.title('Boxplot потужності за роками')
plt.show()
```

Рисунок 2.12 – Побудова boxplot графіка даних датасету генерації електроенергії ФЕС

Boxplot графік датасету генерації електроенергії ФЕС наведена на рисунку 2.13.

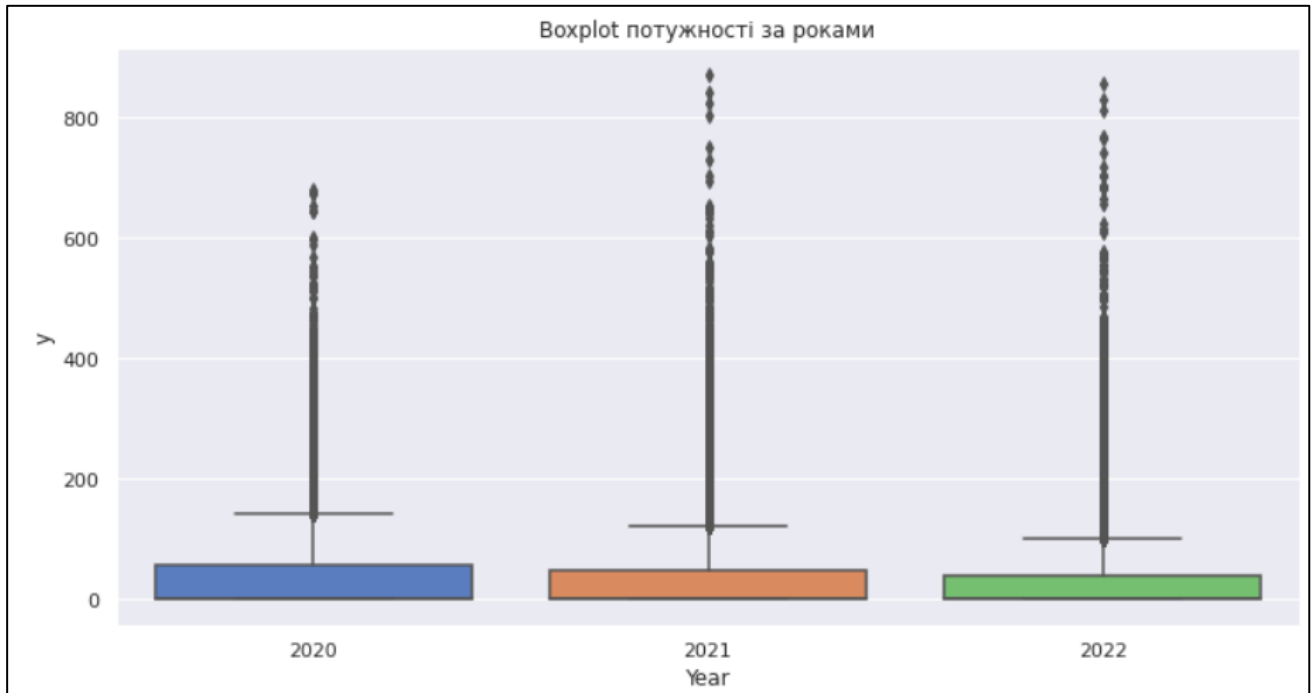


Рисунок 2.13 – Boxplot датасету генерації електроенергії ФЕС

З графіка можна зробити висновок, що генерація за три роки суттєво не відрізняється, а отже суттєвих змін які б могли впливати на прогнозування генерації немає.

Квантильно-квантильний (QQ) графік, або ймовірнісний графік, — це графічний засіб для порівняння змінної з певним теоретичним розподілом або для порівняння з розподілом іншої змінної. Одним із поширених застосувань графіка QQ є перевірка нормального розподілу змінної.

От же побудуємо QQ графік для датасету генерації електроенергії ФЕС (рис. 2.14).

```
sm.qqplot(df_data['y'], line='s')
plt.title('QQ-графік для часового ряду')
plt.show()
```

Рисунок 2.14 – Побудова QQ графіка датасету генерації електроенергії ФЕС

QQ графік датасету генерації електроенергії ФЕС наведена на рисунку 2.15.

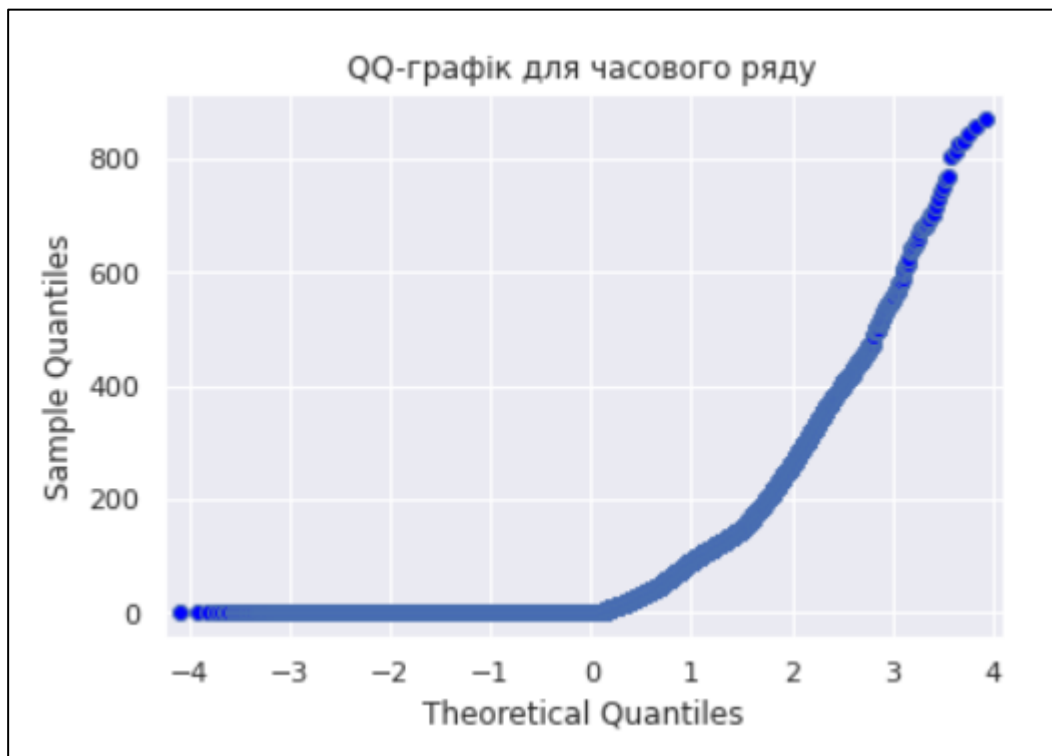


Рисунок 2.15 – QQ графік датасету генерації електроенергії ФЕС

Діаграма розсіювання — це графічне відображення відповідних даних, нанесених на графік з однією змінною на горизонтальній осі та іншою змінною на вертикальній осі. Дані зазвичай наносять на графік із вимірюваннями впливового параметра на горизонтальній осі (незалежна змінна) та мірами атрибута, який може реагувати на впливовий параметр на вертикальній осі (залежна змінна). Діаграми розсіювання є корисним першим кроком у будь-якому аналізі, оскільки вони допомагають візуалізувати взаємозв'язки та визначити можливі проблеми (наприклад, викиди), які можуть вплинути на подальший статистичний аналіз

Діаграма розсіювання для датасету генерації електроенергії ФЕС наведена на рисунку 2.16.

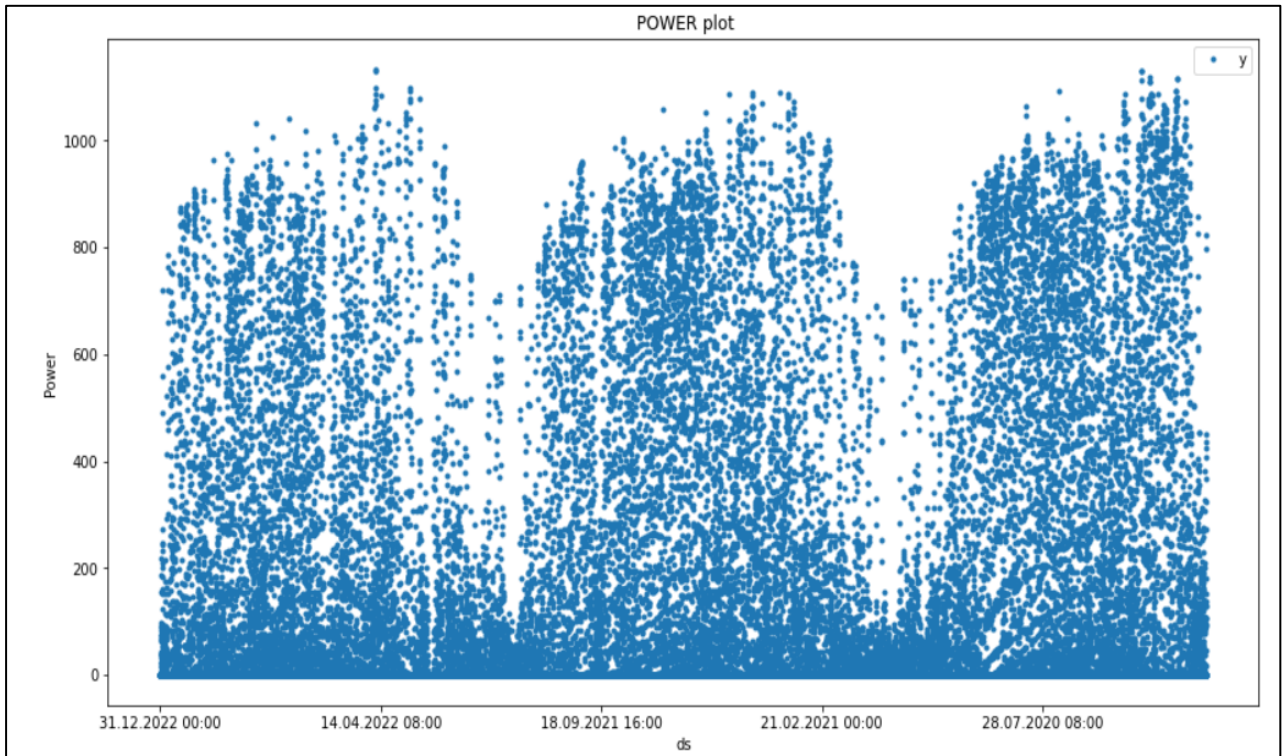


Рисунок 2.16 – Діаграма розсіювання датасету генерації електроенергії ФЕС

Як видно з діаграми розсіювання за період спостережень яскравих викидів не виявлено, проте присутні провали в генерації що варто врахувати при побудові моделі прогнозування

2.2 Огляд методів вирішення поставленої задачі

Генерацію електроенергії від ФЕС можна ефективно прогнозувати, використовуючи різні методи, такі як машинне навчання та аналіз часових рядів. Нижче розглянемо декілька підходів:

Машинне навчання:

Дерева рішень: Моделі на основі дерев рішень можуть аналізувати вплив різних факторів, таких як погодні умови, географічне розташування і технічні параметри станції.

Нейронні мережі: Глибоке навчання може бути використане для аналізу складних залежностей між вхідними факторами та генерацією електроенергії.

Аналіз часових рядів:

ARIMA (авторегресія та проінтегроване ковзне середнє): Модель ARIMA і її аналоги можуть бути ефективними для прогнозування генерації електроенергії, враховуючи тренди та сезонні зміни.

Однією з найпрогресивніших мов програмування для вирішення задач прогнозування є Python.

Python швидко став популярним інструментом для аналізу даних завдяки своїй простоті, універсальності та широкій підтримці спільноти. Ця мова програмування надає не тільки потужність для розв'язання складних завдань, але й вражає користувачів інтуїтивно зрозумілим синтаксисом та розгалуженою бібліотечною екосистемою. Інтуїтивно зрозумілий синтаксис Python робить його доступним для широкого кола користувачів, включаючи тих, хто тільки починає свій шлях у програмуванні. Він дозволяє легко виражати ідеї та концепції, що робить його ідеальним інструментом для аналізу даних, незалежно від рівня досвіду користувача. Широка бібліотечна екосистема Python грає ключову роль у його популярності серед аналітиків даних. Бібліотеки, такі як NumPy, Pandas, Matplotlib та Seaborn, надають потужні інструменти для обробки, аналізу та візуалізації даних. Ця різноманітність бібліотек робить Python ідеальним вибором для різноманітних завдань, від рутинного аналізу до складних моделювань. Незалежно від того, чи ви працюєте над програмою з великим обсягом даних чи співпрацюєте з досвідченим фахівцем із обробки даних, Python надає надійну та стабільну платформу для дослідження, візуалізації та моделювання часових рядів. Завдяки активній спільноті розробників і стеженню за сучасними тенденціями, Python залишається лідером у світі аналітики даних, забезпечуючи користувачам потужність та гнучкість у вирішенні різноманітних завдань.

Розглянемо, як Python може покращити роботу з даними часових рядів. Розглянемо наступний приклад фрагмента коду (рисунок 2.17), який завантажує набір даних часового ряду за допомогою pandas і малює його за допомогою Matplotlib (рис. 2.18)


```

import pandas as pd
import matplotlib.pyplot as plt
import numpy as np

# Generate random time-series data
np.random.seed(42)
dates = pd.date_range(start='2022-01-01', periods=100, freq='D')
values = np.random.randn(100).cumsum()

# Create a DataFrame from the generated data
data = pd.DataFrame({'date': dates, 'value': values})

# Set the 'date' column as the index
data.set_index('date', inplace=True)

# Plot the time-series data
plt.plot(data.index, data['value'])
plt.xlabel('Time')
plt.ylabel('Value')
plt.xticks(rotation = 45)
plt.title('Time Series Data')
plt.show()

```

Рисунок 2.17 – Приклад роботи з Python

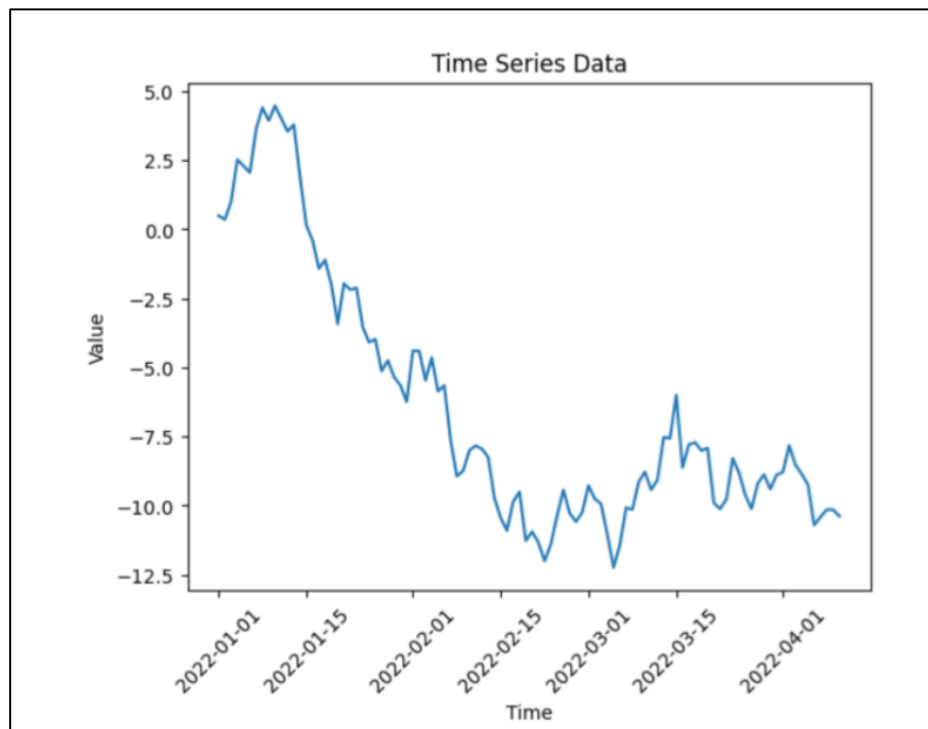


Рисунок 2.18 – Приклад побудови графіка в Python

У цьому прикладі випадковий набір даних часового ряду генерується за допомогою генератора випадкових чисел NumPy . Набір даних складається зі 100

дат, починаючи з 1 січня 2022 року, і відповідних випадкових значень. Потім дані перетворюються на Pandas DataFrame, а стовпець «дата» встановлюється як індекс. Нарешті, дані часових рядів будуються за допомогою Matplotlib, що відображає зміну «значення» з часом.

Прогнозування часових рядів є важливим аспектом аналізу даних, який допомагає підприємствам і організаціям приймати обґрунтовані рішення на основі минулих тенденцій і моделей. Точне передбачення майбутніх тенденцій і моделей у даних часових рядів є важливим для планування, прогнозування та прийняття рішень у багатьох галузях. Хоча існує кілька методів для прогнозування даних часових рядів, Prophet виділяється як популярний алгоритм завдяки своїй точності, гнучкості та простоті використання .

Чітка структура коду Python сприяє ефективним методам кодування, дозволяючи вам зосередитися на аналізі даних часових рядів, а не боротися зі складними концепціями програмування.

Однією з великих переваг Python є те, що це мова з відкритим кодом. Це означає, що він є у вільному доступі для використання та постійно вдосконалюється та підтримується активною спільнотою розробників. Природа Python із відкритим кодом дає змогу спеціалістам із обробки даних отримувати доступ до великої кількості ресурсів, інструментів і бібліотек для аналізу даних часових рядів без додаткових витрат.

Python пропонує велику колекцію спеціалізованих бібліотек і інструментів, спеціально розроблених для аналізу часових рядів. Ці бібліотеки, такі як pandas, NumPy, statsmodels і scikit-learn , надають різні функції та інструменти, адаптовані до унікальних завдань роботи з даними, що залежать від часу. Вони спрощують складні операції, дозволяючи вам зосередитися на отриманні значущої інформації.

Завдяки довговічності та широкому розповсюдженню Python має величезну кодову базу, яку науковці та розробники додатків можуть використовувати для аналізу часових рядів.

Багато поширених завдань, таких як завантаження даних, очищення, перетворення та візуалізація, уже реалізовано та спільно поширено спільнотою Python. Це дає змогу заощадити час і зусилля, спираючись на існуючий код і рішення, прискорюючи процес аналізу.

Python пропонує широкий спектр можливостей для обробки часових рядів, дозволяючи вам виконувати різноманітні завдання для аналізу та отримання інформації з ваших даних. У цьому розділі ми розглянемо деякі типові завдання обробки часових рядів, які можна легко виконати за допомогою Python.

Побудова часових рядів даних є важливим кроком у візуалізації закономірностей, тенденцій і аномалій. Python надає бібліотеку Matplotlib, яка включає модуль Pypplot для створення різних типів графіків, включаючи лінійні діаграми, точкові діаграми та гістограми.

В аналізі часових рядів доступні різні моделі прогнозування для майбутніх значень на основі історичних даних. Кожна модель має свої сильні сторони, обмеження та придатність для різних типів даних часових рядів. Давайте розглянемо деякі поширені типи моделей прогнозування:

Ковзне середнє (MA): модель ковзного середнього обчислює середнє значення минулих спостережень для прогнозування майбутніх значень. Це допомагає усунути короткострокові коливання та визначити базові тенденції в даних.

Просте ковзне середнє (SMA): просте ковзне середнє (SMA) – це форма ковзного середнього (MA), яка використовується для прогнозування часових рядів. Обчислюється шляхом взяття середнього арифметичного даного набору даних за певний період часу. Він займає ковзне вікно протягом заданого періоду часу. Його можна назвати рівнозваженим середнім n записів. Перевага використання SMA полягає в тому, що його легко обчислити та зрозуміти. Однак один недолік полягає в тому, що він базується на минулих даних і не враховує майбутні події. З цієї причини SMA слід використовувати не як єдиний метод прогнозування, а як один із інструментів у ширшому арсеналі прогнозів.

Експоненціальне ковзне середнє (ЕМА). Експоненціальне ковзне середнє (ЕМА) – це тип ковзного середнього, який надає більше ваги останнім точкам даних і допомагає згладити точки даних у часові ряди. На відміну від простих ковзних середніх, які надають однакову вагу всім точкам даних, ЕМА надають більшу вагу останнім точкам даних. Це робить його більш чутливим до нової інформації, ніж просте ковзне середнє. Часто вважається, що ЕМА краще реагують на зміни в базових даних. Існує кілька різних способів обчислення ЕМА, але найпоширенішим підходом є використання вагового коефіцієнта, який експоненціально зменшується з часом. Цей ваговий коефіцієнт може бути використаний для більшого або меншого акцентування останніх даних, залежно від потреб прогнозіста. Прогноз експоненціального ковзного середнього можна використовувати з даними будь-яких часових рядів, зокрема курсами акцій, економічними показниками чи даними про погоду.

Метод ковзного середнього широко використовується в прогнозуванні часових рядів через його гнучкість і простоту. На відміну від інших методів, таких як ARIMA або нейронні мережі, для цього не потрібні глибокі знання математики. Це означає, що навіть ті, хто має базові статистичні знання, можуть використовувати його для отримання надійних результатів.

Основна перевага методу ковзного середнього полягає в тому, що він враховує всі попередні значення при прогнозуванні майбутніх значень. Це допомагає зменшити вплив викидів під час прогнозування, а також полегшує ідентифікацію сезонних моделей у наборі даних часових рядів. Крім того, методологія зважування, яка використовується методом ковзного середнього, надає більше значення останнім значенням, ніж старим, що є корисним для прогнозування короткострокових тенденцій.

Крім того, метод простого ковзного середнього (SMA) зазвичай обчислювально швидший, ніж більш складні методи, такі як експоненціальне ковзне середнє (ЕМА). Він також вимагає менше параметрів і може використовуватися на коротших наборах даних. І, нарешті, ефективність методу

SMA доведена в багатьох додатках, таких як аналіз фондового ринку, а також сезонне прогнозування.

Загалом метод ковзної середньої є ефективним інструментом для короткострокового прогнозування завдяки своїй гнучкості та простоті використання. Його здатність враховувати всі минулі значення під час створення прогнозів забезпечує точність, а його здатність визначати сезонні закономірності означає, що його можна також ефективно використовувати для довгострокового прогнозування. Крім того, його обчислювальна швидкість і мінімальні параметри роблять його популярним вибором для багатьох програм.

Авторегресія (AR): авторегресійна модель використовує минулі спостереження та рівняння лінійної регресії для прогнозування майбутніх значень. Він передбачає, що майбутні значення залежать від попередніх значень із запізненням.

Перевага цієї моделі полягає в тому, що за допомогою функції автокореляції можна визначити, чи немає випадковості. Крім того, модель здатна прогнозувати повторювані закономірності в даних. Також можна передбачити результати з меншою кількістю інформації, використовуючи ряди самозмінних.

Проте модель має кілька обмежень. Коефіцієнт автокореляції в цьому випадку повинен бути не менше 0,5, щоб він був відповідним. Це означає, що якщо він менше 0,5, результат передбачення буде неточним.

Зазвичай модель AR використовується під час прогнозування речей, пов'язаних з економікою, на основі історичних даних. Щось, на що істотно впливають соціальні фактори. Рекомендується замість цього використовувати модель векторної авторегресії. Причина полягає в тому, що одну модель можна використовувати для прогнозування кількох змінних часових рядів одночасно.

Авторегресійне ковзне середнє (ARMA): Модель ARMA поєднує моделі авторегресії та ковзного середнього, враховуючи як минулі спостереження, так і середнє значення минулих помилок для прогнозування. Частина AR передбачає моделювання поточного значення ряду як лінійної комбінації попередніх значень.

Порядок p визначає, скільки попередніх значень використовується. Частина MA моделює поточне значення ряду як лінійну комбінацію минулих помилок (білий шум). Порядок q визначає, скільки включено термінів помилки з відстрочкою.

Модель ARMA припускає, що часовий ряд є стаціонарним, тобто статистичні властивості, такі як середнє значення та дисперсія, є постійними протягом часу. Нестационарні дані часто можна перетворити на стаціонарні за допомогою різницевих або інших перетворень. Передбачається, що терміни помилки є білим шумом, тобто вони некорельовані та мають постійне середнє значення та дисперсію.

Модель ARMA широко використовується в економіці, фінансах, інженерії та інших сферах для:

- прогнозування майбутніх значень часового ряду;
- розуміння базових шаблонів і структур у даних;
- фільтрування або згладжування шумових даних.

Модель ARMA не може безпосередньо обробляти сезонність, і дані повинні бути сезонно скориговані, якщо це є особливістю часових рядів. Вибір належного порядку для частин AR і MA (p та q) вимагає досвіду та може включати такі методи, як інформаційний критерій Akaike (AIC) або байєсівський інформаційний критерій (BIC).

Загалом модель ARMA є гнучким і широко використовуваним інструментом для аналізу часових рядів, що поєднує в собі сильні сторони моделей авторегресії та моделі ковзного середнього. Це вимагає ретельного розгляду основних припущень і відповідного вибору параметрів для оптимальної продуктивності.

Авторегресійне інтегроване ковзне середнє (ARIMA): Модель ARIMA розширює модель ARMA шляхом включення різниці, щоб зробити часовий ряд стаціонарним.

ARIMA поєднує функції авторегресії з функціями ковзних середніх. Наприклад, авторегресійний процес AR(1) — це процес, у якому поточне значення

базується на безпосередньо попередньому значенні, тоді як процес $AR(2)$ — це процес, у якому поточне значення базується на двох попередніх значеннях. Ковзне середнє — це обчислення, яке використовується для аналізу точок даних шляхом створення серії середніх різних підмножин повного набору даних, щоб згладити вплив викидів. Завдяки цій комбінації методів моделі ARIMA можуть враховувати тренди, цикли, сезонність та інші нестатичні типи даних під час прогнозування. Прогнозування ARIMA досягається шляхом підключення даних часових рядів. Статистичне програмне забезпечення визначить відповідну кількість затримок або величину різниці, яка буде застосована до даних, і перевірить їх на стаціонарність. Потім він виведе результати, які часто інтерпретуються подібно до моделі множинної лінійної регресії. Модель ARIMA використовується як інструмент прогнозування, щоб передбачити, як щось буде діяти в майбутньому на основі минулих показників. Він використовується в технічному аналізі для прогнозування майбутніх показників активу. Дана модель підходить для нестационарних даних з тенденціями та сезонністю. Моделювання ARIMA, як правило, не підходить для довгострокових прогнозів, наприклад, на більш ніж шість місяців вперед, оскільки воно використовує минулі дані та параметри, на які впливає людське мислення. З цієї причини його найкраще використовувати з іншими інструментами технічного аналізу, щоб отримати чіткішу картину ефективності активу.

Експоненціальне згладжування: моделі експоненційного згладжування застосовують ваги до минулих спостережень, надаючи більшого значення останнім значенням. Різні варіації експоненціального згладжування, такі як просте експоненціальне згладжування (SES), лінійне експоненціальне згладжування Холта та експоненціальне згладжування Холта-Вінтерса, враховують різні моделі в даних.

Існує кілька типів методів експоненціального згладжування, включаючи просте експоненціальне згладжування (SES), подвійне експоненціальне

згладжування (DES) і потрійне експоненціальне згладжування (TES), також відоме як метод Холта-Уінтерса. Розглянемо кожну з них докладніше.

Просте експоненціальне згладжування (SES): SES є найпростішою формою експоненційного згладжування. Для створення наступного прогнозу він використовує лише поточне спостереження та прогноз за попередній період.

Подвійне експоненціальне згладжування (DES): DES розширює просте експоненціальне згладжування, об'єднуючи компонент тренду та компонент рівня. Він підходить для даних часових рядів, які демонструють тенденцію, тобто систематичну зміну значень часових рядів з часом.

Потрійне експоненціальне згладжування (TES): TES розширює подвійне експоненціальне згладжування, додаючи сезонний компонент на додаток до компонентів рівня та тенденції. Він підходить для даних часових рядів, які виявляють сезонність, тобто повторюваний шаблон у значеннях часових рядів протягом певного часу.

Експоненціальне згладжування має кілька переваг як метод прогнозування. По-перше, він простий у реалізації та обчислювально ефективний, що робить його придатним для великих наборів даних і програм прогнозування в реальному часі. По-друге, він є адаптивним, оскільки оновлює прогнози на основі останніх даних, завдяки чому реагує на зміни в часових рядах. По-третє, він не потребує великої кількості точок історичних даних, оскільки він призначає більш високу вагу останнім спостереженням, що робить його корисним для прогнозування в ситуаціях, коли доступність даних обмежена.

Вибір того, який тип методу експоненційного згладжування використовувати, залежить від характеристик даних часового ряду та конкретних вимог до прогнозування.

Просте експоненціальне згладжування (SES): SES підходить для даних часових рядів, які не мають тенденції чи сезонності та демонструють стабільний рівень у часі. Це основний метод, який можна використовувати, коли дані не показують жодних систематичних закономірностей або порушень, і бажано

простий прогноз, заснований на останньому спостереженні та попередньому прогнозі. SES є обчислювально ефективним і простим у застосуванні, що робить його придатним для прогнозування в реальному часі або ситуацій, коли доступність даних обмежена.

Подвійне експоненціальне згладжування (DES): DES підходить для даних часових рядів, які демонструють тенденцію, тобто систематичну зміну значень часових рядів з часом. Це розширення SES, яке включає компонент тренду на додаток до компонента рівня, що дозволяє фіксувати шаблони трендів у даних. DES є корисним, коли дані показують постійну тенденцію до зростання або зниження, і потрібен прогноз, який враховує як поточний рівень, так і тенденцію. DES також підходить, коли дані не мають сезонності, але демонструють тенденцію.

Потрійне експоненціальне згладжування (TES): TES підходить для даних часових рядів, які демонструють тенденцію та сезонність. Це розширення DES, яке включає сезонний компонент на додаток до компонентів рівня та тенденції, що дозволяє фіксувати сезонні моделі в даних. TES є корисним, коли дані показують як тенденцію, так і повторювану закономірність у значеннях у часі, наприклад місячну чи квартальну сезонність. TES може надавати прогнози, які враховують поточний рівень, тренд і сезонність даних, що робить його придатним для прогнозування в ситуаціях, коли присутні як тренд, так і сезонність.

Важливо зауважити, що вибір методу експоненціального згладжування, який використовувати, також може залежати від конкретних вимог до прогнозування, таких як бажаний горизонт прогнозу, необхідний рівень точності та наявність історичних даних. Можливо, знадобиться поекспериментувати з різними методами та налаштувати параметри згладжування (α , β та γ), щоб знайти найкращий варіант для конкретних даних часового ряду та цілей прогнозування. Крім того, перед прийняттям будь-яких остаточних рішень завжди рекомендується оцінювати точність прогнозів за допомогою відповідних

показників ефективності та перевіряти прогнози за допомогою даних поза вибіркою.

Однак експоненціальне згладжування також має деякі обмеження. Він може не працювати належним чином для даних часових рядів зі складними закономірностями або нерівномірностями, такими як раптові зміни рівня чи тенденції, викиди або різкі зміни в сезонності. У таких випадках інші передові методи прогнозування можуть бути більш доречними. Крім того, вибір параметрів згладжування (α , β і γ) може вплинути на точність прогнозів, і пошук оптимальних значень цих параметрів може вимагати експериментів або методів вибору моделі.

Сезонний ARIMA (SARIMA): SARIMA є розширенням моделі ARIMA, яка враховує сезонні закономірності в даних. Він містить додаткові терміни для врахування сезонності, що робить його придатним для даних часових рядів із повторюваними моделями.

Для подальших досліджень детальніше розглянемо бібліотеку Prophet [113].

Prophet — це алгоритм прогнозування, розроблений командою Facebook у 2017 році. Алгоритм розроблений як масштабований, швидкий і точний, що робить його придатним для широкого спектру застосувань, від прогнозування продажів в електронній комерції до прогнозування погодних умов.

Модель базується на підлаштуванні адитивних регресійних моделей (Generalized Additive Models, GAM) до вигляду:

$$y(t) = g(t) + s(t) + h(t) + \epsilon t, \quad (2.2)$$

де, $g(t)$ – тренд функція,

$s(t)$ – сезонні коливання,

$h(t)$ – функція, що відображає ефекти святкових днів та інших впливових подій,

ϵt – нормально розподілені випадкові збурення.

Моделювання тренду реалізовано на основі кусково-лінійної регресії. Стандартна логістична функція для опису росту в системі з обмеженою ємкістю має вигляд:

$$g(t) = C / (1 + \exp(-k(t-m))), \quad (2.13)$$

де C - ємкість системи,

k - швидкість росту,

t - час,

а m - параметр, що рухає функцію вздовж осі часу.

Однак дана модель не враховує декілька факторів. По-перше ємкість багатьох систем не постійна, наприклад кількість людей що мають доступ до Інтернету, з часом росте. Тому в пакеті prophet ємкість системи C замінена на динамічну $C(t)$. По-друге, не постійною зазвичай буває і швидкість росту k . Для цього в даній моделі водиться поняття точок зламу тренду. Припустимо, що S таких точок знаходяться на часових мітках $s_j, j = 1, 2, \dots, S$. Сукупність усіх змін швидкості росту δ_j можна представити у вигляді вектора $\delta \in \mathbb{R}^S$. Тоді швидкість росту в будь якій точці часу t буде рівнятися сумі базової швидкості k та всіх змін, що передували цій точці: $k + \sum_{j:t > s_j} \delta_j$. Це можна продемонструвати за допомогою вектора $a(t) \in \{0, 1\}^S$, або:

$$a_j(t) = \begin{cases} 1, & \text{if } t \geq s_j \\ 0, & \text{otherwise} \end{cases}, \quad (2.3)$$

В результаті швидкість росту в момент часу t складатиме $k + a(t) T \delta$. При зміні швидкості росту необхідно змінити параметр m , з метою забезпечення гладкого з'єднання сегментів кривої тренду на часовому відрізку (формула 2.4).

$$y_j = (s_j - m - \sum_{l < j} y_l) \left(1 - \frac{\sum_{l < j} \delta_l}{\sum_{l \leq j} \delta_l}\right), \quad (2.4)$$

В результаті функція, що використовується в пакеті Prophet матиме вигляд:

$$g(t) = \frac{c(t)}{1 + \exp(-(k + a(t)T\delta)(t - (m + a(t)T\gamma)))} \quad (2.5)$$

При необмеженому рості системи замість представленої вище функції, в моделі Prophet використовується кусково-лінійна функція:

$$g(t) = (k + a(t)T\delta)t + (m + a(t)T\gamma) \quad (2.6)$$

Річна та тижнева сезонність моделюється за допомогою рядів Фур'є:

$$s(t) = \sum_{n=1}^N (a_n \cos(\frac{2\pi n t}{P}) + b_n \sin(\frac{2\pi n t}{P})) \quad N, n = 1 \quad (2.7)$$

де, для річної сезонності $P=365.25$, $N=10$;

для тижневої – $P=7$, $N=3$;

N – визначає гладкість функції.

Основна ідея Prophet полягає в моделюванні даних часових рядів як комбінації компонентів тенденції, сезонності та шуму. Розкладаючи дані на ці компоненти, алгоритм може створювати точні прогнози, які фіксують базові закономірності в даних.

Компонент **тенденції** фіксує загальний напрямок часового ряду, незалежно від того, зростає він чи зменшується з часом. Цей компонент моделюється за допомогою моделі кусково-лінійної регресії, що забезпечує гнучкість у підгонці тенденції до даних.

Компонент **сезонності** фіксує періодичні моделі в даних, наприклад щотижневі чи місячні тренди. Цей компонент моделюється за допомогою ряду Фур'є, що дозволяє гнучко моделювати різні сезонні моделі.

Компонент **шуму** фіксує випадкові коливання даних, які не можна пояснити компонентами тенденції чи сезонності.

Prophet використовує байєсівську структуру для моделювання даних часових рядів. Це означає, що алгоритм оцінює апостеріорний розподіл параметрів моделі, а не лише точкові оцінки. Таким чином, алгоритм може генерувати імовірнісні прогнози, які забезпечують міру невизначеності навколо точкового прогнозу.

Prophet використовує комбінацію регресійних моделей і байєсівського висновку для моделювання даних часових рядів. Математична база, що лежить в основі алгоритму, поділяється на три компоненти: моделювання трендів, моделювання сезонності та байєсівський висновок.

Моделювання трендів

Трендовий компонент часового ряду моделюється за допомогою моделі кусково-лінійної регресії. Ця модель передбачає, що тенденцію можна представити як послідовність лінійних сегментів, де нахил кожного сегмента є функцією точок зміни в даних. Точки зміни – це точки в часовому ряді, де змінюється напрям тренда.

Математично модель тренду можна записати так:

$$y(t) = g(t) + e(t). \quad (2.8)$$

де $y(t)$ - це значення часового ряду в момент часу t ;

$g(t)$ - компонент тренду;

а $e(t)$ - це помилка.

Компонент тренду моделюється як кусково-лінійна функція, яку можна записати так:

$$g(t) = (k(t) * t) + m(t). \quad (2.9)$$

де $k(t)$ – нахил тренду в момент часу t ;

$m(t)$ – точка перетину тренду в момент часу t .

Нахил і перетин моделюються за допомогою ієрархічної байєсівської моделі, яка дозволяє систематизувати оцінки та фіксує невизначеність навколо оцінок.

Моделювання сезонності

Сезонна складова часового ряду моделюється за допомогою ряду Фур'є. Цей підхід дозволяє гнучко моделювати різні типи сезонних моделей, включаючи тижневі, місячні та річні тренди. Ряд Фур'є можна записати так:

$$S(t) = \sum_{i=1}^N [a(i) * \cos(2\pi i t/P) + b(i) * \sin(2\pi i t/P)]. \quad (2.10)$$

де $S(t)$ – сезонна складова в момент часу t ;

$a(i)$ і $b(i)$ – коефіцієнти Фур'є;

N – кількість членів Фур'є;

P – період сезонної складової.

Коефіцієнти Фур'є моделюються за допомогою ієрархічної байєсівської моделі, яка дозволяє регуляризувати оцінки та враховує невизначеність навколо оцінок.

Байєсівський висновок

Prophet використовує байєсівський висновок для оцінки апостеріорного розподілу параметрів моделі. Це робиться за допомогою алгоритму ланцюга Маркова Монте-Карло, який бере вибірку з апостеріорного розподілу параметрів моделі. Алгоритм оцінює апостеріорний розподіл параметрів моделі за допомогою такої функції правдоподібності:

$$p(y|\theta) = \prod_{t=1}^T N(y(t)|g(t), \sigma). \quad (2.11)$$

де $y(t)$ – спостережене значення часового ряду в момент часу t ;

$g(t)$ – прогнозоване значення часового ряду в момент часу t ;

σ – параметр шуму;

θ – набір параметрів моделі (включаючи компоненти тенденції та сезонності).

Алгоритм використовує функцію ймовірності для обчислення апостеріорного розподілу параметрів моделі за допомогою теореми Байєса:

$$p(\theta|y) = p(y|\theta) * p(\theta) / p(y). \quad (2.12)$$

де $p(\theta|y)$ – апостеріорний розподіл параметрів моделі;

$p(y|\theta)$ – функція ймовірності;

$p(\theta)$ – попередній розподіл параметрів моделі;

$p(y)$ – гранична ймовірність.

Prophet з додатковими регресорами

Prophet також дозволяє включати в модель додаткові регресори. Регресори – це зовнішні змінні, які можуть впливати на часові ряди, наприклад свята, погодні умови чи маркетингові кампанії. Включення регресорів у модель може підвищити точність прогнозів шляхом фіксації впливу цих зовнішніх змінних на часовий ряд.

Щоб включити регресори в модель, потрібно вказати матрицю регресорів і коефіцієнти регресорів. Матриця регресора — це матриця $T \times N$, де T — кількість моментів часу, а N — кількість регресорів. Кожен стовпець матриці представляє різну змінну регресора. Коефіцієнти регресора оцінюються за допомогою моделі лінійної регресії, яка пов'язує часовий ряд із матрицею регресора. Остаточну модель можна записати так:

$$y(t) = g(t) + S(t) + \sum_{j=1}^N X(j,t) * \beta(j) + e(t). \quad (2.13)$$

де $X(j,t)$ — j -та змінна регресора в момент часу t ,

$\beta(j)$ — коефіцієнт j -ї змінної регресора,

$e(t)$ — помилка.

Як аналог Prophet можна використовувати алгоритм прогнозування Amazon SageMaker DeepAR — це контрольований алгоритм навчання для прогнозування скалярних (одновимірних) часових рядів за допомогою рекурентних нейронних мереж (RNN). Класичні методи прогнозування, такі як авторегресійне інтегроване ковзне середнє (ARIMA) або експоненціальне згладжування (ETS), підбирають одну модель для кожного окремого часового ряду. Потім вони використовують цю модель для екстраполяції часових рядів у майбутнє.

Однак у багатьох програмах існує багато подібних часових рядів у наборі одиниць поперечного перерізу. Наприклад, можна групувати часові ряди для попиту на різні продукти, навантаження на сервер і запити на веб-сторінки. Для цього типу застосування можна отримати вигоду від спільного навчання однієї моделі для всіх часових рядів. DeepAR використовує цей підхід. Коли набір даних містить сотні пов'язаних часових рядів, DeepAR перевершує стандартні методи ARIMA та ETS. Також можна використовувати навчену модель для створення прогнозів для нових часових рядів, подібних до тих, на яких було навчено.

Вхідними даними для навчання алгоритму DeepAR є один або, бажано, кілька target часових рядів, які були згенеровано тим самим процесом або подібними процесами. На основі цього вхідного набору даних алгоритм навчає модель, яка вивчає наближення цього процесу/процесів і використовує її для прогнозування розвитку цільового часового ряду. Кожен цільовий часовий ряд можна необов'язково пов'язати з вектором статичних (незалежних від часу) категоріальних ознак, наданих полем, cat1 вектором динамічних (залежних від часу) часових рядів, наданих полем dynamic_feat. SageMaker навчає модель DeepAR шляхом випадкової вибірки навчальних прикладів із кожного цільового часового ряду в навчальному наборі даних. Кожен навчальний приклад складається з пари суміжних вікон контексту та передбачення з фіксованою попередньо визначеною довжиною. Щоб контролювати, наскільки далеко в минулому може бачити мережа, використовується гіперпараметр context_length.

Щоб контролювати, наскільки далеко в майбутньому можна робити прогнози, використовується гіперпараметр `prediction_length`.

2.3 Побудова UML діаграми роботи інформаційної системи прогнозування генерації електроенергії

З метою відображення та моделювання поведінки алгоритму, розроблено UML діаграму станів. Діаграма відображає послідовність кроків, по яким об'єкт проходить протягом свого циклу життя. Діаграма складається областей, кожна з яких містить граф з набором вершин, з'єднаних зв'язками, що являють собою переходи.

UML діаграма станів роботи інформаційної системи прогнозування генерації електроенергії наведено на рисунку 2.19.

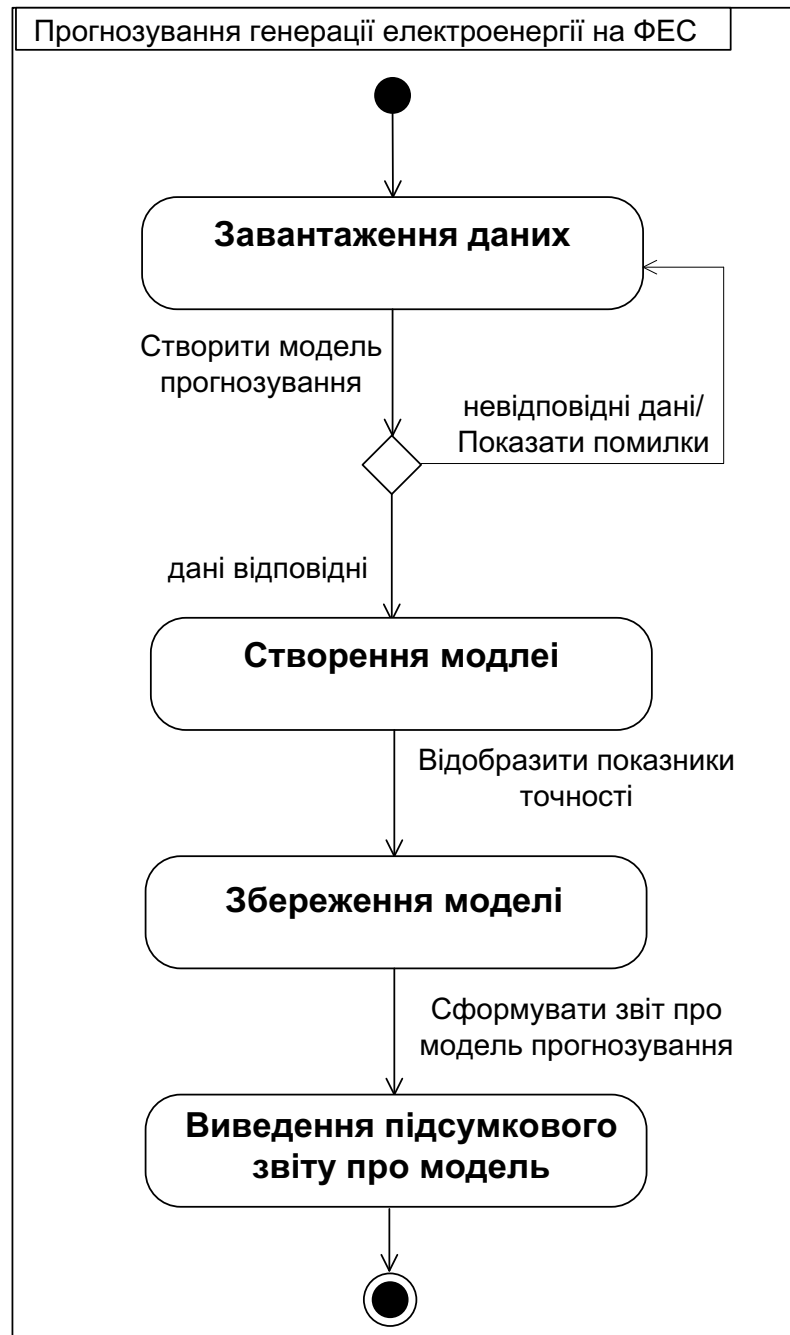


Рисунок 2.19 – UML діаграма станів роботи інформаційної системи прогнозування генерації електроенергії

Діаграма відображає три основні стани, які пов'язані між собою зв'язками.

Стан «Завантаження даних» відтворює випадок коли відпрацьовують підстави обробки даних. Оскільки в процес підготовки даних використовувалася підпрограма конвертації датасету, тому варто стан розглянути детальніше. Після перевірки на відповідність даних відбувається перехід в наступний стан

«Створення моделі». В цьому стані система створює необхідні налаштування моделі, проводить навчання на тренувальних даних, перевірку на тестових та валідаційних даних. Розраховуються похибки прогнозування. Далі здійснюється перехід в стан «Збереження моделі», де відбувається власне збереження моделі прогнозування. Після чого здійснюється формування підсумкового звіту про модель прогнозування стан «Виведення підсумкового звіту про модель».

Розглянемо детальніше стан «Завантаження даних» (рис.2.20).

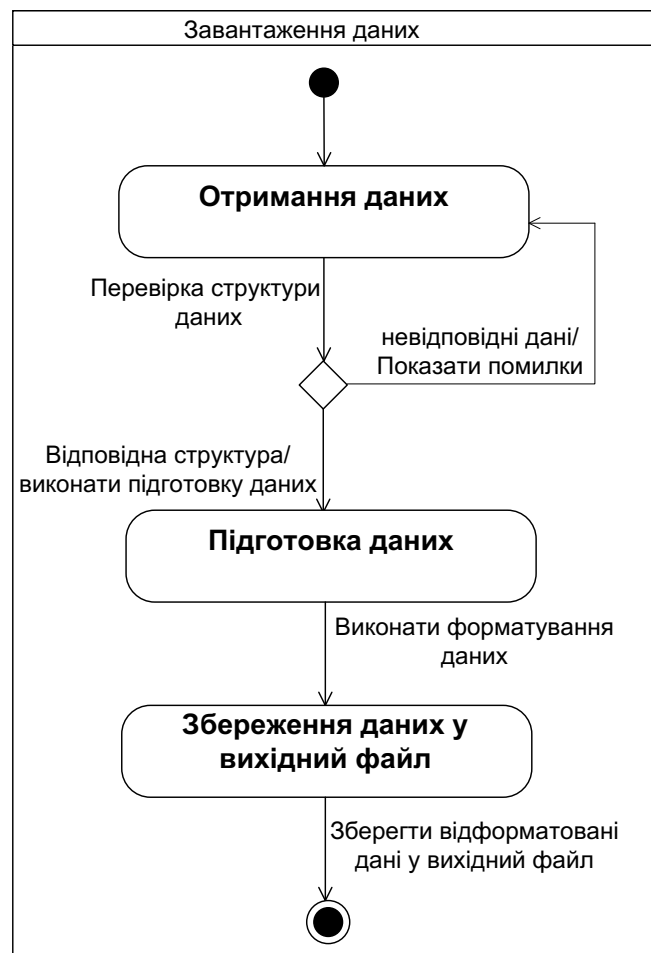


Рисунок 2.20 – UML діаграма підстанів «Завантаження даних» інформаційної системи прогнозування генерації електроенергії

UML діаграма підстанів «Завантаження даних» містить підстани:

– «Отримання даних» – в підстані здійснюється завантаження початкових даних та перевірка їх на відповідність.

– «Підготовка даних» – в підстані здійснюється форматування даних, видалення зайвих колонок, сортування у порядку зростання дати.

– «Збереження даних» – в підстані здійснюється збереження даних, запис їх у вихідний файл.

Розглянемо детальніше стан «Створення моделі» (рис. 2.21).



Рисунок 2.21 – Діаграми складених станів алгоритму прогнозування генерації електроенергії

UML діаграма підстанів «Створення моделі» містить підстани:

– «Отримання підготовлених даних» – в підстані здійснюється завантаження підготовлених даних.

– «Розділення даних» – в підстані здійснюється розділення даних на навчальний, тренувальний та валідаційний набори.

– «Отримання вхідних параметрів» – в підстані здійснюється налаштування моделі прогнозування.

– «Створення моделі» – в підстані здійснюється безпосередньо створення моделі прогнозування.

– «Прогнозування генерації електроенергії» – в підстані здійснюється прогнозування генерації електроенергії на тестовому наборі даних.

– «Обчислення показників точності» – в підстані здійснюється обчислення похибки прогнозування на валідаційному наборі даних.

2.4 Висновки

В даному розділі було проведено огляд вхідних даних щодо генерації електроенергії на ФЕС. Здійснено аналіз датасету та його форматування для забезпечення зручності подальших досліджень. Для оцінки датасету використані статистичні показники, що вказують на особливості генерації електроенергії на ФЕС.

Середнє значення генерації електроенергії за розглянутий період становить 151.24, вказуючи на узагальнене значення. Стандартне відхилення, що дорівнює 260.61, свідчить про значну варіацію в генерації, можливо, пов'язану з різними факторами, такими як погодні умови чи технічні аспекти станції. Медіана, рівна 0.0, вказує на наявність значень, що сконцентровані навколо нуля, що може пояснюватися наявністю багатьох нульових або дуже малих значень у ряді.

Використовуючи автокореляцію, було виявлено виражену сезонність у генерації електроенергії, обумовлену зміною світлового дня залежно від пори року. Графік автокореляції вказує на наявність регулярних циклічних коливань, що є характерним для сонячних електростанцій.

Для візуалізації розподілу даних використано гістограму та boxplot, які свідчать про відсутність суттєвих змін у генерації протягом трьох років. Діаграма розсіювання також не виявила яскравих викидів, проте присутні провали в генерації, що слід врахувати при побудові моделі прогнозування.

Результати розвідувального аналізу свідчать про стабільність генерації електроенергії на ФЕС з можливістю врахування сезонних коливань та провалів у деякі періоди.

Проведено аналіз можливостей роботи з мовою програмування Python для вирішення задач прогнозування та нової технології прогнозування Prophet.

3 РОЗРОБЛЕННЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ ПРОГНОЗУВАННЯ ГЕНЕРАЦІЇ ЕЛЕКТРОЕНЕРГІЇ

3.1 Вибір платформи для побудови моделі прогнозування

В сучасному світі розвитку інформаційних технологій та аналізу даних існує широкий спектр інструментів та середовищ, які дозволяють вченим даних, аналітикам та розробникам працювати ефективно та здійснювати розвідку в області науки про дані та машинного навчання. Найпопулярніших з них Kaggle та Anaconda.

Kaggle – це онлайн-платформа для проведення змагань у галузі аналізу даних, машинного навчання та штучного інтелекту. Особливістю Kaggle є те, що вона зближує глобальну спільноту вчених даних, створюючи унікальне середовище для спільної роботи, обміну знаннями та вирішення реальних завдань у сферах від науки про дані до розробки штучного інтелекту.

Функції та Можливості

Змагання (Competitions): Kaggle надає можливість взяти участь у конкурсах, де учасники можуть використовувати свої навички для розв'язання реальних проблем.

Набори Даних (Datasets): Платформа містить різноманітні набори даних, які можна використовувати для навчання та експериментів.

Кодові Середовища (Kernels): Учасники можуть ділитися своїм кодом та отримувати фідбек від спільноти через інтерактивні блокноти Kaggle.

Форуми та Спільнота (Forums and Community): Групи обговорення дозволяють обмінюватися досвідом та рішеннями проблем.

Anaconda - це відкрите та безкоштовне дистрибутивне середовище для мов програмування Python та R. Воно призначене для обробки та аналізу даних, наукових обчислень, розробки програм та машинного навчання. Основна перевага Anaconda полягає в тому, що воно включає в себе багато популярних пакетів та

бібліотек для роботи з даними, що полегшує встановлення та управління потрібними інструментами.

Основні риси та компоненти Anaconda:

Conda: Це власний менеджер пакетів Anaconda, який дозволяє легко встановлювати, оновлювати та видаляти пакети та їх залежності.

Jupyter Notebooks: Інтерактивні блокноти, які дозволяють об'єднувати код, тексти та графіку в одному документі, що робить роботу з даними більш інтерактивною та зрозумілою.

Spyder: Інтегроване середовище розробки (IDE) для Python, яке надає зручні інструменти для написання, тестування та відлагодження коду.

Pandas, NumPy, Matplotlib, Seaborn і багато інших бібліотек: Включені в Anaconda популярні бібліотеки для роботи з даними, візуалізації та математичних обчислень.

Середовища: З можливістю створення віртуальних середовищ для ізоляції проектів та уникнення конфліктів залежностей.

SciKit-Learn, TensorFlow, PyTorch та інші бібліотеки для машинного навчання: Популярні інструменти для розробки та навчання моделей машинного навчання.

Широкий вибір пакетів: Anaconda містить більше 1500 пакетів, що полегшує використання різноманітних інструментів та бібліотек в області науки про дані та розробки.

Anaconda дозволяє швидко налаштувати середовище для роботи з Python та зосереджуватися на вирішенні задач з аналізу даних, наукових досліджень та машинного навчання.

Поза Kaggle та Anaconda існують інші інструменти та платформи:

Colab (Google Colab): Забезпечує безкоштовний доступ до ресурсів Google для виконання коду на Python. Має підтримку обладнання GPU для прискорення обчислень.

DataCamp: Онлайн-платформа для навчання аналізу даних та програмування, яка також надає інтерактивні вправи для практики.

IBM Watson Studio: Платформа для розробки та виконання проєктів із штучного інтелекту та аналізу даних в середовищі IBM.

Azure Notebooks (Microsoft Azure): Microsoft Azure дозволяє створювати та запускати інтерактивні блокноти у хмарному середовищі.

Ці платформи є лише кількома прикладами інтерпретаторів та інструментів для аналізу даних та машинного навчання, які дозволяють вченим даних та розробникам ефективно працювати над проєктами та навчанням.

Для проведення подальших досліджень було обрано платформу Kaggle.

3.2 Прогнозування генерації електроенергії на ФЕС

Для ефективного старту роботи у нашому проєкті, важливо вчасно та правильно виконати підключення необхідних бібліотек, які будуть використовуватися у подальших етапах розробки. Цей процес є важливим етапом перед початком написання коду, оскільки він визначає набір інструментів та функцій, які будуть доступні для використання в подальших програмах. Правильно підібрані бібліотеки не лише полегшують написання коду, а й роблять його більш структурованим та ефективним. Такий підхід допомагає уникнути зайвого коду та сприяє легшій розробці та обслуговуванню проєкту. Тому вибір та підключення бібліотек - це перший крок до успішної та продуктивної роботи у проєкті.

У попередньому розділі було встановлено, що генерація електроенергії на ФЕС має добову а також сезонну циклічність, для побудови моделі прогнозування було обрано ноутбук [114], в якому автор хоч і досліджує прогнозування захворюваності на COVID-19, проте для побудови моделі автор використовує бібліотеку Prophet, яка гарно підходить для вирішення задачі прогнозування.

Код для підключення бібліотек наведено на рисунку 3.1.

```
# Work with Data - the main Python libraries
import numpy as np
import pandas as pd

# For import data from API
import requests

# Visualization
import matplotlib.pyplot as plt

# Modeling and Prediction
from fbprophet import Prophet
from sklearn.metrics import mean_absolute_error, r2_score

import warnings
warnings.simplefilter('ignore')
```

Рисунок 3.1 – Підключення бібліотек

Для дослідження використовується бібліотека Prophet. Ця бібліотека визначається високою ефективністю та точністю при аналізі даних, особливо в тих випадках, коли присутні сезонні зміни. Вона є інструментом вибору для вирішення задач, пов'язаних із часовими рядами, та володіє великим набором функцій, які дозволяють здійснювати точний та надійний аналіз даних. Prophet є особливо корисним при роботі з даними, де присутні сезонні коливання, оскільки вона здатна адаптуватися до цих змін і надає точний прогноз.

Для початку скачаємо потрібні нам дані. Даний код зображено на рисунку 3.2.

```
# Load the data from the local CSV file
df_data = pd.read_csv('../input/energy/energy_data_set.csv', sep=';')

# Print the dataframe
print(df_data)
```

Рисунок 3.2 – Завантаження даних генерації електричної енергії на ФЕС

Результатом даного коду являється таблиця з статистичними даними де `ds` – дата, `y` – значення потужності в момент вимірювання.

Дана таблиця зображена на рисунку 3.3.

		<code>ds</code>	<code>y</code>
0	31.12.2022	00:00	0
1	31.12.2022	00:30	0
2	31.12.2022	01:00	0
3	31.12.2022	01:30	0
4	31.12.2022	02:00	0
...	
47419	22.02.2020	21:30	0
47420	22.02.2020	22:00	0
47421	22.02.2020	22:30	0
47422	22.02.2020	23:00	0
47423	22.02.2020	23:30	0

Рисунок 3.3– Статистичні дані

Після успішного завантаження даних, наступним важливим етапом в нашому процесі обробки даних буде фільтрація пропущених значень. Це важлива частина підготовки даних, спрямована на забезпечення точності та повноти нашого аналізу.

Пропущені значення можуть виникнути з різних причин, таких як технічні аспекти, помилки вводу, або непередбачувані обставини. Вони можуть впливати на надійність наших аналітичних результатів та прогнозів, тому важливо вчасно їх виявляти та обробляти.

Для фільтрації пропущених значень ми можемо використовувати різноманітні методи, такі як заповнення значеннями за замовчуванням, інтерполяція або видалення рядків з пропущеними даними. Вибір конкретного методу буде залежати від конкретного контексту нашого дослідження та від того, яким чином пропущені значення можуть впливати на наші висновки.

Реалізація фільтрації пропущених значень наведена на рисунку 3.4.

```
# Filtering the missing data
data = df_data[['ds', 'y']].dropna().reset_index(drop=True)
data['y'] = data['y'].astype('int')
data.tail(30)
```

Рисунок 3.4– Фільтрація пропущених значень

Після фільтрації пропущених значень, важливим кроком у подальшій обробці даних є видалення останнього значення, яке рівне нулю. Це рішення може бути виправданим з різних причин, включаючи технічні або методологічні аспекти дослідження. Видалення останнього значення, яке рівне нулю, може бути обґрунтовано з погляду уникнення спотворень у висновках або прогнозах. Нульове значення в кінці часового ряду може бути результатом технічної помилки, аномальності або інших обставин, що не відображають реальний стан даних. Це видалення підвищить точність аналізу та прогнозів, а також сприятиме кращому визначенню змін у динаміці даних.

В результаті цих кроків отримаємо збалансований та надійний набір даних, готовий до подальших досліджень та моделювання.

Код програми, який реалізовує видалення останнього значення яке рівне нулю наведено на рисунок 3.5.

```
# Delete the last zero value
data = data[:-1]
data.tail(30)
```

Рисунок 3.5– Видалення останнього нульового значення

Побудуємо графік для даних з генерації електроенергію ФЕС за період спостережень, як ключовий етап візуалізації та розуміння динаміки даних. Візуалізація графічно демонструє залежності, тренди та можливі аномалії у часових рядах, що є важливим для виявлення патернів та взаємозв'язків (рис. 3.6).

```

# Build the plot
plt.figure(figsize=(20, 6))
plt.plot(data['ds'], data['y'])
plt.legend(['y'])
step = 10000
plt.xticks(range(0, len(data['ds']), step), data['ds'][::step])
plt.xlabel('Date')
plt.ylabel('Power')
plt.show()

```

Рисунок 3.6 – Побудова графіка генерації електроенергію ФЕС за період спостережень

В результаті графічне представлення часового ряду генерації електроенергії ФЕС за період спостережень наведено на рисунку 3.7.

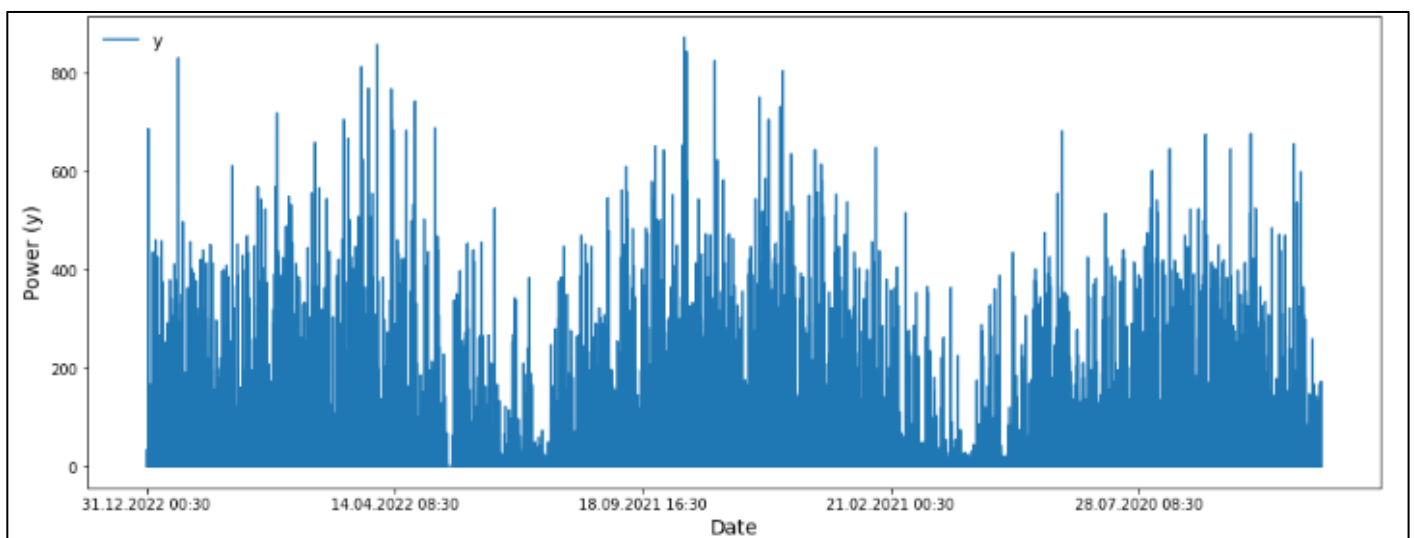


Рисунок 3.7 – Графік генерації електроенергію ФЕС за період спостережень

З аналізу поданого графіка можна зробити важливі висновки, що сприяють розумінню динаміки генерації електроенергії фотоелектростанцією за період

спостережень. Спочатку, чітко виражена сезонність, пов'язана із змінами пори року, відзначається на графіку. Це свідчить про те, що електроенергія генерується відповідно до природних коливань та умов, які змінюються впродовж року.

Додатково, графік відзначається аномаліями у вигляді провалів у генерації електроенергії. Ці провали можуть бути зв'язані з різноманітними факторами, такими як технічні проблеми, погодні умови, або інші непередбачувані обставини. Аналіз цих аномалій може допомогти виокремити ключові моменти, що впливають на генерацію, і врахувати їх у моделі прогнозування. Отже в моделі прогнозування важливо враховувати сезонні коливання та аномалії, що дозволить створити більш точну та надійну модель, яка здатна прогнозувати генерацію електроенергії в умовах змінливого середовища.

Розглянемо часовий ряд у більш детальний спосіб, аналізуючи генерацію електроенергії в межах доби. Це дозволить виявити внутрішні зміни та патерни, які можуть бути пов'язані з щоденними циклами та визначити часи максимальної та мінімальної генерації. Виділимо з часового ряду дні з максимальною та мінімальною генерацією та проведемо їх аналіз.

Для цього потрібно виконати наступні команди (рис. 3.8).

```
[8]: df_solar['DATE_TIME'] = pd.to_datetime(df_solar['DATE_TIME'])
daily_avg_power = df_solar.groupby(df_solar['DATE_TIME'].dt.date)['DC_POWER'].mean()
max_avg_date = daily_avg_power.idxmax()
print(f"The date with the highest average DC_POWER is: {max_avg_date}")
```

The date with the highest average DC_POWER is: 2021-04-28

Рисунок 3.8 – Вибір дня з максимальною генерацією з часового ряду за весь період спостережень

Програмна реалізація побудови графіка та графік наведені на рисунках 3.9, 3.10.

```
[10]: plt.figure(figsize=(16,16))
date=["2021-04-28"]
plt.subplot(411)
sns.lineplot(df_solar[df_solar["DATE_STRING"].isin(date)].DATE_TIME, df_solar[df_solar["DATE_STRING"].isin(date)].DC_POWER, label="DC_Power_Best", color='green')
plt.title("DC Power Generation: {}".format(date[0]))
plt.tight_layout()
plt.show()
```

Рисунок 3.9 – Побудова графіка генерації в межах доби

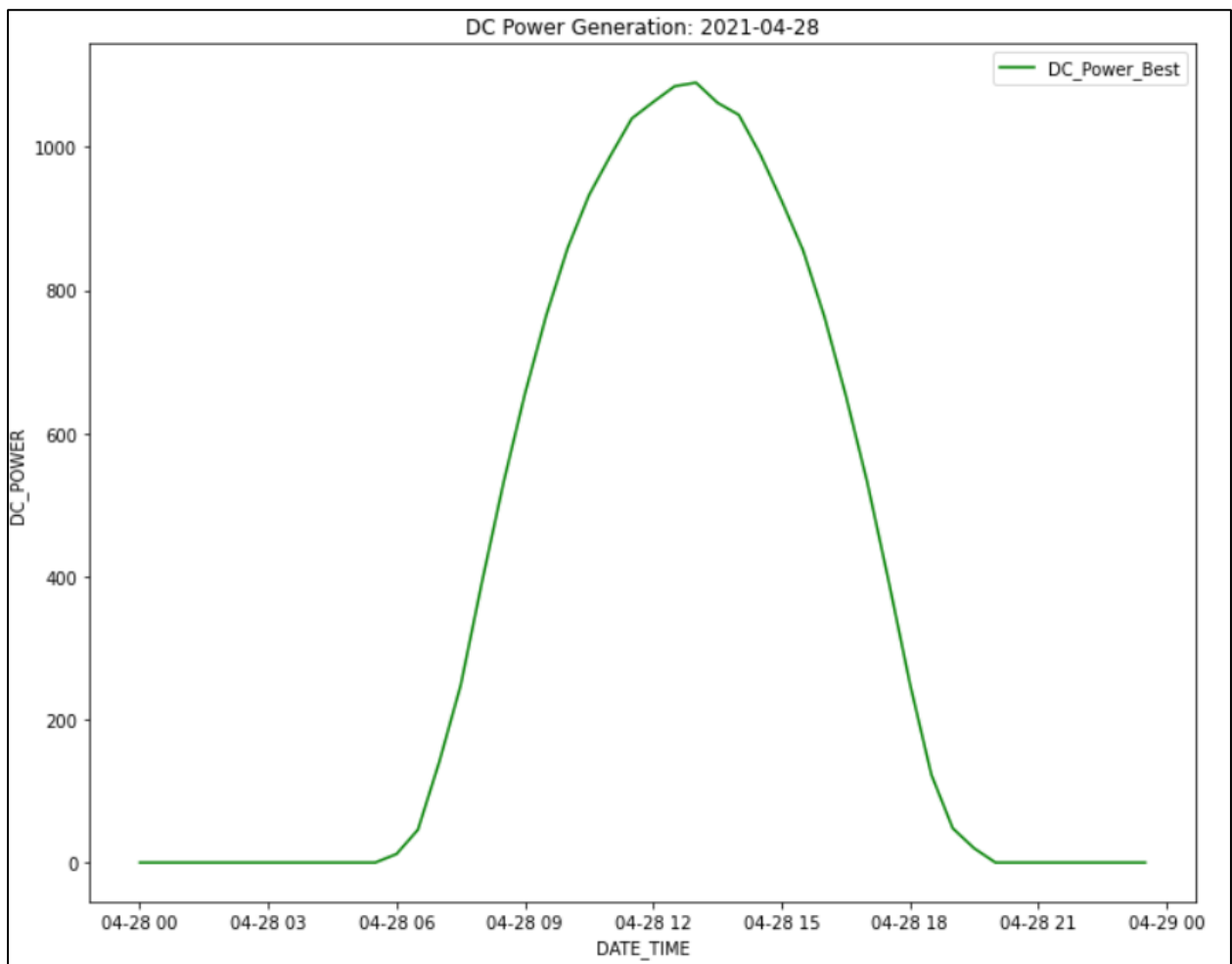


Рисунок 3.10 – Графік максимальною генерацією в межах доби з часового ряду за весь період спостережень

Аналогічним чином будемо будемо графік з найменшою генерацією за період спостережень. Оскільки в датасеті можуть міститися дні коли електростанція взагалі була відключена від мережі, то потрібно відфільтрувати такі дні як аномальні значення. Для цього виконується наступний код (рис. 3.11)

```
[12]:
df_solar['DATE_TIME'] = pd.to_datetime(df_solar['DATE_TIME'])
daily_avg_power = df_solar.groupby(df_solar['DATE_TIME'].dt.date)['DC_POWER'].mean()
zero_power_dates = daily_avg_power[daily_avg_power == 0].index
df_solar_filtered = df_solar[~df_solar['DATE_TIME'].dt.date.isin(zero_power_dates)]
```

Рисунок 3.11 – Видалення днів з нульовою генерацією за добу з датасету

Далі виконується побудова графіка з найменшою генерацією за період спостережень (рис. 3.12).

```
[13]:
df_solar_filtered['DATE_TIME'] = pd.to_datetime(df_solar['DATE_TIME'])
daily_avg_power = df_solar.groupby(df_solar['DATE_TIME'].dt.date)['DC_POWER'].mean()
max_avg_date = daily_avg_power.idxmin()
print(f"The date with the lower average DC_POWER is: {max_avg_date}")
```

The date with the lower average DC_POWER is: 2022-12-19

Рисунок 3.12 – Вибір дня з мінімальною генерацією з часового ряду за весь період спостережень

Графік доби з найменшою генерацією електроенергії за період спостережень наведено на рисунку 3.13.

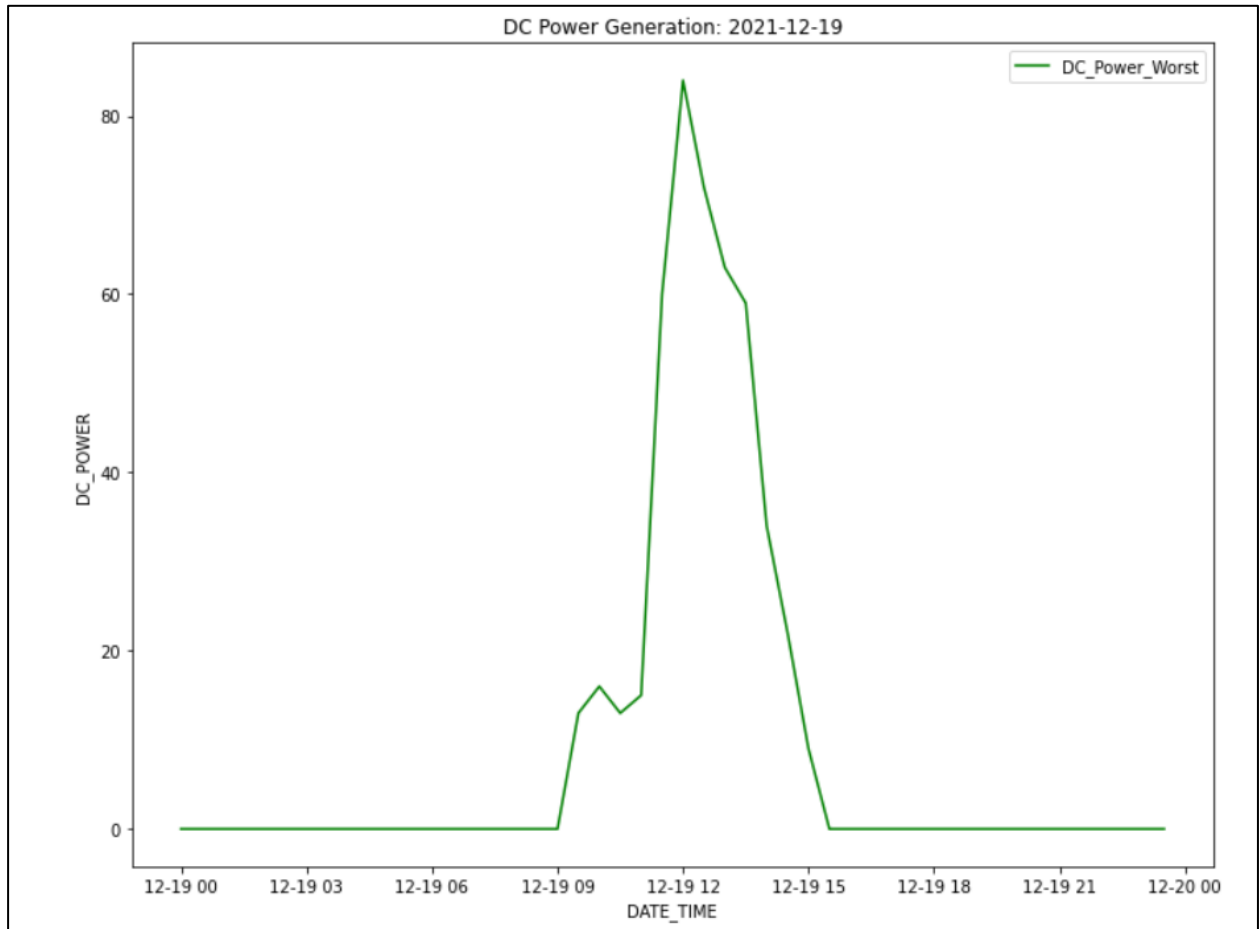


Рисунок 3.13 – Графік доби з мінімальною генерацією з часового ряду за весь період спостережень

Графіки чітко демонструють добову сезонність у генерації електроенергії фотоелектростанцією. Це виражається у періодичних змінах у рівні генерації, які повторюються впродовж 24 годин. Зазвичай відзначаються піками генерації в практично однаковий час кожного дня, відповідно до руху сонця та його впливу на ефективність сонячних панелей. Крім того, на графіках може бути помітно певні аномалії чи відхилення, які можуть бути пов'язані з різноманітними факторами, такими як технічні проблеми чи непередбачені події. Такий детальний аналіз цих графіків є важливим для підбору параметрів налаштування моделі прогнозування.

Далі, в процесі нашої роботи, переходимо до важливого етапу – подальшої підготовки даних для моделювання з використанням високоефективної бібліотеки

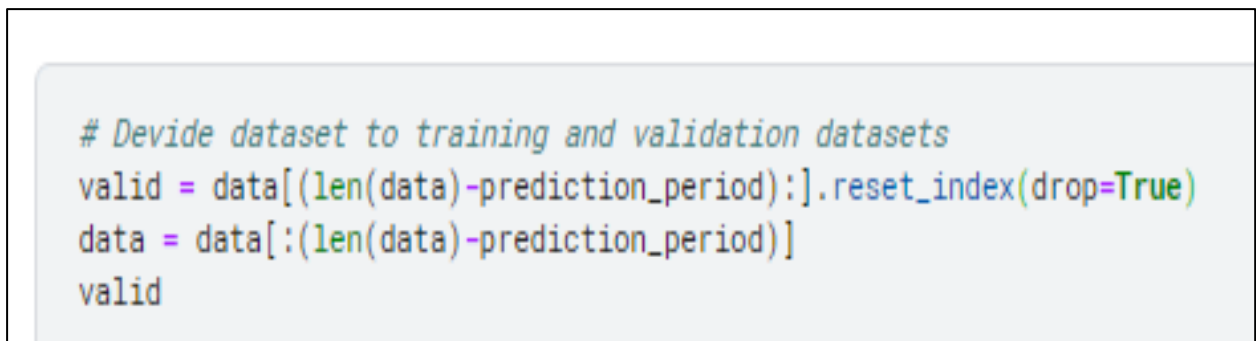
Prophet (рис. 3.14). Цей етап є важливим для створення точної та надійної прогностичної моделі, спроможної передбачати майбутні зміни в генерації електроенергії.



```
# Preparing data for modeling with Prophet
data.columns = ['ds', 'y']
data.tail()
```

Рисунок 3.14 – Моделювання з використанням бібліотеки Prophet

Далі у процесі роботи, проводиться розділення даних на два основні набори: набір для навчання та набір для перевірки, щоб використовувати їх при моделюванні з використанням бібліотеки Prophet (рисунок 3.15).



```
# Devide dataset to training and validation datasets
valid = data[(len(data)-prediction_period):].reset_index(drop=True)
data = data[:(len(data)-prediction_period)]
valid
```

Рисунок 3.15 – Розділення даних на навчальний та тренувальний набір

Набір для навчання:

Цей набір використовується для навчання прогностичної моделі. Він включає історичні дані про генерацію електроенергії та відповідає за основні знання, які модель використовує для усвідомлення та пристосування до закономірностей вхідних даних.

Набір для перевірки:

Цей набір служить для оцінки ефективності нашої моделі на нових, раніше невиданих даних. Він використовується для валідації точності та надійності прогнозів, зроблених моделлю, і дозволяє нам визначити, наскільки добре модель може узагальнювати свої знання на нових областях часового ряду.

Цей етап розділення даних допомагає уникнути перенавчання моделі та забезпечує її здатність ефективно працювати з реальними, невідомими даними. Такий підхід сприяє створенню більш надійної та адаптованої прогностичної моделі для генерації електроенергії.

Наступним етапом у аналізі є побудова моделі за допомогою бібліотеки Prophet та її навчання на усіх доступних даних. Цей процес включає в себе налаштування моделі, визначення її параметрів та здійснення оптимального навчання.

Процедура навчання моделі на всіх даних дозволяє врахувати весь спектр історичних змін у генерації електроенергії, щоб отримати більш точні та узагальнені прогнози на майбутнє. Важливо визначити параметри моделі, такі як чутливість до змін (changepoints), а також додаткові сезонні та підтримуючі компоненти.

Зокрема, додавання параметрів сезонності є ключовим етапом. Врахування сезонних факторів, таких як добова, тижнева та річна сезонності, дозволяє моделі більш точно адаптуватися до різних циклів та патернів у генерації електроенергії. Це підвищує точність та достовірність прогнозів моделі.

Отже, цей етап включає в себе не лише навчання моделі, але й оптимізацію її параметрів, забезпечуючи максимальну адаптивність до умов та закономірностей, які можуть впливати на генерацію електроенергії на фотоелектростанції (рис. 3.16).

Побудова моделі Prophet з параметрами за замовчуванням наведена на рисунку 3.16. Побудова моделі Prophet з налаштуванням параметрів сезонності наведена на рисунку 3.17.

```

# Build Prophet model
model = Prophet()

# Training model
model.fit(train)

```

Рисунок 3.16 – Побудова моделі Prophet з параметрами за замовчуванням

```

# Build Prophet model with parameters and structure
model2 = Prophet(daily_seasonality=True, weekly_seasonality=True, yearly_seasonality=True,
                 changepoint_range=1, changepoint_prior_scale = 0.3)
model2.add_seasonality(name='weekly', period=7, fourier_order=12,
                      mode = 'multiplicative', prior_scale = 0.24)
model2.add_seasonality(name='triplly', period=3, fourier_order=2,
                      mode = 'multiplicative', prior_scale = 0.15)

# Training model
model2.fit(train)

```

<fbprophet-forecaster-Prophet-at-0x7c515869d910>

Рисунок 3.17 – Побудова моделі Prophet з налаштуванням параметрів сезонності

На цьому етапі ми переходимо до фази прогнозування, використовуючи навчену модель Prophet (рис. 3.18). Для цього визначаємо період прогнозування (`prediction_period`), який визначає, на який час ми хочемо здійснити прогноз.

Модель Prophet враховує різні фактори, такі як сезонність, зміни тренду та свята, для отримання точних прогнозів. Після завершення прогнозування ми отримуємо прогнозовані значення на кожен день у визначений період часу.

Цей етап є важливим для подальшого планування та прийняття рішень. Прогнозовані дані можуть служити основою для оптимізації виробничих процесів та планування використання електроенергії, допомагаючи забезпечити ефективне використання сонячної енергії на фотоелектростанції.

```

# Make a forecast for prediction_period days ahead
future = model.make_future_dataframe(periods = prediction_period)
forecast = model.predict(future)

# Make values integer, and replace negative values with zero
feature_all = ['yhat_lower', 'yhat', 'yhat_upper']
forecast[feature_all] = forecast[feature_all].astype('int')
for feature in feature_all:
    forecast.loc[forecast[feature] < 0, feature] = 0

```

Рисунок 3.18 – Прогнозування на prediction_period днів вперед за допомогою моделі Prophet

В результаті графік прогнозування наведено на рисунку 3.19

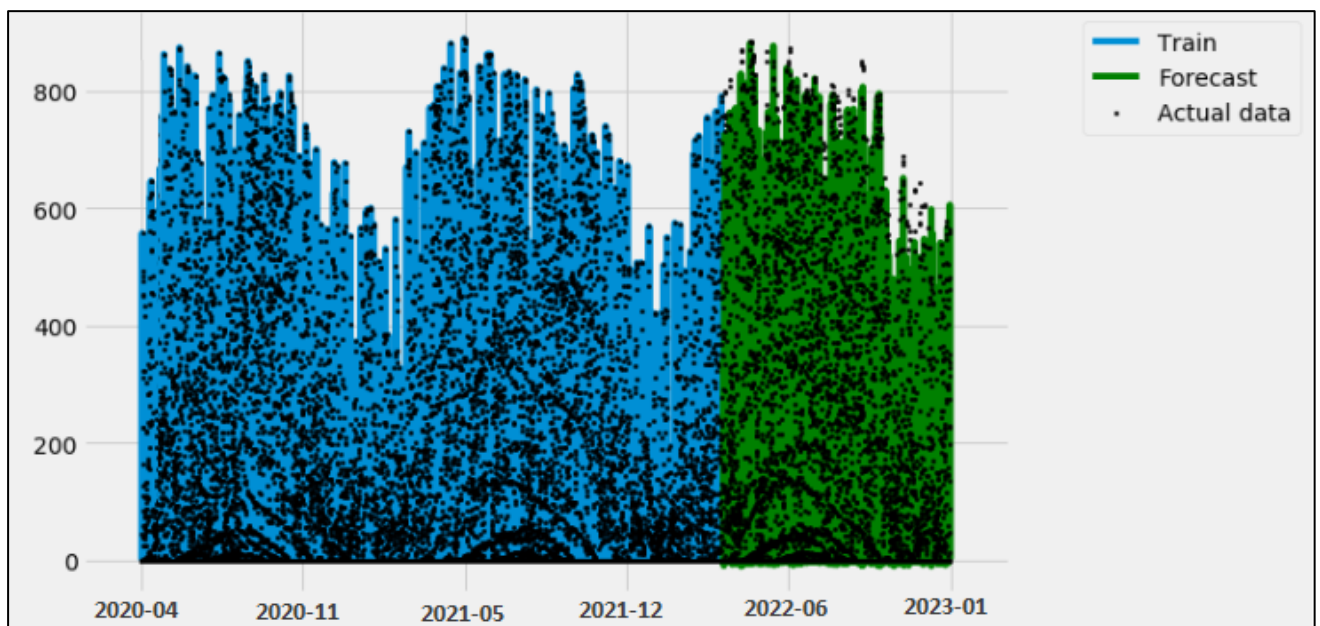


Рисунок 3.19 – Графік прогнозування генерацій електроенергії

Виконаємо прогнозування генерації електроенергії поза межами тренувального набору. Оскільки задача прогнозування передбачає короткостроковий прогноз на добу, тому для цього датасету було відібрано одну добу яка не приймала участь у навчанні та тренуванні моделі. Графік прогнозування генерацій електроенергії на добу з параметрами за замовчуванням

наведено на рисунку 3.20. Графік прогнозування генерацій електроенергії на добу з налаштуванням параметрів сезонності наведено на рисунку 3.21

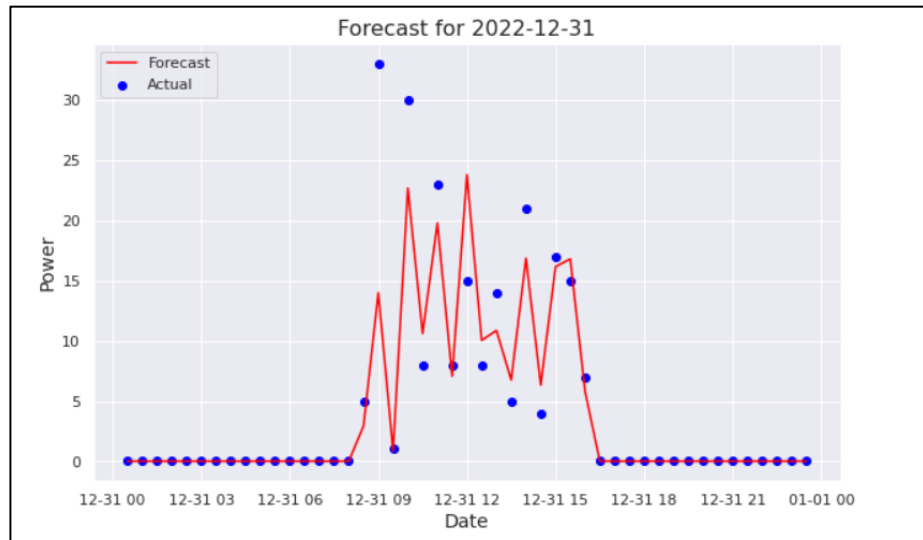


Рисунок 3.20 – Графік прогнозування генерацій електроенергії на добу

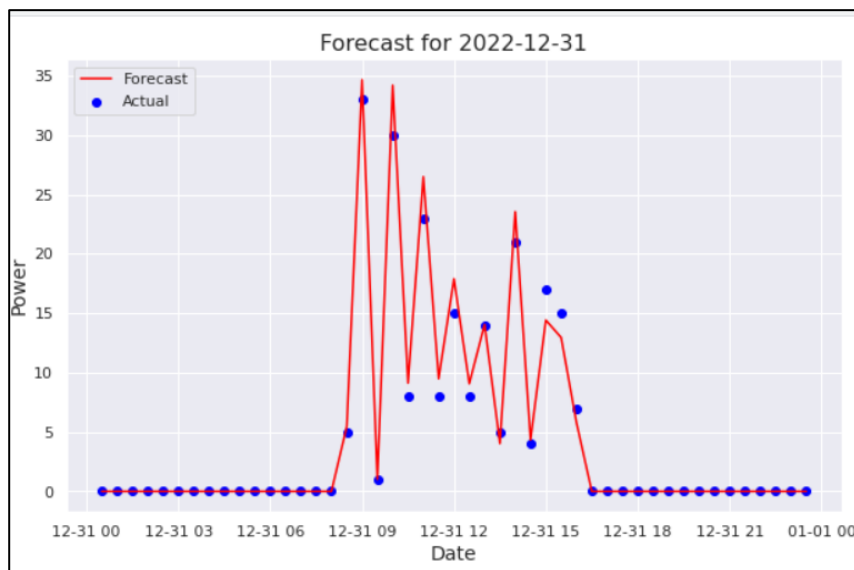


Рисунок 3.21 – Графік прогнозування генерацій електроенергії на добу

Графіки річної, тижневої, сезонної, а також денної циклічності наведено на рисунку 3.22.

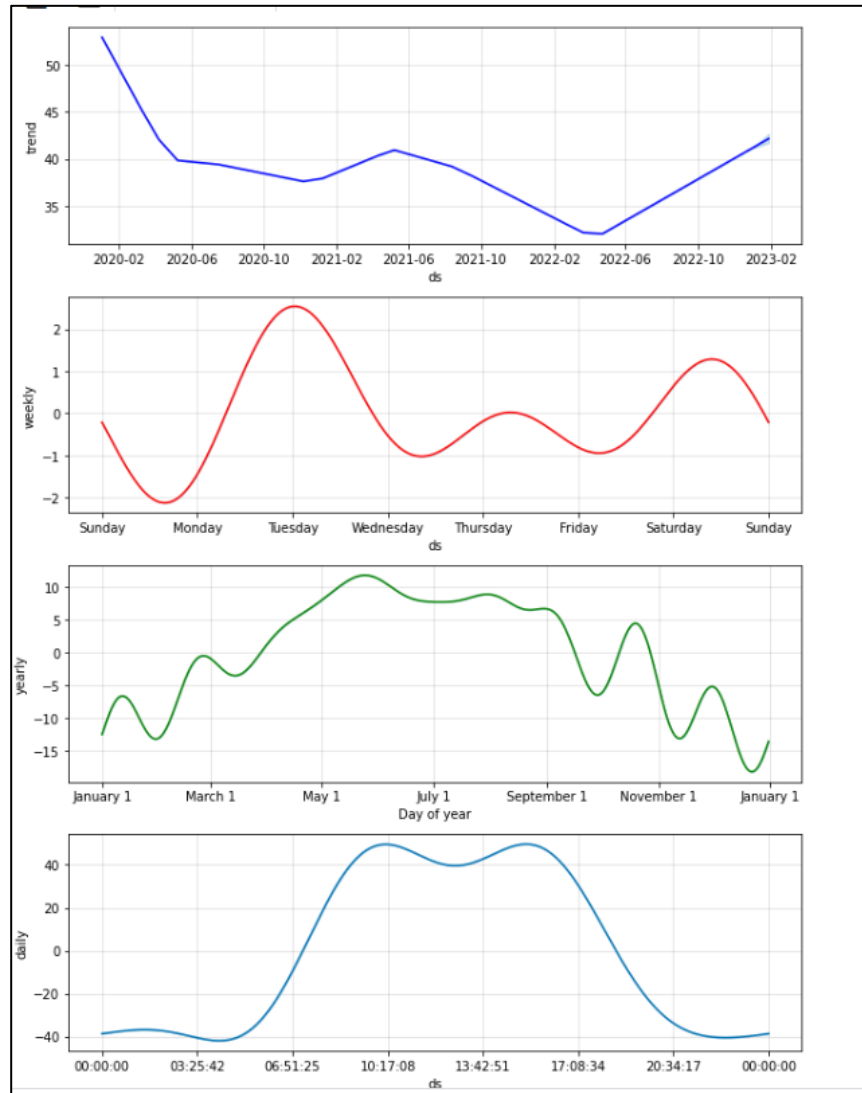


Рисунок 3.22 – Графіки річної, тижневої, сезонної, а також денної циклічності

Розрахунок похибки прогнозування на Prophet з параметрами за замовчуванням наведено на рисунку 3.23. Розрахунок похибки прогнозування на валідаційному наборі даних з налаштуванням параметрів сезонності наведено на рисунку 3.24.

```
print(f"Relative error (WAPE) for model = {prediction_accuracy(valid, model)}%")
```

Relative error (WAPE) for model = 33.78%

Рисунок 3.23 – Похибка моделі прогнозування з параметрами за замовчуванням

```
:  
print(f"Relative error (WAPE) for model = {prediction_accuracy(valid, model)}%")  
Relative error (WAPE) for model = 6.98%
```

Рисунок 3.24 – Похибка моделі прогнозування з налаштуванням параметрів сезонності

Отримана похибка прогнозування складає 6.98% що є чудовим та переконливим доказом високої точності та ефективності моделі Prophet у передбаченні генерації електроенергії на фотоелектростанції.

3.3 Висновки

В розділі було проведено огляд можливостей Kaggle для аналізу та прогнозування генерації електроенергії на ФЕС, було використано бібліотеку Prophet для побудови моделі. Основною метою дослідження було покращення точності та достовірності прогнозів.

Під час попереднього етапу обробки даних виконана фільтрація нульових значень, що вважаються аномальними. Далі дані були розділені на навчальний, тренувальний та валідаційний набори для побудови та оцінки моделі. Здійснено підбір оптимальних параметрів, включаючи параметри сезонності.

Було проведено налаштування моделі з врахуванням сезонності, такої як добова, тижнева та річна, що є ключовим етапом в побудові моделі.

По результатам досліджень було обрана модель Prophet з налаштуванням параметрів сезонності, яка у порівнянні з моделлю з параметрами за замовчуванням видає чудовий результат прогнозування з похибкою 6.98%. Низький рівень похибки говорить про те, що модель успішно враховує різноманітні аспекти, такі як сезонні коливання, зміни тренду та інші важливі фактори, що впливають на генерацію електроенергії. Отримані точні прогнози є вкрай корисними для

стратегічного планування та управління енергетичними ресурсами фотоелектростанції. З ними можна ефективно регулювати роботу станції, максимізувати використання сонячної енергії та оптимізувати виробничі процеси для досягнення максимальної продуктивності. Даний результат свідчить про правильний вибір налаштувань та правильний вибір методу прогнозування, доводячи, що модель Prophet виявляється ефективною для точного визначення генерації електроенергії на фотоелектростанціях. Такий результат підкреслює доцільність використання такого підходу у сфері прогнозування генерації електроенергії нестабільними відновлювальними джерелами енергії.

4 ЕКОНОМІЧНА ЧАСТИНА

4.1 Проведення комерційного та технологічного аудиту науково-технічної розробки

Питання генерації зеленої енергії надзвичайно актуальне не тільки для технологічно розвинених країн з величезною промисловістю а також для країн, що розвиваються зокрема і для України. Щороку спостерігається величезний приріст відновлюваних джерел енергії, більша частка яких є нестабільними, оскільки залежить від погодних умов. Враховуючи вплив нестабільних джерел енергії на енергосистему та формування ціни на електроенергію важливим є створення методів та засобів прогнозування генерації електроенергії, що дозволить операторам якісно розподіляти генерацію та споживання, а також дозволить чітко формувати ціну на електроенергію для усіх учасників ринку.

Тому питання прогнозування генерації електроенергії такими джерелами є актуальним не тільки з технічної точки зору, але й з економічної.

Особлива увага приділяється аудиту технологій та розробок, пов'язаних з виробництвом сонячної енергії, зокрема, прогнозуванню генерації електроенергії фотоелектростанціями.

Проведення комерційного аудиту передбачає оцінку економічної доцільності та фінансової стійкості науково-технічної розробки, її впливу на ринкові показники, а також визначення стратегічних переваг для бізнесу. У контексті фотоелектрогенерації, цей аспект включає в себе аналіз та прогнозування генерації сонячної електроенергії, яка є ключовим параметром для визначення рентабельності та конкурентоспроможності фотоелектростанцій.

Технологічний аспект аудиту передбачає оцінку сучасних та перспективних технологічних рішень у галузі фотоелектрогенерації, включаючи вдосконалення сонячних панелей, систем зберігання енергії та систем моніторингу. Прогнозування генерації електроенергії відображається в оцінці технічної ефективності інновацій та можливостях їхнього впровадження на ринку.

Зазначена область аудиту визначається важливістю та актуальністю розвитку відновлювальних джерел енергії в сучасному енергетичному ландшафті. Генерація електроенергії від сонячних джерел стає не лише стратегічно важливим напрямом розвитку, але й ключовим чинником для забезпечення сталого енергетичного майбутнього.

Критерії оцінювання наукового-технічного рівня та комерційного потенціалу розробки наведено в таблиці 4.1 [115].

Таблиця 4.1 – Критерії оцінювання наукового-технічного рівня та комерційного потенціалу розробки та бальна оцінка (в таблиці проблеми з пробілами хоча це може так у мене відображається)

Бали (за п'ятибальною шкалою)					
	0	1	2	3	4
1	2	3	4	5	6
<i>Технічна здійсненність концепції</i>					
1	Достовірність Концепції не підтверджена	Концепція підтверджена експертними висновками	Концепція підтверджена розрахунками	Концепція перевірена на практиці	Перевірено роботоздатність продукту в реальних умовах
<i>Ринкові переваги (недоліки)</i>					
2	Багато аналогів на малому ринку	Мало аналогів на малому ринку	Кілька аналогів на великому ринку	Один аналог на великому ринку	Продукт не має аналогів на великому ринку
3	Ціна продукту значно вища за ціни аналогів	Ціна продукту дещо вища за ціни аналогів	Ціна продукту приблизно дорівнює цінам аналогів	Ціна продукту дещо нижча за ціни аналогів	Ціна продукту значно нижча за ціни аналогів
4	Технічні та споживчі властивості продукту значно гірші, ніж в аналогів	Технічні та споживчі властивості продукту трохи гірші, ніж в аналогів	Технічні та споживчі властивості продукту нарівні аналогів	Технічні та споживчі властивості продукту трохи кращі, ніж в аналогів	Технічні та споживчі властивості продукту значно кращі, ніж в аналогів
5	Експлуатаційні витрати значно вищі, ніж в аналогів	Експлуатаційні витрати дещо вищі, ніж в аналогів	Експлуатаційні витрати на рівні експлуатацій-них витрат аналогів	Експлуатаційні витрати трохи нижчі, ніж аналогів	Експлуатаційні витрати значно нижчі, ніж аналогів
<i>Ринкові перспективи</i>					

Продовження таблиці 4.1

6	Ринок малий і не має позитивної динаміки	Ринок малий, але має позитивну динаміку	Середній ринок позитивною динамікою	Великий стабільний ринок	Великий ринок з позитивною динамікою
7	Активна конкуренція великих компаній на ринку	Активна конкуренція	Помірна конкуренція	Незначна конкуренція	Конкурентів немає
<i>Практична здійсненність</i>					
8	Відсутні фахівці як з технічної, так і з комерційної реалізації ідеї	Необхідно наймати фахівців або витратити значні кошти та час на навчання наявних фахівців	Необхідне незначне навчання фахівців та збільшення їх штату	Необхідне незначне навчання фахівців	Є фахівці з питань як з технічної, так і з комерційної реалізації ідеї
9	Потрібні значні фінансові ресурси, які відсутні. Джерела фінансування ідеї відсутні	Потрібні незначні фінансові ресурси. Джерела фінансування відсутні	Потрібні значні фінансові ресурси. Джерела фінансування є	Потрібні незначні фінансові ресурси. Джерела фінансування є	Не потребує додаткового фінансування
10	Необхідна розробка нових матеріалів	Потрібні матеріали, що використовуються у військово-промисловому комплексі	Потрібні дорогі матеріали	Потрібні досяжні та дешеві матеріали	Всі матеріали для реалізації ідеї відомі та давно використовуються у виробництві
11	Термін реалізації ідеї більший за 10 років	Термін реалізації ідеї більший за 5 років. Термін окупності інвестицій більше 10-ти років	Термін реалізації ідеї від 3-х до 5-ти років. Термін окупності інвестицій більший 5-ти років	Термін реалізації ідеї менший 3-х років. Термін окупності інвестицій від 3-х до 5-ти років	Термін реалізації ідеї менший 3-х років. Термін окупності інвестицій менший 3-х років
12	Необхідна розробка регламентних документів та отримання великої кількості дозвільних документів на виробництво та реалізацію продукту	Необхідно отримання великої кількості дозвільних документів на виробництво та реалізацію продукту, що потребує значних коштів та часу	Процедура отримання дозвільних документів для виробництва та реалізації продукту потребує незначних коштів та часу	Необхідно тільки повідомлення відповідним органам про виробництво та реалізацію продукту	Відсутні будь-які регламентні обмеження на виробництво та реалізацію продукту

Результати оцінювання комерційного потенціалу розробки наведено в таблиці 4.2

Таблиця 4.2 Результат оцінювання комерційного потенціалу розробки

Критерії	Експерт		
	1	2	3
	Бали:		
1. Технічна здійсненність концепції	3	4	4
2. Ринкові переваги (наявність аналогів)	4	3	4
3. Ринкові переваги (ціна продукту)	3	4	4
4. Ринкові переваги (технічні властивості)	3	4	3
5. Ринкові переваги (експлуатаційні витрати)	2	3	3
6. Ринкові перспективи (розмір ринку)	4	4	3
7. Ринкові перспективи (конкуренція)	4	3	3
8. Практична здійсненність (наявність фахівців)	3	4	3
9. Практична здійсненність (наявність фінансів)	3	3	3
10. Практична здійсненність (необхідність нових матеріалів)	3	2	3
11. Практична здійсненність (термін реалізації)	3	3	3
12. Практична здійсненність (розробка документів)	3	3	4
Сума балів	СБ ₁ =38	СБ ₂ =40	СБ ₃ =40
Середньоарифметична сума балів СБ _с	$СБ_c = \frac{\sum_{i=1}^3 СБ_i}{3} = \frac{118}{3} = 39,3$		

За даними таблиці 4.2 можна зробити висновок щодо науково-технічного рівня і рівня комерційного потенціалу розробки. Для цього доцільно скористатись рекомендаціями, наведеними в таблиці 4.3 [115].

Таблиця 4.3 Науково-технічні рівні та комерційні потенціали розробки

Середньоарифметична сума балів СБ, розрахована на основі висновків експертів	Науково-технічний рівень та комерційний потенціал розробки
41...48	Високий
31...40	Вищий середнього
21...30	Середній
11...20	Нижчий середнього
0...10	Низький

Як видно з таблиці, науково-технічний рівень та комерційний потенціал розробки є вищим середнього, що досягається за рахунок того, що оператор кожної сонячної електростанції зобов'язаний подавати прогноз генерації електроенергії на добу наперед погодинно, у випадку відхилення від графіка більше ніж на 10%, підприємство з виробництва електроенергії буде сплачувати штрафи, відповідно точний прогноз генерації є надзвичайно важливим Вищий середнього рівень комерційного потенціалу розробленої інформаційної технології досягається за рахунок того, що остання використовує методи машинного навчання, які дозволяють підвищити точність передбачення у порівнянні з аналогічними технологіями [115].

4.2 Розрахунок витрат на проведення науково-технічної роботи

4.2.1 Витрати на оплату праці

Основна заробітна плата дослідників.

Витрати на основну заробітну плату дослідників (Z_0) розраховують відповідно до посадових окладів працівників, за формулою [113]:

$$Z_0 = \sum_{i=1}^k \frac{M_{ni} \cdot t_i}{T_p}, \quad (4.1)$$

де k – кількість посад дослідників, залучених до процесу досліджень;

$M_{\text{ш}}$ – місячний посадовий оклад конкретного дослідника, грн;

t_i – кількість днів роботи конкретного дослідника, дн.;

T_p – середня кількість робочих днів в місяці, $T_p=21\dots23$ дні.

Проведені розрахунки бажано звести до таблиці 4.4.

Таблиця 4.4 – Витрати на заробітну плату дослідників

Найменування посади	Місячний посадовий оклад, грн	Оплата за робочий день, грн	Кількість днів роботи	Витрати на заробітну плату, грн
Керівник проекту	33000	1500	50	75000
Програміст	28600	1300	50	65000
Всього				140000

Основна заробітна плата робітників

Витрати на основну заробітну плату робітників (Z_p) за відповідними найменуваннями робіт на тему «Аналіз та прогнозування добового генерування електроенергії» розраховується за формулою:

$$Z_p = \sum_{i=1}^n C_i * t_i, \quad (4.2)$$

де C_i – погодинна тарифна ставка робітника відповідного розряду, за виконану відповідну роботу, грн/год;

t_i – час роботи робітника на виконання певної роботи, год.

Погодинну тарифну ставку робітника відповідного розряду C_i можна визначити за формулою:

$$C_i = \frac{M_M \cdot K_i \cdot K_C}{T_p \cdot t_{3M}}, \quad (4.3)$$

де M_M – розмір прожиткового мінімуму працездатної особи або мінімальної місячної заробітної плати (залежно від діючого законодавства), приймемо значення 6700 грн;

K_i – коефіцієнт міжкваліфікаційного співвідношення для встановлення тарифної ставки робітнику відповідного розряду [115];

K_c – мінімальний коефіцієнт співвідношень місячних тарифних ставок робітників першого розряду з нормальними умовами праці виробничих об'єднань і підприємств до законодавчо встановленого розміру мінімальної заробітної плати [115];

T_p – середня кількість робочих днів в місяці, приблизно $T_p = 21 \dots 23$ дні;

$t_{зм}$ – тривалість зміни, год.

$$C_I = \frac{6700 \cdot 1,1 \cdot 1,65}{21 \cdot 8} = 72,38 \text{ (грн)},$$

$$Z_p = 72,38 \cdot 8,2 = 593,52 \text{ (грн)}.$$

Проведені розрахунки бажано звести до таблиці 4.5.

Таблиця 4.5 – Величина витрат на основну заробітну плату робітників

Найменування робіт	Тривалість роботи, год	Розряд роботи	Тарифний коефіцієнт	Погодинна тарифна ставка, грн	Величина оплати на робітника, грн
Підготовка робочого місця	8,2	2	1,1	72,38	593,52
Інсталяція програмного забезпечення	6,7	3	1,35	88,83	595,161
Компіляція програмних частин	8,1	5	1,7	111,87	906,147
Всього					2094,83

Додаткова заробітна плата дослідників та робітників.

Додаткову заробітну плату розраховуємо як 10 ... 12% від суми основної заробітної плати дослідників та робітників за формулою:

$$Z_{\text{дод}} = (Z_o + Z_p) \cdot \frac{N_{\text{дод}}}{100\%}, \quad (4.4)$$

де $N_{\text{дод}}$ – норма нарахування додаткової заробітної плати.

$$Z_{\text{дод}} = (140000 + 2094,83) \cdot \frac{12}{100\%} = 17051,38 \text{ (грн)}.$$

4.2.2 Відрахування на соціальні заходи

Нарахування на заробітну плату дослідників та робітників розраховуємо як 22% від суми основної та додаткової заробітної плати дослідників і робітників за формулою:

$$Z_{\text{н}} = (Z_o + Z_p + Z_{\text{дод}}) \cdot \frac{N_{\text{зп}}}{100\%}, \quad (4.5)$$

де $N_{\text{зп}}$ – норма нарахування на заробітну плату.

$$Z_{\text{н}} = (140000 + 2094,83 + 17051,38) \cdot \frac{22}{100\%} = 35013,49 \text{ (грн)}.$$

4.2.3 Сировина та матеріали

Витрати на матеріали (M) у вартісному вираженні розраховуємо окремо для кожного виду матеріалів за формулою:

$$M = \sum_{j=1}^n H_j \cdot C_j - \sum_{j=1}^n B_j \cdot C_j, \quad (4.6)$$

де H_j – норма витрат матеріалу j -го найменування, кг;

n – кількість видів матеріалів;

Ц_j – вартість матеріалу j -го найменування, грн/кг;

K_j – коефіцієнт транспортних витрат, ($K_j = 1,1 \dots 1,15$);

V_j – маса відходів j -го найменування, кг;

Ц_{vj} – вартість відходів j -го найменування, грн/кг.

Проведені розрахунки зведемо в таблицю 4.6.

Таблиця 4.6 – Витрати на матеріали

Найменування матеріалу, марка, тип, сорт	Ціна за 1 кг, грн	Норма витрат, кг	Величина відходів, кг	Ціна відходів, грн/кг	Вартість витраченого матеріалу, грн
Flash-пам'ять 16GB	250	1,0	-	-	275
Офісний папір А4 500	230,0	3,0	-	-	759
Картридж для принтера	890,0	1,0	-	-	979
Набір канцелярський офісний	230,0	2,0	-	-	506,0
Органайзер офісний	530,0	3,0	-	-	1749,0
Папір для записів А5 250	90,0	3,0	-	-	297
Всього					4565

4.2.4 Розрахунок витрат на комплектуючі

Витрати на комплектуючі відсутні.

4.2.5 Спецустаткування для наукових (експериментальних) робіт

Витрати на спецустаткування для наукових (експериментальних) робіт відсутні.

4.2.6 Програмне забезпечення на наукових (експериментальних) робіт

Витрати на програмне забезпечення на наукових (експериментальних) робіт відсутні.

4.2.7 Амортизація обладнання, програмних засобів та приміщень

В спрощеному вигляді амортизаційні відрахування по кожному виду обладнання, приміщень та програмному забезпеченню тощо, розраховуємо з використанням прямолінійного методу амортизації за формулою:

$$A_{\text{обл}} = \frac{Ц_{\text{б}}}{T_{\text{в}}} \cdot \frac{t_{\text{вик}}}{12}, \quad (4.7)$$

де $Ц_{\text{б}}$ – балансова вартість обладнання, програмних засобів, приміщень тощо, які використовувались для проведення досліджень, грн;

$t_{\text{вик}}$ – термін використання обладнання, програмних засобів, приміщень під час досліджень, місяців;

$T_{\text{в}}$ – строк корисного використання обладнання, програмних засобів, приміщень тощо, років.

$$A_{\text{обл}} = \frac{22500 \cdot 3}{2 \cdot 12} = 2812,5 \text{ (грн)}.$$

Проведені розрахунки зведено в таблицю 4.7.

Таблиця 4.7 – Амортизація відрахування по кожному виду обладнання

Найменування обладнання	Балансова вартість, грн	Строк корисного використання, років	Термін використання обладнання, місяців	Амортизаційні відрахування, грн
Персональний комп'ютер проведення розробки та моделювання	22500	2	3	2812,5
Робоче місце інженера розробника	9600	5	3	480
Офісне приміщення	1200000	20	3	15000
ОС Windows 11	7300	2	3	912,5
Прикладний пакет Microsoft Office 2016	7100	2	3	887,5
Всього				20092,5

4.2.8 Паливо та енергія для науково-виробничих цілей

Витрати на силову електроенергію (B_e) розраховуємо за формулою:

$$B_e = \sum_{i=1}^n \frac{W_{yi} \cdot t_i \cdot C_e \cdot K_{впi}}{\eta_i}, \quad (4.8)$$

де W_{yi} – встановлена потужність обладнання на визначеному етапі розробки, кВт;

t_i – тривалість роботи обладнання на етапі дослідження, год;

C_e – вартість 1 кВт-години електроенергії, грн; прийmemo $C_e = 7,50$ грн;

$K_{впi}$ – коефіцієнт, що враховує використання потужності, $K_{впi} < 1$;

η_i – коефіцієнт корисної дії обладнання, $\eta_i < 1$.

$$B_e = 0,25 \cdot 360,0 \cdot 7,50 \cdot 0,96 / 0,97 = 668,04 \text{ грн.}$$

Проведені розрахунки зведемо до таблиці 4.8.

Таблиця 4.8 – Витрати на електроенергію

Найменування обладнання	Встановлена потужність, кВт	Тривалість роботи, год	Сума, грн
Персональний комп'ютер проведення розробки та моделювання	0,25	360,0	668,04
Робоче місце інженера	0,12	310,0	276,12
Всього			944,16

4.2.9 Службові відрядження

Витрати за статтею «Службові відрядження» розраховуємо як 20...25% від суми основної заробітної плати дослідників та робітників за формулою:

$$V_{\text{св}} = (Z_0 + Z_p) \cdot \frac{H_{\text{св}}}{100\%}, \quad (4.9)$$

де $H_{\text{св}}$ – норма нарахування за статтею «Службові відрядження», прийmemo $H_{\text{св}} = 25\%$.

$$V_{\text{св}} = (140000 + 2094,83) \cdot \frac{25}{100\%} = 35523,7 \text{ (грн)}.$$

4.2.10 Витрати на роботи, які виконують сторонні підприємства, установи і організації

Витрати на роботи, які виконують сторонні підприємства, установи і організації відсутні.

4.2.11 Інші витрати

Витрати за статтею «Інші витрати» розраховуємо як 50...100% від суми основної заробітної плати дослідників та робітників за формулою:

$$I_{\text{в}} = (Z_0 + Z_p) \cdot \frac{H_{\text{ів}}}{100\%}, \quad (4.10)$$

де N_{iv} – норма нарахування за статтею «Інші витрати», приймемо $N_{iv} = 50\%$.

$$I_B = (140000 + 2094,83) \cdot \frac{50}{100\%} = (71047,42) \text{ грн.}$$

4.2.12 Накладні (загальновиробничі) витрати

Витрати за статтею «Накладні (загальновиробничі) витрати» розраховуємо як 100...150% від суми основної заробітної плати дослідників та робітників за формулою:

$$V_{H3B} = (Z_o + Z_p) \cdot \frac{N_{H3B}}{100\%}, \quad (4.11)$$

де N_{H3B} – норма нарахування за статтею «Накладні (загальновиробничі) витрати», приймемо $N_{H3B} = 100\%$.

$$V_{H3B} = (140000 + 2094,83) \cdot 100 / 100\% = 142094,83 \text{ (грн).}$$

Витрати на проведення науково-технічної роботи на тему «Аналіз та прогнозування добового генерування електроенергії» розраховуємо як суму всіх попередніх статей витрат за формулою:

$$V_{zag} = Z_o + Z_p + Z_{дод} + Z_H + M + K_B + V_{спец} + V_{прг} + A_{обл} + V_e + V_{св} + V_{сп} + I_B + V_{H3B}, \quad (4.12)$$

$$V_{zag} = 140000 + 2094,83 + 17051,38 + 35013,49 + 4565 + 0 + 0 + 0 + 20092,5 + 944,16 + 35523,7 + 0 + 71047,42 + 142094,83 = 468427,31$$

Загальні витрати ЗВ на завершення науково-технічної роботи та оформлення її результатів розраховується за формулою:

$$ЗВ = \frac{В_{заг}}{\eta}, \quad (4.13)$$

де η - коефіцієнт, який характеризує етап (стадію) виконання науково-технічної роботи, прийmemo $\eta=0,90$.

$$ЗВ = \frac{468427,31}{0,90} = 520474,79 \text{ грн.}$$

4.3 Розрахунок економічної ефективності науково-технічної розробки за її можливої комерціалізації потенційним інвестором

В умовах ринку значущим позитивним вигодою для потенційного інвестора, пов'язаною із можливим впровадженням результатів науково-технічної розробки, є збільшення чистого прибутку. Розширення чистого прибутку сприятиме надходженню додаткових коштів для потенційного інвестора, поліпшить його фінансові показники, підвищить конкурентоспроможність та може позитивно вплинути на його рішення щодо комерціалізації даної розробки.

Для того, щоб розрахувати можливе зростання чистого прибутку у потенційного інвестора від можливого впровадження науково-технічної розробки, слід виконати наступні кроки

а) вказати, з якого часу можуть бути впроваджені результати науково-технічної розробки;

б) зазначити, протягом скількох років після впровадження цієї науково-технічної розробки очікуються основні позитивні результати для потенційного інвестора (наприклад, протягом 4-х років після її впровадження);

в) кількісно оцінити величину існуючого та майбутнього попиту на цю або аналогічні чи подібні науково-технічні розробки та назвати основних суб'єктів (зацікавлених осіб) цього попиту;

г) визначити ціну реалізації на ринку науково-технічних розробок з аналогічними чи подібними функціями.

Так, врахування зміни вартості грошей у часі є критичним аспектом при розрахунку економічної ефективності. Цей принцип відомий як "часова вартість грошей" і базується на ідеї, що гроші, які витрачаються чи заробляються у майбутньому, мають іншу вартість порівняно з грошми, які витрачаються чи заробляються сьогодні.

При оцінюванні ефективності інноваційних проектів передбачається розрахунок таких важливих показників:

- абсолютного економічного ефекту (чистого дисконтованого доходу);
- внутрішньої економічної дохідності (внутрішньої норми дохідності);
- терміну окупності (дисконтованого терміну окупності).

Проведення науково-технічних розробок та оцінка їхньої економічної ефективності в контексті можливої комерціалізації потенційним інвестором може бути здійснена через аналіз різних напрямків, зокрема в контексті розробки чи суттєвого вдосконалення програмного засобу (програмного забезпечення, програмного продукту) для використання масовим споживачем.

У випадку розробки або вдосконалення програмного засобу, ключовою метою є створення продукту, який задовольнить потреби масового ринку. Потенційний інвестор може розглядати таку розробку як інвестицію з високим потенціалом прибутку.

В цьому випадку майбутній економічний ефект буде формуватися на основі таких даних [115]:

$$\Delta\Pi_i = (\pm\Delta\Pi_0 + N + \Pi_0 \cdot \Delta N_i)_i \cdot \lambda \cdot \rho \cdot \left(1 - \frac{\rho}{100}\right), \quad (4.15)$$

де $\pm\Delta\Pi_0$ – зміна основного якісного показника від впровадження результатів науково-технічної розробки в аналізованому році.

Зазвичай, таким показником може бути зміна ціни реалізації одиниці нової розробки в аналізованому році (відносно року до впровадження цієї розробки); $\pm\Delta C_0$ може мати як додатне, так і від'ємне значення (від'ємне – при зниженні ціни відносно року до впровадження цієї розробки, додатне – при зростанні ціни). Прийmemo зростання на 175,50 грн;

N – основний кількісний показник, який визначає величину попиту на аналогічні чи подібні розробки у році до впровадження результатів нової науково-технічної розробки. Кількість споживачів прийmemo 35000 осіб;

C_0 – основний якісний показник, який визначає ціну реалізації нової науково-технічної розробки в аналізованому році, $C_0 = C_6 \pm \Delta C_0$;

C_6 – основний якісний показник, який визначає ціну реалізації існуючої (базової) науково-технічної розробки у році до впровадження результатів. Прийmemo 2350,00 грн;

ΔN – зміна основного кількісного показника від впровадження результатів науково-технічної розробки в аналізованому році. Зазвичай таким показником може бути зростання попиту на науково-технічну розробку в аналізованому році (відносно року до впровадження цієї розробки);

λ – коефіцієнт, який враховує сплату потенційним інвестором податку на додану вартість. У 2023 році ставка податку на додану вартість становить 18%, а коефіцієнт $\lambda = 0,8333$;

ρ – коефіцієнт, який враховує рентабельність інноваційного продукту (послуги). Рекомендується брати $\rho = 0,2 \dots 0,5$. Прийmemo $\rho = 0,35$;

ϑ – ставка податку на прибуток, який має сплачувати потенційний інвестор, у 2023 році $\vartheta = 18\%$

Збільшення чистого прибутку 1-го року:

$$\Delta\Pi_1 = (175,50 \cdot 30000,0 + 2525,5 \cdot 3000) \cdot 0,83 \cdot 0,35 \cdot \left(1 - \frac{0,18}{100\%}\right) =$$

3723740,93 (грн).

Збільшення чистого прибутку 2-го року:

$$\Delta\Pi_2 = (175,50 \cdot 30000,0 + 2525,5 \cdot 8000) \cdot 0,83 \cdot 0,35 \cdot \left(1 - \frac{0,18}{100\%}\right) =$$

7385426,76 (грн).

Збільшення чистого прибутку 3-го року:

$$\Delta\Pi_3 = (175,50 \cdot 30000,0 + 2525,5 \cdot 15000) \cdot 0,83 \cdot 0,35 \cdot \left(1 - \frac{0,18}{100\%}\right) =$$

12511786,9 (грн).

Збільшення чистого прибутку 4-го року:

$$\Delta\Pi_4 = (175,50 \cdot 30000,0 + 2525,5 \cdot 17000) \cdot 0,83 \cdot 0,35 \cdot \left(1 - \frac{0,18}{100\%}\right) =$$

13976461,3 (грн).

Далі розраховуємо приведену вартість збільшення всіх чистих прибутків ПП, що їх може отримати потенційний інвестор від можливого впровадження та комерціалізації науково-технічної розробки:

$$ПП = \sum_{i=1}^T \frac{\Delta\Pi_i}{(1+\tau)^t}, \quad (4.16)$$

де $\Delta\Pi_i$ – збільшення чистого прибутку у кожному з років, протягом яких виявляються результати впровадження науково-технічної розробки, грн;

T – період часу, протягом якого очікується отримання позитивних результатів від впровадження та комерціалізації науково-технічної розробки, роки;

τ – ставка дисконтування, за яку можна взяти щорічний прогнозований рівень інфляції в країні, $\tau = 0,05 \dots 0,15$. Прийmemo $\tau = 0,15$;

t – період часу (в роках) від моменту початку впровадження науково-технічної розробки до моменту отримання потенційним інвестором додаткових чистих прибутків у цьому році.

$$ПП = \frac{3723740,93}{(1+0,15)^1} + \frac{7385426,76}{(1+0,15)^2} + \frac{12511786,9}{(1+0,15)^3} + \frac{13976461,3}{(1+0,15)^4} = 323221,44 +$$

5584443,67 + 8231438,75 + 7986549,31 = 22125653,2грн.

Далі розраховують величину початкових інвестицій PV , які потенційний інвестор має вкласти для впровадження і комерціалізації науково-технічної розробки. Для цього можна використати формулу:

$$PV = k_{\text{інв}} \cdot ЗВ, \quad (4.17)$$

де $k_{\text{інв}}$ – коефіцієнт, що враховує витрати інвестора на впровадження науково-технічної розробки та її комерціалізацію. Це можуть бути витрати на підготовку приміщень, розробку технологій, навчання персоналу, маркетингові заходи тощо; зазвичай $k_{\text{інв}} = 2 \dots 5$, але може бути і більшим. Прийmemo $k_{\text{інв}} = 2,1$;
 $ЗВ$ – загальні витрати на проведення науково-технічної розробки та оформлення її результатів, приймаємо 520474,79 грн.

$$PV = k_{\text{інв}} \cdot ЗВ = 2,1 \cdot 520474,79 = 1092997,06 \text{ (грн)}.$$

Тоді абсолютний економічний ефект $E_{\text{абс}}$ або чистий приведений дохід (NPV, NetPresentValue) для потенційного інвестора від можливого впровадження та комерціалізації науково-технічної розробки становитиме:

$$E_{\text{абс}} = \text{ПП} - PV, \quad (4.18)$$

де ПП – приведена вартість зростання всіх чистих прибутків від можливого впровадження та комерціалізації науково-технічної розробки, 22125653,2 грн;
 PV – теперішня вартість початкових інвестицій, 1092997,06 грн.

$$E_{\text{абс}} = 22125653,2 - 1092997,06 = 21032656,1 \text{ грн}.$$

Внутрішня економічна дохідність інвестицій E_B , які можуть бути вкладені потенційним інвестором у впровадження та комерціалізацію науково-технічної розробки, розраховується за формулою:

$$E_B = \sqrt[T_{ж}]{1 + \frac{E_{абс}}{PV} - 1}, \quad (4.19)$$

де $E_{абс}$ – абсолютний економічний ефект вкладених інвестицій, 20209937.76 грн;

PV – теперішня вартість початкових інвестицій, 1092997,06 грн;

$T_{ж}$ – життєвий цикл науково-технічної розробки, тобто час від початку

її розробки до закінчення отримання позитивних результатів від її впровадження, 4 роки.

$$E_B = \sqrt[T_{ж}]{1 + \frac{E_{абс}}{PV} - 1} =$$

$$(1 + 21032656,1/1092997,06)^{1/4} = (20,24)^{1/4} = 2,12.$$

Мінімальна внутрішня економічна дохідність вкладених інвестицій $\tau_{мін}$ визначається за формулою:

$$\tau_{мін} = d + f, \quad (4.20)$$

де d – середньозважена ставка за депозитними операціями в комерційних банках; в 2023 році в Україні $d = 0,16$;

f – показник, що характеризує ризикованість вкладення інвестицій; зазвичай величина $f = 0,05 \dots 0,5$, але може бути і значно вищою. Прийmemo 0,35.

$\tau_{мін} = d + f = 0,16 + 0,35 = 0,51 < 2,12$, що свідчить про те, що внутрішня економічна дохідність інвестицій E_B , вища мінімальної внутрішньої дохідності.

Тобто, потенційних інвестор може бути зацікавлений у фінансуванні науково-технічної розробки за темою «Аналіз та прогнозування добового генерування електроенергії» та виведенні її на ринок, тобто в її комерціалізації.

Період окупності інвестицій $T_{ок}$ (DPP, DiscountedPaybackPeriod), які можуть бути вкладені потенційним інвестором у впровадження та комерціалізацію науково-технічної розробки:

$$T_{ок} = \frac{1}{E_B}, \quad (4.21)$$

де E_B – внутрішня економічна дохідність вкладених інвестицій.

$$T_{ок} = \frac{1}{2,12} = 0,47 \text{ р.}$$

Якщо $T_{ок} < 3$ -х років, то це свідчить про комерційну привабливість науково-технічної розробки і може спонукати потенційного інвестора профінансувати впровадження даної розробки та виведення її на ринок [115].

4.4 Висновки

Здійснені дослідження вказують на високий рівень комерційного потенціалу розробки, присвяченої темі "Аналіз та прогнозування добового генерування електроенергії", оцінений на рівні 39,3 балу. Це свідчить про значущі можливості для успішної комерціалізації даної науково-технічної розробки.

Особливо обнадійливим є той факт, що термін окупності даної розробки складає 0,47 року, що менше 3-х років. Це вказує на високу комерційну привабливість та швидке повернення інвестицій. Зазначений термін окупності може слугувати стимулом для потенційного інвестора розглядати можливість фінансування впровадження розробки та введення її на ринок.

Отже, на підставі отриманих результатів можна зробити висновок про доцільність проведення науково-технічної роботи в рамках теми " Аналіз та прогнозування добового генерування електроенергії". Успішна комерціалізація може призвести до значного успіху на ринку та забезпечити інвестора високим рівнем чистого прибутку у найкоротший термін.

ВИСНОВКИ

Роботу присвячено аналізу та прогнозуванню добового генерування електроенергії.

В роботі проведено аналіз проблематики генерацій електроенергії фотоелектричними електростанціями. Окреслено важливість прогнозування генерації електроенергії фотоелектричними електростанціями при роботі на об'єднану енергосистему. Розглянуто чинники які впливають на генерацію.

Також проведений огляд існуючих підходів та моделей машинного навчання в питанні прогнозування генерації електроенергії нестабільними джерелами енергії зокрема фотоелектростанціями.

Отже, враховуючи технологічну необхідність в балансуванні енергосистеми та імовірнісний характер генерації електроенергії фотоелектростанціями, які працюють в об'єднаній енергосистемі, виникає необхідність у точному короткостроковому прогнозі генерації електроенергії.

Відомі методи прогнозування варіюються в залежності від задачі прогнозування а саме довгострокове прогнозування та короткострокове прогнозування. Переважна більшість існуючих моделей прогнозування базуються зокрема: ARMA, ARIMA, ARMAX, штучна нейронна мережа ANN та адаптивна нейронечітка система логічного висновку ANFIS, рекурентні нейронні мережі (RNN). Недоліком існуючих моделей є те, що для навчання автори використовують дуже велику кількість даних, іншому випадку моделі не забезпечують необхідну точність прогнозу, а саме в межах 10%. У випадку зі збільшенням кількості даних, модель стає менш гнучкою, оскільки переважна більшість фотоелектричних станцій є новими або за період у 5-10 років неодноразово були модернізовані.

В подальшому проведено огляд вхідних даних щодо генерації електроенергії на ФЕС. Здійснено аналіз датасету та його форматування для забезпечення зручності подальших досліджень. Для оцінки датасету використані

статистичні показники, що вказують на особливості генерації електроенергії на ФЕС. Середнє значення генерації електроенергії за розглянутий період становить 151.24, вказуючи на узагальнене значення. Стандартне відхилення, що дорівнює 260.61, свідчить про значну варіацію в генерації, можливо, пов'язану з різними факторами, такими як погодні умови чи технічні аспекти станції. Медіана, рівна 0.0, вказує на наявність значень, що сконцентровані навколо нуля, що може пояснюватися наявністю багатьох нульових або дуже малих значень у ряді.

Використовуючи автокореляцію, було виявлено виражену сезонність у генерації електроенергії, обумовлену зміною світлового дня залежно від пори року. Графік автокореляції вказує на наявність регулярних циклічних коливань, що є характерним для сонячних електростанцій.

Для візуалізації розподілу даних використано гістограму та boxplot, які свідчать про відсутність суттєвих змін у генерації протягом трьох років. Діаграма розсіювання також не виявила яскравих викидів, проте присутні провали в генерації, що слід врахувати при побудові моделі прогнозування.

Результати розвідувального аналізу свідчать про стабільність генерації електроенергії на ФЕС з можливістю врахування сезонних коливань та провалів у деякі періоди.

На наступному етапі проведено налаштування моделі з врахуванням сезонності, такої як добова, тижнева та річна, що є ключовим етапом в побудові моделі.

Під час тренування моделі розраховано похибку прогнозування на валідаційному наборі. Отримана похибка прогнозування на рівні 6.98%. Низький рівень похибки говорить про те, що модель успішно враховує різноманітні аспекти, такі як сезонні коливання, зміни тренду та інші важливі фактори, що впливають на генерацію електроенергії. Отримані точні прогнози є вкрай корисними для стратегічного планування та управління енергетичними ресурсами фотоелектростанції. З ними можна ефективно регулювати роботу станції, максимізувати використання сонячної енергії та оптимізувати виробничі процеси

для досягнення максимальної продуктивності. Даний результат свідчить про правильний вибір налаштувань та правильний вибір методу прогнозування, доводячи, що модель Prophet виявляється ефективною для точного визначення генерації електроенергії на фотоелектростанціях. Результати роботи можуть бути універсалізовані та використані для прогнозування генерації на будь-яких нестабільних відновлюваних джерелах генерації електроенергії.

У розділі економічної оцінки проведено глибокий аналіз доцільності науково-дослідної роботи. Здійснені дослідження вказують на високий рівень комерційного потенціалу розробки, присвяченої темі "Аналіз та прогнозування добового генерування електроенергії", оцінений на рівні 39,3 балу. Це свідчить про значущі можливості для успішної комерціалізації даної науково-технічної розробки. Особливо обнадійливим є той факт, що термін окупності даної розробки складає 0,47 року, що менше 3-х років. Це вказує на високу комерційну привабливість та швидке повернення інвестицій. Зазначений термін окупності може слугувати стимулом для потенційного інвестора розглядати можливість фінансування впровадження розробки та введення її на ринок.

Отже, на підставі отриманих результатів можна зробити висновок про доцільність проведення науково-технічної роботи в рамках теми " Аналіз та прогнозування добового генерування електроенергії". Успішна комерціалізація може призвести до значного успіху на ринку та забезпечити інвестора високим рівнем чистого прибутку у найкоротший термін.

Отже, завдання магістерської кваліфікаційної роботи було виконано в повному обсязі.

Результати роботи доповідались з публікацією тез доповіді на ЛІІ Науково-технічній конференції факультету інтелектуальних інформаційних технологій та автоматизації ВНТУ (м. Вінниця, 2023 р.)

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. До уваги громадськості та експертів - проект Дорожньої карти розвитку відновлюваної енергетики України на період до 2020 року. *Державне агентство з енергоефективності та енергозбереження України*. URL: <http://saee.gov.ua/uk/pressroom/1133> (дата звернення: 01.11.2023).

2. Chowdhury A. A., Reliability Modeling of Distributed Generation in Conventional Distribution Systems Planning and Analysis. *IEEE Transactions on Industry Application*. 2003. Vol.39. No.5. P. 1493.1498. doi:10.1109/TIA.2003.816554. URL: https://www.researchgate.net/publication/3171659_Reliability_Modeling_of_Distributed_Generation_in_Conventional_Distribution_Systems_Planning_and_Analysis (дата звернення: 01.11.2023).

3. Bae I., J. Kim, Reliability Evaluation of Distributed Generation Based on Operation Mode. *IEEE Transactions on Power Systems*. 2007. Vol.22. No.2. P. 785.790. doi:10.1109/TPWRS.2007.894842. URL: https://www.researchgate.net/publication/3267813_Reliability_Evaluation_of_Distributed_Generation_Based_on_Operation_Mode (дата звернення: 03.11.2023).

4. R. Medeiros, X. Xu, E. Makram, Assessment of Operating Condition Dependent Reliability Indices in Microgrids. *Journal of Power and Energy Engineering*. 2016. No. 4. P. 56-66. doi:10.4236/jpee.2016.44006. URL: https://www.researchgate.net/publication/301740067_Assessment_of_Operating_Condition_Dependent_Reliability_Indices_in_Microgrids (дата звернення: 15.8.2023).

5. Kondo T, Jumpei Baba, Akihiko Yokoyama. Voltage control of distribution network with a large penetration of photovoltaic generations using facts devices. *Electrical Engineering in Japan*. 2008. Vol. 165. № 3. P. 16-28. doi:10.1002/ej.20499. URL: https://www.researchgate.net/publication/229904659_Voltage_Control_of_Distribution_Network_with_a_Large_Penetration_of_Photovoltaic_Generations_using_FACTS_Devices (дата звернення: 15.11.2023).

6. Закон України про "Про ринок електричної енергії" від 13.04.2017 № 2019.VIII. URL: [http://zakon5.rada.gov.Ua / laws / show/2019-19](http://zakon5.rada.gov.Ua/laws/show/2019-19)_(дата звернення: 15.11.2023).

7. Постанова НКРЕКП «Про затвердження Правил ринку «на добу наперед» та внутрішньодобового ринку» від 14.03.2018 № 308. URL: <https://zakon.rada.gov.ua/laws/show/v0308874-18#n9>_(дата звернення: 15.11.2023).

8. Бартецький А.А., Гнатюк М.Ю., Кручай І.С. Передбачення погодинної генерації електроенергії ФЕС з врахуванням прогнозу хмарності // *Матеріали ІІІ Науково-технічної конференції факультету інтелектуальних інформаційних технологій та автоматизації Вінницького національного технічного університету*, Вінниця, 21 – 23 червня 2023 р. – Електрон. текст. дані. – 2023. – URL: <https://conferences.vntu.edu.ua/index.php/all-fksa/all-fksa-2023/paper/view/18504/15>_(дата звернення: 25.11.2023).

9. Лежнюк П. Д., Комар В. О., Кравчук С. В., Дідіченко Є. С. Аналіз метеопараметрів для погодинного прогнозування виробітку електроенергії фотовольтаїчними електростанціями на добу наперед. *Енергетика та комп'ютерно-інтегровані технології в АПК*. – 2017. - № 1 (6). – С. 27-31. URL: <https://repo.btu.kharkov.ua/handle/123456789/6866>_(дата звернення: 25.11.2023).

10. Лежнюк, П. Д. Визначення оптимальної потужності резерву для забезпечення балансової надійності локальної електричної системи. *Вісник Нац. техн. ун-ту "ХПІ" : зб. наук. пр. Сер. : Нові рішення в сучасних технологіях*. Харків: НТУ "ХПІ", 2016. № 42 (1214). С. 69-75. URL: https://www.researchgate.net/publication/352758650_MATHEMATICAL_MODELING_OF_BATTERY_ENERGY_STORAGE_SYSTEMS_IN_THE_ADDITIONAL_SERVICE_MARKET_OF_THE_UNITED_ELECTRIC_POWER_SYSTEM_OF_UKRAINE_(дата звернення: 15.8.2023).

11. Д.Г. Дерев'янюк, Д.С. Горенко. Особливості побудови та функціонування віртуальних електростанцій в умовах розвитку ОЕС України. *Енергетика: економіка, технології, екологія*. 2016. № 3. С. 61-69. URL:

<https://ela.kpi.ua/handle/123456789/18721?locale=uk> (дата звернення: 25.11.2023).

12. Лежнюк П. Д., Бартецький А.А Бартецька І.А. Оптимізація роботи фотоелектричних станцій для забезпечення балансової стійкості енергосистеми. *Енергетика та комп'ютерно-інтегровані технології в АПК.* – 2017. - № 1 (6). – С. 27-31. URL:

<https://ir.lib.vntu.edu.ua/bitstream/handle/123456789/23353/%D0%9E%D0%BF%D1%82%D0%B8%D0%BC%D1%96%D0%B7%D0%B0%D1%86%D1%96%D1%8F%20%D1%80%D0%BE%D0%B1%D0%BE%D1%82%D0%B8%20%D0%A4%D0%95%D0%A1.pdf?sequence=1&isAllowed=y> (дата звернення: 25.11.2023).

13. Закон України про ринок електричної енергії (Відомості Верховної Ради (ВВР), 2017, №27-28, ст. 312). URL: <http://zakon2.rada.gov.ua/laws/show/2019-19>.

14. Характеристики напруги електропостачання в електричних мережах загальної призначеності. Національний стандарт України. Київ. Мінекономрозвитку України. 2014. 27 с. URL: <https://www.nerc.gov.ua/sferi-diyalnosti/elektroenergiya/yakist-elektropostachannya/yakist-elektrichnoyi-energiyi> (дата звернення: 25.11.2023).

15. Cody A. Hill ; Matthew Clayton Such ; Dongmei Chen ; Juan Gonzalez ; W. Mack Grady. Battery Energy Storage for Enabling Integration of Distributed Solar Power Generation. *IEEE Transactions on Smart Grid.* Vol:3, Issue:2, June 2012. Pp. 850- 857. URL: https://www.researchgate.net/publication/260581959_Battery_Energy_Storage_for_Enabling_Integration_of_Distributed_Solar_Power_Generation (дата звернення: 25.11.2023).

16. Б. І. Мокін, М. М. Чепурний, О. Б. Мокін. Повітряна акумулююча електростанція з двома повітросховищами різного тиску. *Наукові праці ВНТУ.* №1. 2008. С. 1-7. URL: <http://praci.vntu.edu.ua/index.php/praci/article/view/45> (дата звернення: 15.8.2023).

17. Syed Afroz. Notebook Kaggle "Solar Power Generation Forecast with 99% AUC". 2022. URL: <https://www.kaggle.com/code/pythonafroz/solar-power-generation-forecast-with-99-auc> (дата звернення: 25.11.2023).

18. Reikard, G. Hansen, C. Forecasting solar irradiance at short horizons: Frequency and time domain models. *Renew. Energy* 2019, 135, 1270–1290. URL: https://www.researchgate.net/publication/327506962_Forecasting_solar_irradiance_at_short_horizons_Frequency_and_time_domain_models (дата звернення: 15.8.2023).
19. Wang, J.; Li, Q.; Zeng, B. Multi-layer cooperative combined forecasting system for short-term wind speed forecasting. *Sustain. Energy Technol. Assess.* 2021, 43, 100946. URL: https://www.researchgate.net/publication/347888838_Multi-layer_cooperative_combined_forecasting_system_for_short-term_wind_speed_forecasting (дата звернення: 26.10.2023).
20. Akhter, M.N.; Mekhilef, S.; Mokhlis, H.; Shah, N.M. Review on forecasting of photovoltaic power generation based on machine learning and metaheuristic techniques. *IET Renew. Power Gener.* 2019, 13, 1009–1023. URL: https://www.researchgate.net/publication/330932621_A_review_on_Forecasting_of_Photovoltaic_Power_Generation_based_on_Machine_Learning_and_Metaheuristic_Techniques (дата звернення: 25.10.2023).
21. Pierro, M.; Bucci, F.; De Felice, M.; Maggioni, E.; Perotto, A.; Spada, F.; Moser, D.; Cornaro, C. Deterministic and stochastic approaches for day-ahead solar power forecasting. *J. Sol. Energy Eng.* 2017, 139, 021010. URL: <https://www.matteodefelice.name/files/pierro-deterministic.pdf> (дата звернення: 25.10.2023).
22. Prema, V.; Rao, K.U. Development of statistical time series models for solar power prediction. *Renew. Energy* 2015, 83, 100–109. URL: https://www.researchgate.net/publication/275719077_Development_of_statistical_time_series_models_for_solar_power_prediction (дата звернення: 25.10.2023).
23. Hirata, Y.; Aihara, K. Improving time series prediction of solar irradiance after sunrise: Comparison among three methods for time series prediction. *Sol. Energy* 2017, 149, 294–301. URL: https://www.researchgate.net/publication/316356067_Improving_time_series_prediction_of_solar_irradiance_after_sunrise_Comparison_among_three_methods_for_time_s

[eries_prediction](#)(дата звернення: 25.10.2023).

24. Shireen, T.; Shao, C.; Wang, H.; Li, J.; Zhang, X.; Li, M. Iterative multi-task learning for time-series modeling of solar panel PV outputs. *Appl. Energy* 2018, 212, 654–662. URL: https://www.researchgate.net/publication/322199306_Iterative_multi-task_learning_for_time-series_modeling_of_solar_panel_PV_outputs(дата звернення: 25.10.2023).

25. Bigdeli, N.; Salehi Borujeni, M.; Afshar, K. Time series analysis and short-term forecasting of solar irradiation, a new hybrid approach. *Swarm Evol. Comput.* 2017, 34, 75–88. URL: https://www.researchgate.net/publication/311960015_Time_Series_Analysis_and_Short-Term_Forecasting_of_Solar_Irradiation_a_New_Hybrid_Approach(дата звернення: 25.10.2023).

26. Wang, F.; Zhen, Z.; Liu, C.; Mi, Z.; Hodge, B.-M.; Shafie-khah, M.; Catalão, J.P.S. Image phase shift invariance based cloud motion displacement vector calculation method for ultra-short-term solar PV power forecasting. *Energy Convers. Manag.* 2018, 157, 123–135. URL: https://www.researchgate.net/publication/322195426_Image_phase_shift_invariance_based_cloud_motion_displacement_vector_calculation_method_for_ultra-short-term_solar_PV_power_forecasting(дата звернення: 25.10.2023).

27. Zaher, A.; Thil, S.; Nou, J.; Traoré, A.; Grieu, S. Comparative study of algorithms for cloud motion estimation using sky-imaging data. *IFAC-PapersOnLine* 2017, 50, 5934–5939. URL: https://www.researchgate.net/publication/320492776_Comparative_study_of_algorithms_for_cloud_motion_estimation_using_sky-imaging_data(дата звернення: 25.10.2023).

28. Cheng, H.-Y. Cloud tracking using clusters of feature points for accurate solar irradiance nowcasting. *Renew. Energy* 2017, 104, 281–289. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0960148116310710>(дата звернення: 25.10.2023).

29. Lima, F.J.L.; Martins, F.R.; Pereira, E.B.; Lorenz, E.; Heinemann, D. Forecast for surface solar irradiance at the Brazilian Northeastern region using NWP model and artificial neural networks. *Renew. Energy* 2016, 87, 807–818. URL: https://www.researchgate.net/publication/284167643_Forecast_for_surface_solar_irradiance_at_the_Brazilian_Northeastern_region_using_NWP_model_and_artificial_neural_networks(дата звернення: 25.10.2023).

30. Perez, R.; Lorenz, E.; Pelland, S.; Beauharnois, M.; Van Knowe, G.; Hemker, K.; Heinemann, D.; Remund, J.; Müller, S.C.; Traunmüller, W.; et al. Comparison of numerical weather prediction solar irradiance forecasts in the US, Canada and Europe. *Sol. Energy* 2013, 94, 305.326. URL: https://www.researchgate.net/publication/305482662_Comparison_of_global_horizontal_irradiance_forecasts_based_on_numerical_weather_prediction_models_with_different_spatio-temporal_resolutions_Comparison_of_irradiance_forecasts_based_on_NWP_models(дата звернення: 25.10.2023).

31. Verzijlbergh, R.A.; Heijnen, P.W.; de Roode, S.R.; Los, A.; Jonker, H.J.J. Improved model output statistics of numerical weather prediction based irradiance forecasts for solar power applications. *Sol. Energy* 2015, 118, 634–645. URL: https://www.researchgate.net/publication/281086370_Improved_model_output_statistics_of_numerical_weather_prediction_based_irradiance_forecasts_for_solar_power_applications(дата звернення: 25.10.2023).

32. Mathiesen, P.; Collier, C.; Kleissl, J. A high-resolution, cloud-assimilating numerical weather prediction model for solar irradiance forecasting. *Sol. Energy* 2013, 92, 47–61. URL: https://escholarship.org/content/qt1rm8p16v/qt1rm8p16v_noSplash_9e689e8c81526d5eedf905a9719b032a.pdf(дата звернення: 25.10.2023).

33. Hussain, S.; AlAlili, A. A hybrid solar radiation modeling approach using wavelet multiresolution analysis and artificial neural networks. *Appl. Energy* 2017, 208, 540–550. URL:

<https://www.sciencedirect.com/science/article/abs/pii/S030626191731379X?via%3Dihub> (дата звернення: 25.10.2023).

34. Хуе, Х. Prediction of daily diffuse solar radiation using artificial neural networks. *Int. J. Hydrog. Energy* 2017, 42, 28214–28221. URL: https://www.researchgate.net/publication/320440545_Prediction_of_daily_diffuse_solar_radiation_using_artificial_neural_networks (дата звернення: 25.10.2023).

35. Alzahrani, A.; Shamsi, P.; Dagli, C.; Ferdowsi, M. Solar Irradiance Forecasting Using Deep Neural Networks. *Procedia Comput. Sci.* 2017, 114, 304–313. URL: <https://www.sciencedirect.com/science/article/pii/S1877050917318392> (дата звернення: 25.10.2023).

36. Bou-Rabee, M.; Sulaiman, S.A.; Saleh, M.S.; Marafi, S. Using artificial neural networks to estimate solar radiation in Kuwait. *Renew. Sustain. Energy Rev.* 2017, 72, 434–438. URL: <https://www.sciencedirect.com/science/article/abs/pii/S1364032117300138> (дата звернення: 25.10.2023).

37. Olatomiwa, L.; Mekhilef, S.; Shamshirband, S.; Petković, D.J.R.; Reviews, S.E. Adaptive neuro-fuzzy approach for solar radiation prediction in Nigeria. *Renew. Sustain. Energy Rev.* 2015, 51, 1784–1791. URL: https://www.researchgate.net/publication/277214041_Adaptive_neuro-fuzzy_approach_for_solar_radiation_prediction_in_Nigeria (дата звернення: 25.10.2023).

38. Dolara, A.; Grimaccia, F.; Leva, S.; Mussetta, M.; Ogliari, E. A physical hybrid artificial neural network for short term forecasting of PV plant power output. *Energies* 2015, 8, 1138–1153. URL: https://www.researchgate.net/publication/272415201_A_Physical_Hybrid_Artificial_Neural_Network_for_Short_Term_Forecasting_of_PV_Plant_Power_Output (дата звернення: 25.10.2023).

39. De Giorgi, M.; Malvoni, M.; Congedo, P. Comparison of strategies for multi-step ahead photovoltaic power forecasting models based on hybrid group method of data

handling networks and least square support vector machine. *Energy* 2016, 107, 360–373. URL: https://www.researchgate.net/publication/301797929_Comparison_of_strategies_for_multi-step_ahead_photovoltaic_power_forecasting_models_based_on_hybrid_group_method_of_data_handling_networks_and_least_square_support_vector_machine (дата звернення: 25.10.2023).

40. Jang, H.S.; Bae, K.Y.; Park, H.-S.; Sung, D.K. Solar power prediction based on satellite images and support vector machine. *IEEE Trans. Sustain. Energy* 2016, 7, 1255–1263. URL: <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9702145> (дата звернення: 29.10.2023).

41. Da Silva Fonseca, J.G., Jr.; Oozeki, T.; Ohtake, H.; Shimose, K.-i.; Takashima, T.; Ogimoto, K. Forecasting regional photovoltaic power generation—a comparison of strategies to obtain one-day-ahead data. *Energy Procedia* 2014, 57, 1337–1345. URL: https://www.researchgate.net/publication/285834802_Forecasting_Regional_Photovoltaic_Power_Generation_-_A_Comparison_of_Strategies_to_Obtain_One-Day-Ahead_Data (дата звернення: 29.10.2023).

42. Wolff, B.; Kühnert, J.; Lorenz, E.; Kramer, O.; Heinemann, D. Comparing support vector regression for PV power forecasting to a physical modeling approach using measurement, numerical weather prediction, and cloud motion data. *Sol. Energy* 2016, 135, 197–208. URL: https://www.researchgate.net/publication/303819783_Comparing_support_vector_regression_for_PV_power_forecasting_to_a_physical_modeling_approach_using_measurement_numerical_weather_prediction_and_cloud_motion_data (дата звернення: 29.10.2023).

43. Tang, P.; Chen, D.; Hou, Y. Entropy method combined with extreme learning machine method for the short-term photovoltaic power generation forecasting. *Chaos Solitons Fractals* 2016, 89, 243–248. URL: https://www.researchgate.net/publication/287120101_Entropy_method_combined_with_extreme_learning_machine_method_for_the_short-term_photovoltaic_power_generation_forecasting

[term photovoltaic power generation forecasting](#) (дата звернення: 29.10.2023).

44. Hossain, M.; Mekhilef, S.; Danesh, M.; Olatomiwa, L.; Shamshirband, S. Application of extreme learning machine for short term output power forecasting of three grid-connected PV systems. *J. Clean. Prod.* 2017, 167, 395–405. URL: https://www.researchgate.net/publication/319064122_Application_of_Extreme_Learning_Machine_for_short_term_output_power_forecasting_of_three_grid-connected_PV_systems (дата звернення: 29.10.2023).

45. Deo, R.C.; Downs, N.; Parisi, A.V.; Adamowski, J.F.; Quilty, J.M. Very short-term reactive forecasting of the solar ultraviolet index using an extreme learning machine integrated with the solar zenith angle. *Environ. Res.* 2017, 155, 141–166. URL: https://www.researchgate.net/publication/313888809_Very_short-term_reactive_forecasting_of_the_solar_ultraviolet_index_using_an_extreme_learning_machine_integrated_with_the_solar_zenith_angle (дата звернення: 29.10.2023).

46. Youssef, A.; El-Telbany, M.; Zekry, A. The role of artificial intelligence in photo-voltaic systems design and control: A review. *Renew. Sustain. Energy Rev.* 2017, 78, 72–79. URL: <https://www.sciencedirect.com/science/article/abs/pii/S1364032117305555> (дата звернення: 29.10.2023).

47. Wang, H.; Lei, Z.; Zhang, X.; Zhou, B.; Peng, J. A review of deep learning for renewable energy forecasting. *Energy Convers. Manag.* 2019, 198, 111799. URL: https://www.researchgate.net/publication/336184094_A_review_of_deep_learning_for_renewable_energy_forecasting (дата звернення: 29.10.2023).

48. Narvaez, G.; Giraldo, L.F.; Bressan, M.; Pantoja, A. Machine learning for site-adaptation and solar radiation forecasting. *Renew. Energy* 2021, 167, 333–342. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0960148120318395> (дата звернення: 29.10.2023).

49. Blaga, R.; Sabadus, A.; Stefu, N.; Dughir, C.; Paulescu, M.; Badescu, V. A current perspective on the accuracy of incoming solar energy forecasting. *Energy Combust. Sci.* 2019, 70, 119–144. URL:

<https://www.sciencedirect.com/science/article/abs/pii/S0360128518300303> (дата звернення: 29.10.2023).

50. Turchenko, V.A.; Trukhanov, S.V.; Kostishin, V.G.; Damay, F.; Porcher, F.; Klygach, D.S.; Vakhitov, M.G.; Lyakhov, D.; Michels, D.; Bozzo, B.; et al. Features of structure, magnetic state and electrodynamic performance of SrFe_{12-x}In_xO₁₉. *Sci. Rep.* 2021, 11, 18342. URL: <https://www.nature.com/articles/s41598-021-97684-8> (дата звернення: 29.10.2023).

51. Kozlovskiy, A.L.; Alina, A.; Zdorovets, M.V. Study of the effect of ion irradiation on increasing the photocatalytic activity of WO₃ microparticles. *J. Mater. Sci. Mater. Electron.* 2021, 32, 3863–3877. URL: https://www.researchgate.net/publication/348428993_Study_of_the_effect_of_ion_irradiation_on_increasing_the_photocatalytic_activity_of_WO3_microparticles (дата звернення: 29.10.2023).

52. Zambrano, A.F.; Giraldo, L.F. Solar irradiance forecasting models without on-site training measurements. *Renew. Energy* 2020, 152, 557–566. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0960148120301142> (дата звернення: 05.11.2023).

53. Dorado-Moreno, M.; Navarin, N.; Gutiérrez, P.A.; Prieto, L.; Sperduti, A.; Salcedo-Sanz, S.; Hervás-Martínez, C. Multi-task learning for the prediction of wind power ramp events with deep neural networks. *Neural Netw.* 2020, 123, 401–411. URL: https://www.researchgate.net/publication/338427143_Multi-task_learning_for_the_prediction_of_wind_power_ramp_events_with_deep_neural_networks (дата звернення: 05.11.2023).

54. Aslam, M.; Lee, J.-M.; Kim, H.-S.; Lee, S.-J.; Hong, S. Deep learning models for long-term solar radiation forecasting considering microgrid installation: A comparative study. *Energies* 2019, 13, 147. URL: https://www.researchgate.net/publication/338206094_Deep_Learning_Models_for_Long-Term_Solar_Radiation_Forecasting_Considering_Microgrid_Installation_A_Compara

[tive Study](#) (дата звернення: 05.11.2023).

55. Rana, M.; Rahman, A. Multiple steps ahead solar photovoltaic power forecasting based on univariate machine learning models and data re-sampling. *Sustain. Energy Grids Netw.* 2020, 21, 100286. URL: <https://www.sciencedirect.com/science/article/abs/pii/S2352467719304023> (дата звернення: 05.11.2023).

56. Al-Dahidi, S.; Ayadi, O.; Adeeb, J.; Louzazni, M. Assessment of artificial neural networks learning algorithms and training datasets for solar photovoltaic power production prediction. *Front. Energy Res.* 2019, 7, 130. URL: <https://www.sciencedirect.com/science/article/pii/S0960148123002811> (дата звернення: 05.11.2023).

57. Wuwei, W.; Jianfeng, C.; Bo, W.; Shuangli, F. A solar power prediction using support vector machines based on multi-source data fusion. In Proceedings of the 2018 International Conference on Power System Technology (POWERCON), *Guangzhou, China*, 6.8 November 2018; pp. 4573-4577. URL: https://www.researchgate.net/publication/330595757_A_Solar_Power_Prediction_Using_Support_Vector_Machines_Based_on_Multi-source_Data_Fusion (дата звернення: 05.11.2023).

58. Ahmad, M.W.; Mourshed, M.; Rezgui, Y. Tree-based ensemble methods for predicting PV power generation and their comparison with support vector regression. *Energy* 2018, 164, 465–474. URL: https://www.researchgate.net/publication/327332475_Tree-based_ensemble_methods_for_predicting_PV_power_generation_and_their_comparison_with_support_vector_regression (дата звернення: 05.11.2023).

59. Munawar, U.; Wang, Z. A framework of using machine learning approaches for short-term solar power forecasting. *J. Electr. Eng. Technol.* 2020, 15, 561–569. URL: https://www.researchgate.net/publication/338557008_A_Framework_of_Using_Machine_Learning_Approaches_for_Short-Term_Solar_Power_Forecasting

(дата звернення: 05.11.2023).

60. Andrade, J.R.; Bessa, R.J. Improving renewable energy forecasting with a grid of numerical weather predictions. *IEEE Trans. Sustain. Energy* 2017, 8, 1571–1580.

URL:

https://www.researchgate.net/publication/316355623_Improving_Renewable_Energy_Forecasting_With_a_Grid_of_Numerical_Weather_Predictions

(дата звернення: 05.11.2023).

61. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* 2015, 521, 436–444.

URL: https://www.researchgate.net/publication/277411157_Deep_Learning

(дата звернення: 05.11.2023).

62. Sun, Y.; Venugopal, V.; Brandt, A.R. Short-term solar power forecast with deep learning: Exploring optimal input and output configuration. *Sol. Energy* 2019, 188, 730–741. URL:

https://www.researchgate.net/publication/334841983_Short-term_solar_power_forecast_with_deep_learning_Exploring_optimal_input_and_output_configuration (дата звернення: 05.11.2023).

63. Ghimire, S.; Deo, R.C.; Raj, N.; Mi, J. Deep learning neural networks trained with MODIS satellite-derived predictors for long-term global solar radiation prediction. *Energies* 2019, 12, 2407. URL:

https://www.researchgate.net/publication/333976099_Deep_Learning_Neural_Networks_Trained_with_MODIS_Satellite-Derived_Predictors_for_Long-Term_Global_Solar_Radiation_Prediction (дата звернення: 08.11.2023).

64. Zhang, C.; Liang, M.; Song, X.; Liu, L.; Wang, H.; Li, W.; Shi, M. Generative adversarial network for geological prediction based on TBM operational data. *Mech. Syst. Signal Processing* 2020, 162, 108035. URL:

https://www.researchgate.net/publication/351870082_Generative_adversarial_network_for_geological_prediction_based_on_TBM_operational_data

(дата звернення: 08.11.2023).

65. Zhang, J.; Verschae, R.; Nobuhara, S.; & Lalonde, J.F. Deep photovoltaic nowcasting. *Sol. Energy* 2018, 176, 267–276. URL:

https://www.researchgate.net/publication/329343823_Deep_photovoltaic_nowcasting

66. Qing, X.; Niu, Y. Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. *Energy* 2018, 148, 461–468. URL:

https://www.researchgate.net/publication/322911289_Hourly_day-ahead_solar_irradiance_prediction_using_weather_forecasts_by_LSTM

(дата звернення: 08.11.2023).

67. He, W. Load forecasting via deep neural networks. *Procedia Comput. Sci.* 2017.

URL: <https://www.sciencedirect.com/science/article/pii/S1877050917326170>

(дата звернення: 08.11.2023).

68. Taieb, S.B.; Bontempi, G.; Atiya, A.F.; Sorjamaa, A. A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition. *Expert Syst. Appl.* 2012, 39, 7067–7083. URL:

<https://www.sciencedirect.com/science/article/abs/pii/S0957417412000528>

(дата звернення: 08.11.2023).

69. Santhosh, M.; Venkaiah, C.; Kumar, D.V. Short-term wind speed forecasting approach using ensemble empirical mode decomposition and deep Boltzmann machine. *Sustain. Energy Grids Netw.* 2019, 19, 100242. URL:

https://www.researchgate.net/publication/335182153_Short-term_wind_speed_forecasting_approach_using_Ensemble_Empirical_Mode_Decomposition_and_Deep_Boltzmann_Machine (дата звернення: 08.11.2023).

70. Wang, L.; Li, X.; Bai, Y. Short-term wind speed prediction using an extreme learning machine model with error correction. *Energy Convers. Manag.* 2018, 162, 239–250. URL:

https://www.researchgate.net/publication/323387819_Short-term_wind_speed_prediction_using_an_extreme_learning_machine_model_with_error_correction (дата звернення: 08.11.2023).

71. Golyandina, N.; Nekrutkin, V.; Zhigljavsky, A.A. *Analysis of Time Series. SSA and Related Techniques.* URL:

https://www.researchgate.net/publication/27648878_Analysis_of_Time_Series_Struct

[ure SSA and Related Techniques](#) (дата звернення: 08.11.2023).

72. Broomhead, D.S.; King, G.P. Extracting qualitative dynamics from experimental data. *Phys. D Nonlinear Phenom.* 20, 217-236. URL: https://www.researchgate.net/publication/216638258_Broomhead_DS_King_GP_Extracting_qualitative_dynamics_from_experimental_data_Physica_D_20_217-236 (дата звернення: 08.11.2023).

73. Moreno, S.R.; dos Santos Coelho, L. Wind speed forecasting approach based on singular spectrum analysis and adaptive neuro fuzzy inference system. *Renew. Energy* 2018, 126, 736-754. URL: https://www.researchgate.net/publication/321582781_Wind_speed_forecasting_approach_based_on_Singular_Spectrum_Analysis_and_Adaptive_Neuro_Fuzzy_Inference_System (дата звернення: 08.11.2023).

74. Liu, H.; Mi, X.; Li, Y.; Duan, Z.; Xu, Y. Smart wind speed deep learning based multi-step forecasting model using singular spectrum analysis, convolutional Gated Recurrent Unit network and Support Vector Regression. *Renew. Energy* 2019, 143, 842–854. URL: https://www.researchgate.net/publication/333137465_Smart_wind_speed_deep_learning_based_multi-step_forecasting_model_using_singular_spectrum_analysis_convolutional_Gated_Recurrent_Unit_network_and_Support_Vector_Regression (дата звернення: 08.11.2023).

75. Wang, C.; Zhang, H.; Ma, P. Wind power forecasting based on singular spectrum analysis and a new hybrid Laguerre neural network. *Appl. Energy* 2020, 259, 114139. URL: https://www.researchgate.net/publication/337469003_Wind_power_forecasting_based_on_singular_spectrum_analysis_and_a_new_hybrid_Laguerre_neural_network (дата звернення: 08.11.2023).

76. Moreno, S.R.; da Silva, R.G.; Mariani, V.C.; dos Santos Coelho, L. Multi-step wind speed forecasting based on hybrid multi-stage decomposition model and long short-term memory neural network. *Energy Convers. Manag.* 2020, 213, 112869. URL:

<https://www.sciencedirect.com/science/article/pii/S2352484722024374>

(дата звернення: 08.11.2023).

77. Chen, C.; Duan, S.; Cai, T.; Liu, B. Online 24-h solar power forecasting based on weather type classification using artificial neural network. *Sol. Energy* 2011, 85, 2856–2870. URL: https://www.researchgate.net/publication/259575631_Online_24-h_solar_power_forecasting_based_on_weather_type_classification_using_artificial_neural_network (дата звернення: 08.11.2023).

78. Wang, F.; Zhen, Z.; Wang, B.; Mi, Z. Comparative study on KNN and SVM based weather classification models for day ahead short term solar PV power forecasting. *Appl. Sci.* 2017, 8, 28. URL: https://www.researchgate.net/publication/322067009_Comparative_Study_on_KNN_and_SVM_Based_Weather_Classification_Models_for_Day_Ahead_Short_Term_Solar_PV_Power_Forecasting (дата звернення: 08.11.2023).

79. Lin, P.; Peng, Z.; Lai, Y.; Cheng, S.; Chen, Z.; Wu, L. Short-term power prediction for photovoltaic power plants using a hybrid improved Kmeans-GRA-Elman model based on multivariate meteorological factors and historical power datasets. *Energy Convers. Manag.* 2018, 177, 704–717. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0196890418311130> (дата звернення: 08.11.2023).

80. Gu, B.; Shen, H.; Lei, X.; Hu, H.; Liu, X. Forecasting and uncertainty analysis of day-ahead photovoltaic power using a novel forecasting method. *Appl. Energy* 2021, 299, 117291. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0306261921007054> (дата звернення: 08.11.2023).

81. Zhang, N.; Ren, Q.; Liu, G.; Guo, L.; Li, J. Short-term PV Output Power Forecasting Based on CEEMDAN-AE-GRU. *J. Electr. Eng. Technol.* 2022, 17, 1183–1194. URL: https://www.researchgate.net/publication/370925371_Short-Term_Prediction_of_PV_Power_Based_on_Combined_Modal_Decomposition_and_NARX-LSTM-LightGBM (дата звернення: 08.11.2023).

82. Nam, K.; Hwangbo, S.; Yoo, C. A deep learning-based forecasting model for renewable energy scenarios to guide sustainable energy policy: A case study of Korea. *Renew. Sustain. Energy Rev.* 2020, 122, 109725. URL: <https://www.sciencedirect.com/science/article/abs/pii/S1364032120300228> (дата звернення: 08.11.2023).

83. Hong, T.; Fan, S. Probabilistic electric load forecasting: A tutorial review. *Forecast.* 2016, 32, 914.938. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0169207015001508> (дата звернення: 08.11.2023).

84. Вох, G.E.; Jenkins, G.M.; Reinsel, G.C.; Ljung, G.M. *Time Series Analysis: Forecasting and Control.* 2015. URL: http://repo.darmajaya.ac.id/4781/1/Time%20Series%20Analysis_%20Forecasting%20and%20Control%20%28%20PDFDrive%20%29.pdf (дата звернення: 08.11.2023).

85. Pappas, S.S.; Ekonomou, L.; Moussas, V.; Karampelas, P.; Katsikas, S. Adaptive load forecasting of the Hellenic electric grid. *Zhejiang Univ. Sci. A* 2008, 9, 1724.1730. URL: https://www.researchgate.net/publication/225534972_Adaptive_load_forecasting_of_the_Hellenic_electric_grid (дата звернення: 08.11.2023).

86. Hippert, H.S.; Pedreira, C.E.; Souza, R.C. Neural networks for short-term load forecasting: A review and evaluation. *IEEE Trans. Power Syst.* 2001, 16, 44–55. URL: <https://ieeexplore.ieee.org/document/910780> (дата звернення: 11.11.2023).

87. Karampelas, P.; Vita, V.; Pavlatos, C.; Mladenov, V.; Ekonomou, L. Design of artificial neural network models for the prediction of the Hellenic energy consumption. In Proceedings of the 10th Symposium on Neural Network Applications in Electrical Engineering. *IEEE.* 2010; pp. 41–44. URL: https://www.researchgate.net/publication/270452389_Design_of_artificial_neural_network_models_for_the_prediction_of_the_Hellenic_energy_consumption (дата звернення: 11.11.2023).

88. Ekonomou, L.; Christodoulou, C.; Mladenov, V. A short-term load forecasting

method using artificial neural networks and wavelet analysis. *Int. J. Power Syst* 2016, 1, 64–68. URL: <https://www.ias.ac.org/iasac/filedownloads/ijps/2016/010-0010.pdf> (дата звернення: 11.11.2023).

89. Chen, B.J.; Chang, M.W. Load forecasting using support vector machines: A study on EUNITE competition 2001. *IEEE Trans. Power Syst.* 2004, 19, 1821–1830. URL: https://www.researchgate.net/publication/3267168_Load_forecasting_using_support_vector_machines_A_study_on_EUNITE_Competition_2001 (дата звернення: 11.11.2023).

90. Taylor, J.W.; McSharry, P.E. Short-term load forecasting methods: An evaluation based on european data. *IEEE Trans. Power Syst.* 2007, 22, 2213–2219. URL: <https://ieeexplore.ieee.org/document/4349130> (дата звернення: 11.11.2023).

91. Papaioannou, G.; Dikaiakos, C.; Dramountanis, A.; Papaioannou, P. Analysis and modeling for short-to medium-term load forecasting using a hybrid manifold learning principal component model and comparison with classical statistical models (SARIMAX, Exponential Smoothing) and artificial intelligence models (ANN, SVM): The case of Greek electricity market. *Energies* 2016, 9, 635. URL: <https://ideas.repec.org/a/gam/jeners/v9y2016i8p635-d76042.html> (дата звернення: 11.11.2023).

92. Kuo, P.H.; Huang, C.J. A high precision artificial neural networks model for short-term energy load forecasting. *Energies* 2018, 11, 213. URL: https://www.researchgate.net/publication/322547645_A_High_Precision_Artificial_Neural_Networks_Model_for_Short-Term_Energy_Load_Forecasting (дата звернення: 11.11.2023).

93. Kiprijanovska, I.; Stankoski, S.; Ilievski, I.; Jovanovski, S.; Gams, M.; Gjoreski, H. Houseec: Day-ahead household electrical energy consumption forecasting using deep learning. *Energies* 2020, 13, 2672. URL: https://www.researchgate.net/publication/341645597_HouseEC_Day-Ahead_Household_Electrical_Energy_Consumption_Forecasting_Using_Deep_Learn

[ng](#) (дата звернення: 11.11.2023).

94. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* 1997, 9, 1735–1780. URL:

https://www.researchgate.net/publication/13853244_Long_Short-term_Memory

(дата звернення: 11.11.2023).

95. Bianchi, F.M.; Maiorino, E.; Kampffmeyer, M.C.; Rizzi, A.; Jenssen, R. Recurrent Neural Networks for Short-Term Load Forecasting. *An Overview and Comparative Analysis*; Germany, 2017. URL:

https://www.researchgate.net/publication/320970557_Recurrent_Neural_Networks_for_Short-Term_Load_Forecasting_An_Overview_and_Comparative_Analysis

(дата звернення: 11.11.2023).

96. Kong, W.; Dong, Z.Y.; Jia, Y.; Hill, D.J.; Xu, Y.; Zhang, Y. Short-term residential load forecasting based on LSTM recurrent neural network. *IEEE Trans. Smart Grid* 2017, 10, 841.851. URL: <https://ieeexplore.ieee.org/document/8039509>

97. Bouktif, S.; Fiaz, A.; Ouni, A.; Serhani, M.A. Single and multi-sequence deep learning models for short and medium term electric load forecasting. *Energies* 2019, 12, 149. URL: <https://ideas.repec.org/a/gam/jeners/v12y2019i1p149-d194483.html>

(дата звернення: 11.11.2023).

98. Tian, C.; Ma, J.; Zhang, C.; Zhan, P. A deep neural network model for short-term load forecast based on long short-term memory network and convolutional neural network. *Energies* 2018, 11, 3493. URL:

https://www.researchgate.net/publication/329661648_A_Deep_Neural_Network_Model_for_Short-Term_Load_Forecast_Based_on_Long_Short-Term_Memory_Network_and_Convolutional_Neural_Network

(дата звернення: 11.11.2023).

99. He, W. Load forecasting via deep neural networks. *Procedia Comput. Sci.* 2017, 122, 308.314. URL: <https://www.sciencedirect.com/science/article/pii/S1877050917326170> (дата звернення: 11.11.2023).

100. Taylor, J.W. Short-term electricity demand forecasting using double seasonal

exponential smoothing. *J. Oper. Res. Soc.* 2003, 54, 799-805. URL: https://www.researchgate.net/publication/228392213_Short-term_electricity_demand_forecasting_using_double_seasonal_exponential_smoothing (дата звернення: 11.11.2023).

101. Hyndman, R.J.; Athanasopoulos, G. *Forecasting: Principles and Practice*; OTexts: Melbourne, Australia, 2018. URL: <https://robjhyndman.com/uwafiles/fpp-notes.pdf> (дата звернення: 11.11.2023).

102. Zhang, G.P.; Qi, M. Neural network forecasting for seasonal and trend time series. *Eur. Oper. Res.* 2005, 160, 501–514. URL: https://www.researchgate.net/publication/4871997_Neural_network_forecasting_for_seasonal_and_trend_time_series (дата звернення: 11.11.2023).

103. Taieb, S.B.; Bontempi, G.; Atiya, A.F.; Sorjamaa, A. A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition. *Expert Syst. Appl.* 2012, 39, 7067–7083. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0957417412000528> (дата звернення: 12.11.2023).

104. Cleveland, R.B.; Cleveland, W.S.; McRae, J.E.; Terpenning, I. STL: A seasonal.trend decomposition. *Off. Stat.* 3.73. URL: <https://www.wessa.net/download/stl.pdf> (дата звернення: 12.11.2023).

105. De Livera, A.M.; Hyndman, R.J.; Snyder, R.D. Forecasting time series with complex seasonal patterns using exponential smoothing. *J. Am. Stat. Assoc.* 2011, 106, 1513-1527. URL: <https://www.jstor.org/stable/23239555> (дата звернення: 12.11.2023).

106. Elsner, J.B.; Tsonis, A.A. *Singular Spectrum Analysis: A New Tool in Time Series Analysis*. Springer Science & Business Media: New York, NY, USA, 2013. URL: https://www.researchgate.net/publication/260124592_Singular_Spectrum_Analysis_for_Time_Series (дата звернення: 12.11.2023).

107. Bandara, K.; Bergmeir, C.; Hewamalage, H. LSTM-MSNet: Leveraging forecasts on sets of related time series with multiple seasonal patterns. *IEEE Trans.*

Neural Netw. Learn. Syst. 2020, 32, 1586.1599. URL: https://www.researchgate.net/publication/335737890_LSTM-MSNet_Leveraging_Forecasts_on_Sets_of_Related_Time_Series_with_Multiple_Seasonal_Patterns (дата звернення: 12.11.2023).

108. Wang, C.; Zhang, H.; Ma, P. Wind power forecasting based on singular spectrum analysis and a new hybrid Laguerre neural network. *Appl. Energy* 2020, 259, 114139. URL: https://www.researchgate.net/publication/337469003_Wind_power_forecasting_based_on_singular_spectrum_analysis_and_a_new_hybrid_Laguerre_neural_network (дата звернення: 12.11.2023).

109. Moreno, S.R.; dos Santos Coelho, L. Wind speed forecasting approach based on singular spectrum analysis and adaptive neuro fuzzy inference system. *Renew. Energy* 2018, 126, 736.754. URL: <https://www.sciencedirect.com/science/article/pii/S2352484723015834> (дата звернення: 12.11.2023).

110. Liu, H.; Mi, X.; Li, Y.; Duan, Z.; Xu, Y. Smart wind speed deep learning based multi-step forecasting model using singular spectrum analysis, convolutional Gated Recurrent Unit network and Support Vector Regression. *Renew. Energy* 2019, 143, 842.854. URL: https://www.researchgate.net/publication/333137465_Smart_wind_speed_deep_learning_based_multi-step_forecasting_model_using_singular_spectrum_analysis_convolutional_Gated_Recurrent_Unit_network_and_Support_Vector_Regression (дата звернення: 12.11.2023).

111. Liu, H.; Mi, X.; Li, Y. Smart multi-step deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysis, LSTM network and ELM. *Energy Convers. Manag.* 2018, 159, 54–64. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0960148119306998> (дата звернення: 12.11.2023).

112. Moreno, S.R.; da Silva, R.G.; Mariani, V.C.; dos Santos Coelho, L. Multi-step wind speed forecasting based on hybrid multi-stage decomposition model and long

short-term memory neural network. *Energy Convers. Manag.* 2020, 213, 112869. URL: https://www.researchgate.net/publication/340853440_Multi-step_wind_speed_forecasting_based_on_hybrid_multi-stage_decomposition_model_and_long_short-term_memory_neural_network (дата звернення: 12.11.2023).

113. Taylor, S. J., & Letham, B. (2018). Forecasting at scale. *The American Statistician*, 72(1), 37–45.: <http://lethalletham.com/ForecastingAtScale.pdf> (дата звернення: 14.11.2023).

114. Vitalii Mokin. Kaggle Notebook "COVID in UA: Prophet with 4, Nd seasonality". 2022. https://www.kaggle.com/code/vbmokin/covid-in-ua-prophet-with-4-nd-seasonality/notebook?scriptVersionId=60244069__Prophet URL: – Режим доступу до ресурсу: <https://facebook.github.io/prophet/> (дата звернення: 14.11.2023).

115. Козловський В. О., Лесько О. Й., Кавецький В. В. Методичні вказівки до виконання економічної частини магістерських кваліфікаційних робіт : уклад. Вінниця : ВНТУ, 2021. 42 с. URL: https://pdf.lib.vntu.edu.ua/books/2022/Kozlovskii_2021_43.pdf (дата звернення: 16.11.2023).

Додаток А
Міністерство освіти і науки України
Вінницький національний технічний університет
Факультет інтелектуальних інформаційних технологій та автоматизації

ЗАТВЕРДЖУЮ
Завідувач кафедри САІТ
_____ д.т.н., проф. Віталій МОКІН
«__» _____ 2023 року

ТЕХНІЧНЕ ЗАВДАННЯ
на комплексну магістерську кваліфікаційну роботу
АНАЛІЗ ТА ПЕРЕДБАЧЕННЯ ДОБОВОГО ГЕНЕРУВАННЯ
ЕЛЕКТРОЕНЕРГІЇ ФОТОЕЛЕКТРОСТАНЦІЄЮ. ЧАСТИНА 1. АНАЛІЗ ТА
ПРОГНОЗУВАННЯ ДОБОВОГО ГЕНЕРУВАННЯ ЕЛЕКТРОЕНЕРГІЇ
08-34.МКР.001.00.000 ТЗ

Керівник: д.т.н., проф.
_____ Олександр МОКІН
«__» _____ 2023 р.
Розробив студент гр. 2ІСТ-22м
_____ Андрій БАРТЕЦЬКИЙ
«__» _____ 2023 р.

1. Підстава для проведення робіт

Підставою для виконання роботи є наказ № __ по ВНТУ від «__» _____ 2023 р., та індивідуальне завдання на МКР, затверджене протоколом № __ засідання кафедри САІТ від «__» _____ 2023 р.

2. Джерела розробки:

– Liu, H.; Mi, X.; Li, Y. Smart multi-step deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysis, LSTM network and ELM. *Energy Convers. Manag.* 2018, 159, 54–64. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0960148119306998>

– Cleveland, R.B.; Cleveland, W.S.; McRae, J.E.; Terpenning, I. STL: A seasonal-trend decomposition. *J. Off. Stat.* 1990, 6, 3–73. URL: <https://www.wessa.net/download/stl.pdf>

– De Livera, A.M.; Hyndman, R.J.; Snyder, R.D. Forecasting time series with complex seasonal patterns using exponential smoothing. *J. Am. Stat. Assoc.* 2011, 106, 1513–1527. URL: <https://www.jstor.org/stable/23239555>

– Vitalii Mokin. Kaggle Notebook "COVID in UA: Prophet with 4, Nd seasonality". 2022. <https://www.kaggle.com/code/vbmokin/covid-in-ua-prophet-with-4-nd-seasonality/notebook?scriptVersionId=60244069> Prophet URL: <https://facebook.github.io/prophet/>

3. Метою дослідження є підвищення точності прогнозування генерування електроенергії ФЕС за рахунок використання методів машинного навчання.

4. Вихідні дані для проведення робіт: Kaggle Dataset “energy” <https://www.kaggle.com/datasets/andriibartetskyi/energy> наданий, з дозволу адміністрації фотоелектричної станції, розташованої у Вінницькій області.

5. Методи дослідження: Використання моделей машинного навчання, а саме моделі Prophet від Facebook

6. Етапи роботи і терміни їх виконання:

1. Загальна характеристика об’єкту досліджень 01.09–30.09

2. Вибір оптимальних налаштувань для розв’язання поставленої задачі 01.10 – 09.10

3. Розроблення інформаційної технології прогнозування генерації електроенергії на ФЕС 10.10 – 28.10

4. Економічна частина 29.10 – 03.11

5. Оформлення пояснювальної записки 04.11 – 25.11

6. Очікувані результати та порядок реалізації:

Аналіз та прогнозування добового генерування електроенергії, результатом якої є підвищення точності прогнозування генерування електроенергії ФЕС за рахунок використання методів машинного навчання.

8. Вимоги до розробленої документації

Текстова та ілюстративна частини роботи оформлені у відповідності до вимог «Методичних вказівок до виконання магістерських кваліфікаційних робіт для студентів спеціальності 126 «Інформаційні системи та технології» (освітня програма «Інформаційні технології аналізу даних та зображень»).

9. Порядок приймання роботи

Публічний захист«19» грудня 2023 р.
Початок розробки «01» вересня 2023 р.
Граничні терміни виконання МКР «30» листопада 2023 р.

Розробив студент групи 2ІСТ–22м _____ Андрій БАРТЕЦЬКИЙ

Додаток Б

Протокол перевірки магістерської кваліфікаційної роботи на наявність текстових запозичень

Назва роботи: «Аналіз та передбачення добового генерування електроенергії фотоелектростанцією. Частина 1. Аналіз та прогнозування добового генерування електроенергії»

Тип роботи: магістерська кваліфікаційна робота

Підрозділ: кафедра САІТ

Показники звіту подібності Unicheck

Оригінальність 95,25%

Схожість 4,75%

Аналіз звіту подібності (відмітити потрібне)

- Запозичення, виявлені у роботі, оформлені коректно і не містять ознак плагіату.
 - Виявлені у роботі запозичення не мають ознак плагіату, але їх надмірна кількість викликає сумніви щодо цінності роботи і самостійності її автора. Роботу направити на розгляд експертної комісії кафедри.
 - Виявлені у роботі запозичення є недобросовісними і мають ознаки плагіату та/або в ній містяться навмисні спотворення тексту, що вказують на спроби приховування недобросовісних запозичень.

Особа, відповідальна за перевірку

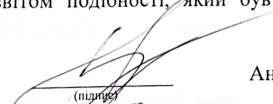


(підпис)

Сергій ЖУКОВ

Ознайомлені з повним звітом подібності, який був згенерований системою Unicheck щодо роботи.

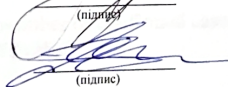
Автор роботи



(підпис)

Андрій БАРТЕЦЬКИЙ

Керівник роботи



(підпис)

Олександр МОКІН

Додаток В
Лістинг програми

```
# Work with Data - the main Python libraries
import numpy as np
import pandas as pd

# For import data from API
import requests

# Visualization
import matplotlib.pyplot as plt

# Modeling and Prediction
from fbprophet import Prophet
from sklearn.metrics import mean_absolute_error, r2_score

import warnings
warnings.simplefilter('ignore')

prediction_period = 48 # Period for prediction, half an hour

# Load the data from the local CSV file
df_data = pd.read_csv('./input/energy/energy_data_set.csv', sep=';')

# Print the dataframe
print(df_data)

# Display the last 10 rows of the dataframe "df_data"
df_data.tail(50)

# Calculation daily data for number of confirmed cases
df_data['y'] = df_data['y'].diff().abs()
df_data['y'] = df_data['y']

# Delete the last zero value
data = data[:-1]
data.tail(30)
# Build the plot
plt.figure(figsize=(16, 6))
plt.plot(data['ds'], data['y'])
plt.legend(['y'], loc='upper left', fontsize=12, frameon=False)
step = 10000
```

```

plt.xticks(range(0, len(data['ds']), step), data['ds'][::step])
plt.xlabel('Date',fontsize=14)
plt.ylabel('Power (y)',fontsize=14)
plt.show()

# Preparing data for modeling with Prophet
data.columns = ['ds', 'y']
data.tail()

# Devide dataset to training and validation datasets
valid = data[(len(data)-prediction_period):].reset_index(drop=True)
data = data[:(len(data)-prediction_period)]
valid

data.tail(prediction_period)

# Build Prophet model
model = Prophet()

# Training model for all data
model.fit(data)

# Build Prophet model with parameters ans structure
model2 = Prophet(daily_seasonality=True, weekly_seasonality=True,
yearly_seasonality=True,
    changepoint_range=1, changepoint_prior_scale = 0.3)
model2.add_seasonality(name='weekly', period=7, fourier_order=12,
    mode = 'multiplicative', prior_scale = 0.24)
model2.add_seasonality(name='triplly', period=3, fourier_order=2,
    mode = 'multiplicative', prior_scale = 0.15)

def prediction_accuracy(valid, model):
    # Calculation prediction accuracy for valid dataset by model

    future = model.make_future_dataframe(periods = prediction_period) # Valid data
    prediction
    forecast = model.predict(future)
    forecast = forecast[(len(forecast)-prediction_period):]
    forecast[forecast['yhat'] < 0]['yhat'] = 0
    y_val = forecast['yhat'].round() # Prediction
    y_target = valid['y'] # Real data

    return round(mean_absolute_error(y_val, y_target)*100/valid['y'].sum(), 2)

```

```

# Make a forecast for prediction_period days ahead
future = model.make_future_dataframe(periods = prediction_period)
forecast = model.predict(future)

# Make values integer, and replace negative values with zero
feature_all = ['yhat_lower', 'yhat', 'yhat_upper']
forecast[feature_all] = forecast[feature_all].astype('int')
for feature in feature_all:
    forecast.loc[forecast[feature] < 0, feature] = 0

# Draw plot of the values with forecasting data
data['ds'] = pd.to_datetime(data['ds'])
forecast['ds'] = pd.to_datetime(forecast['ds'])

plt.figure(figsize=(10, 6))
plt.plot(data['ds'], data['y'], color='blue', label='Actual')
plt.plot(forecast['ds'], forecast['yhat'], color='red', label='Forecast')
plt.fill_between(forecast['ds'], forecast['yhat_lower'], forecast['yhat_upper'],
color='red', alpha=0.2)
plt.xlabel('Date',fontsize=14)
plt.ylabel('Power',fontsize=14)
plt.legend(loc='upper left')
plt.show()

# Draw plot with the components (trend and weekly seasonality) of the forecasts
figure_component = model.plot_components(forecast)
figure_component.axes[0].lines[0].set_color('blue')
figure_component.axes[1].lines[0].set_color('red')
figure_component.axes[2].lines[0].set_color('green')

# Output the prediction for the next prediction_period days
forecast[['ds', 'yhat_lower', 'yhat', 'yhat_upper']].tail(prediction_period)

print(f'Relative error (WAPE) for model = {prediction_accuracy(valid, model)}%")

```

Додаток Г

ІЛЮСТРАТИВНА ЧАСТИНА

АНАЛІЗ ТА ПЕРЕДБАЧЕННЯ ДОБОВОГО ГЕНЕРУВАННЯ
ЕЛЕКТРОЕНЕРГІЇ ФОТОЕЛЕКТРОСТАНЦІЄЮ. ЧАСТИНА 3. АНАЛІЗ
ВПЛИВУ СОНЯЧНОЇ АКТИВНОСТІ

Нормоконтроль: к.т.н., доцент

_____ Сергій ЖУКОВ

«__» _____ 2023 р.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	
	id	DATE	P	h13	h14	h15	h16	h17	h18	h19	h20	h21	h22	h23	h24	h25	h26	h27	h28
2	1080113	31.12.2022	812	0	0	0	0	0	5	38	37	67	75	98	90	75	83	69	
3	1661195	07.11.2022	3170	0	0	3	18	14	22	91	187	219	155	33	719	490	247	558	
4	1659626	06.11.2022	1803	0	0	0	3	9	23	54	77	70	94	153	232	202	273	185	
5	1658057	05.11.2022	938	0	0	0	4	30	59	60	52	69	71	73	87	74	50	34	
6	1656450	04.11.2022	982	0	0	0	38	93	139	175	7	0	0	0	0	0	0	0	
7	1654843	03.11.2022	1739	0	0	2	18	28	54	111	293	412	165	523	88	42	0	0	
8	1652863	02.11.2022	582	0	0	0	6	32	36	36	35	69	49	65	77	42	0	21	
9	1651294	01.11.2022	5149	0	0	22	83	301	472	577	659	720	762	785	460	0	0	0	
0	1649687	31.10.2022	605	0	0	0	21	133	0	0	0	0	0	0	0	0	0	0	
1	1648080	30.10.2022	4674	0	0	0	33	108	83	83	121	157	180	262	198	534	754	328	
2	1646435	29.10.2022	5867	0	0	0	0	30	116	278	261	508	536	677	690	717	505	240	
3	1644828	28.10.2022	7693	0	0	0	1	37	72	221	234	326	515	527	673	628	548	694	
4	1643221	27.10.2022	4677	0	0	0	0	1	15	30	45	64	96	126	584	771	662	256	
5	1641614	26.10.2022	3237	0	0	0	0	0	19	53	104	82	120	100	125	211	196	235	
6	1640007	25.10.2022	562	0	0	0	0	6	48	59	49	63	83	32	28	21	16	6	
7	1638400	24.10.2022	5756	0	0	0	0	47	125	139	246	453	487	426	638	411	394	636	

Рисунок Г.1 – Вхідний датасет генерації електроенергії на ФЕС

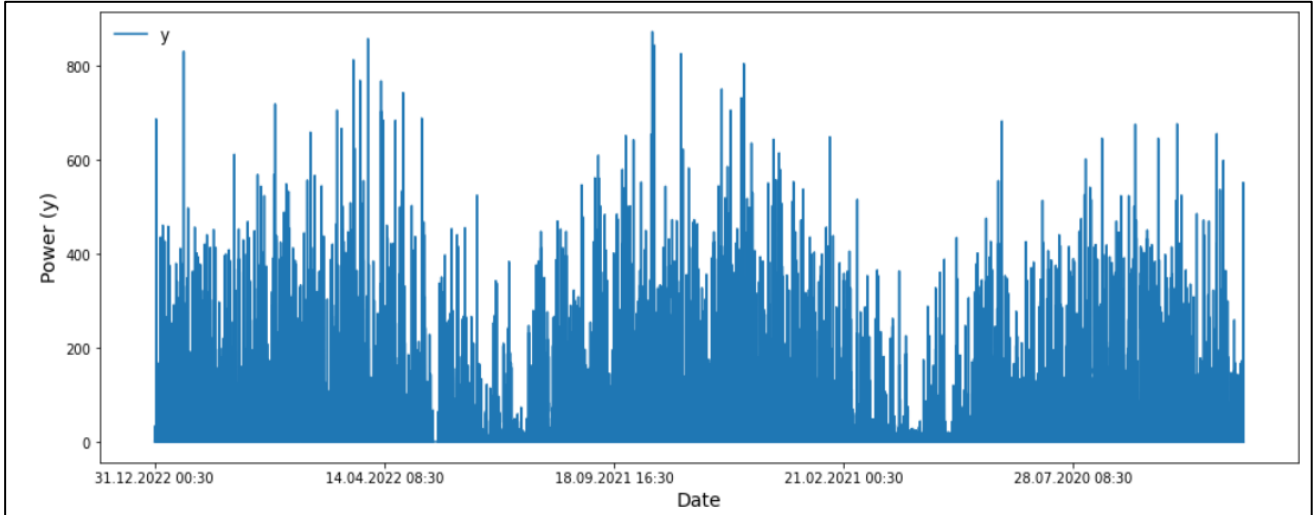


Рисунок Г.2 – Графік генерації електроенергії ФЕС за увесь період спостережень

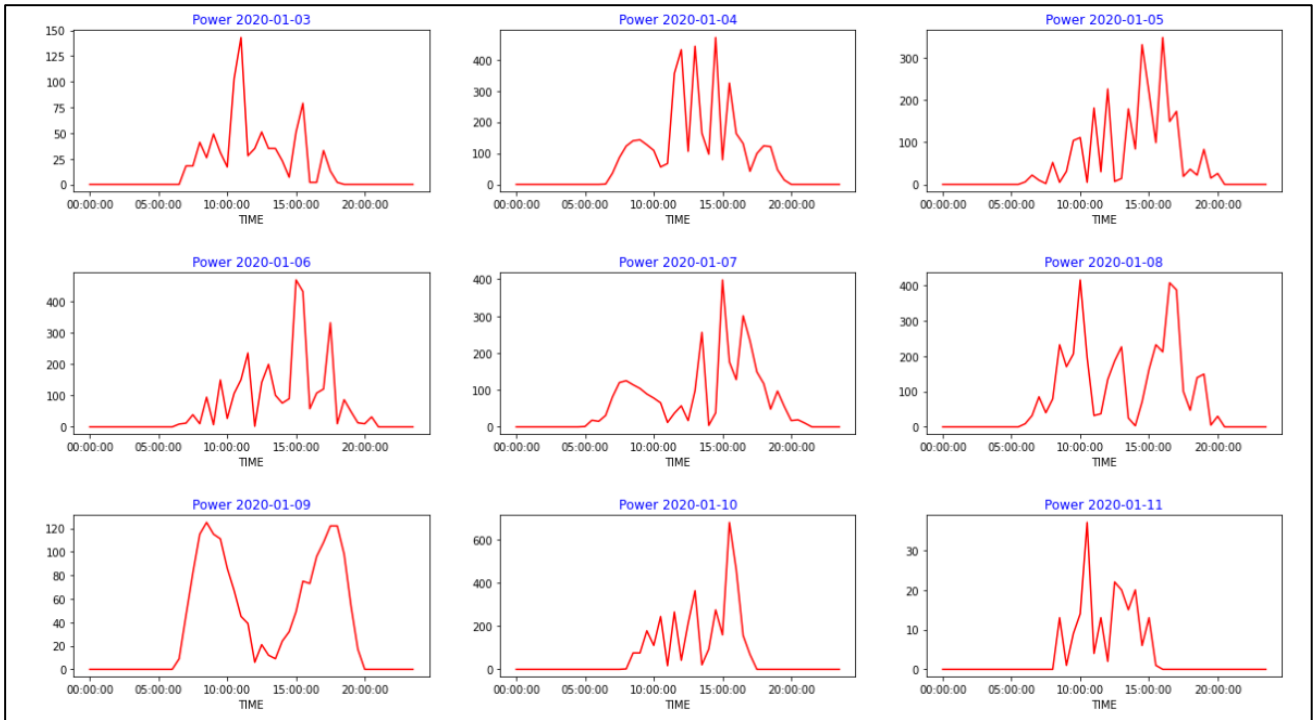


Рисунок Г.3 – Графіки генерації електроенергії ФЕС в межах доби

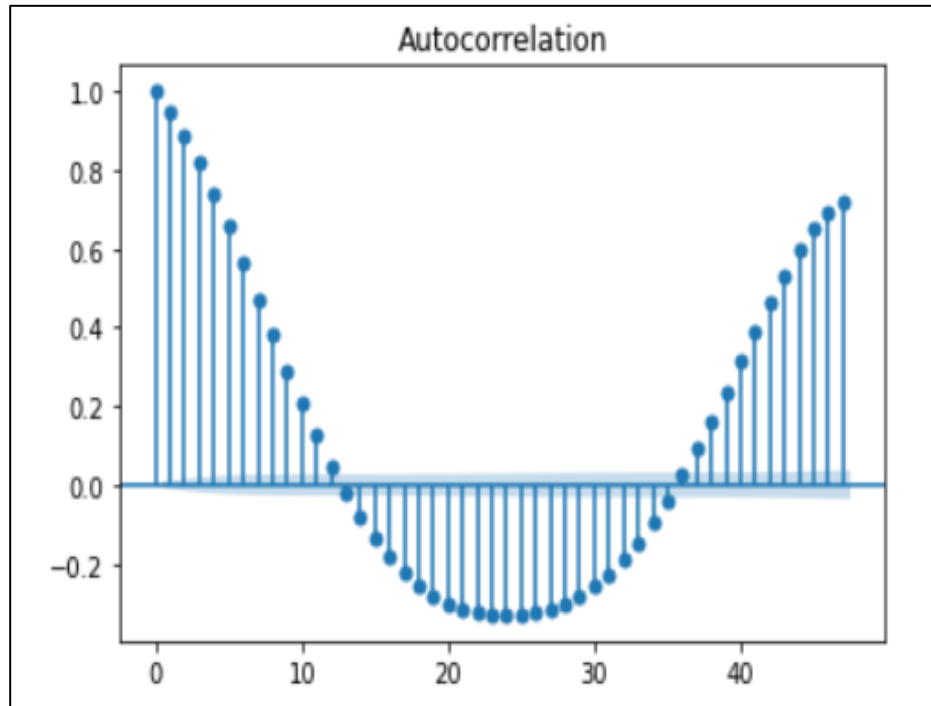


Рисунок Г.4 – Графік автокореляції датасету генерації електроенергії ФЕС

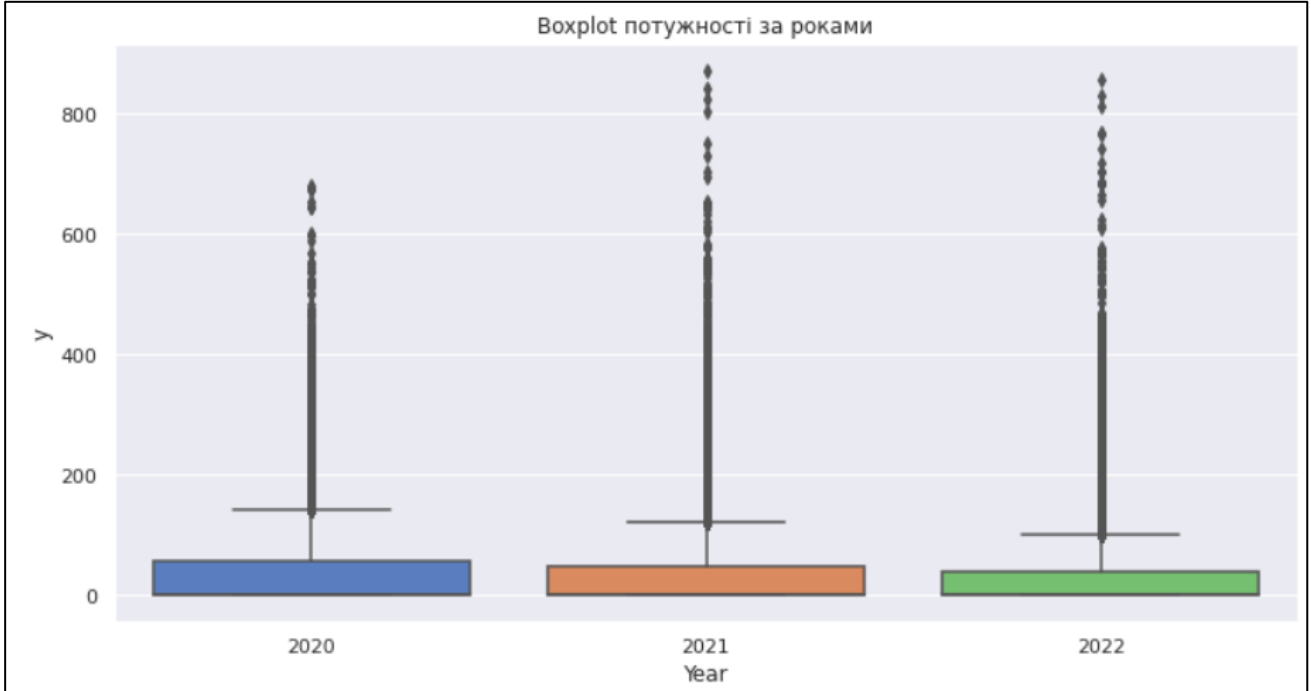


Рисунок Г.5 – Вохplot датасету генерації електроенергії ФЕС

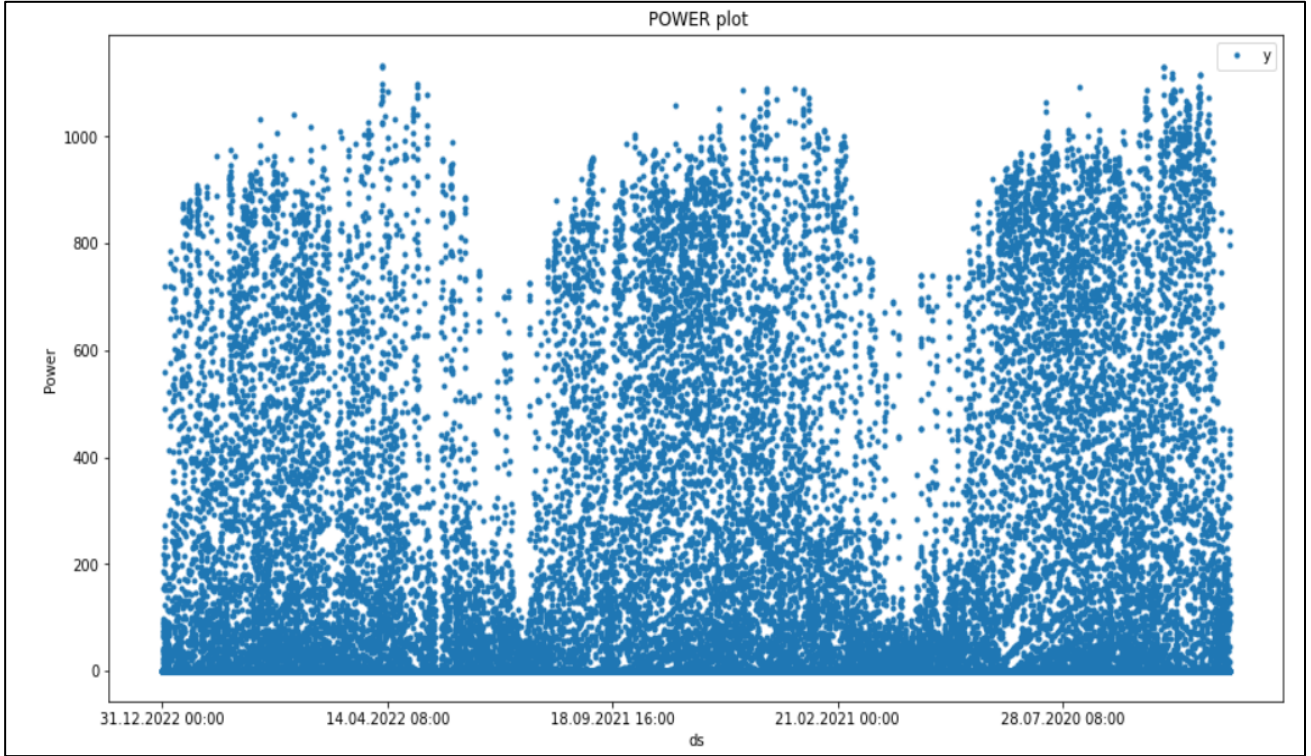


Рисунок Г.6 – Діаграма розсіювання датасету генерації електроенергії ФЕС

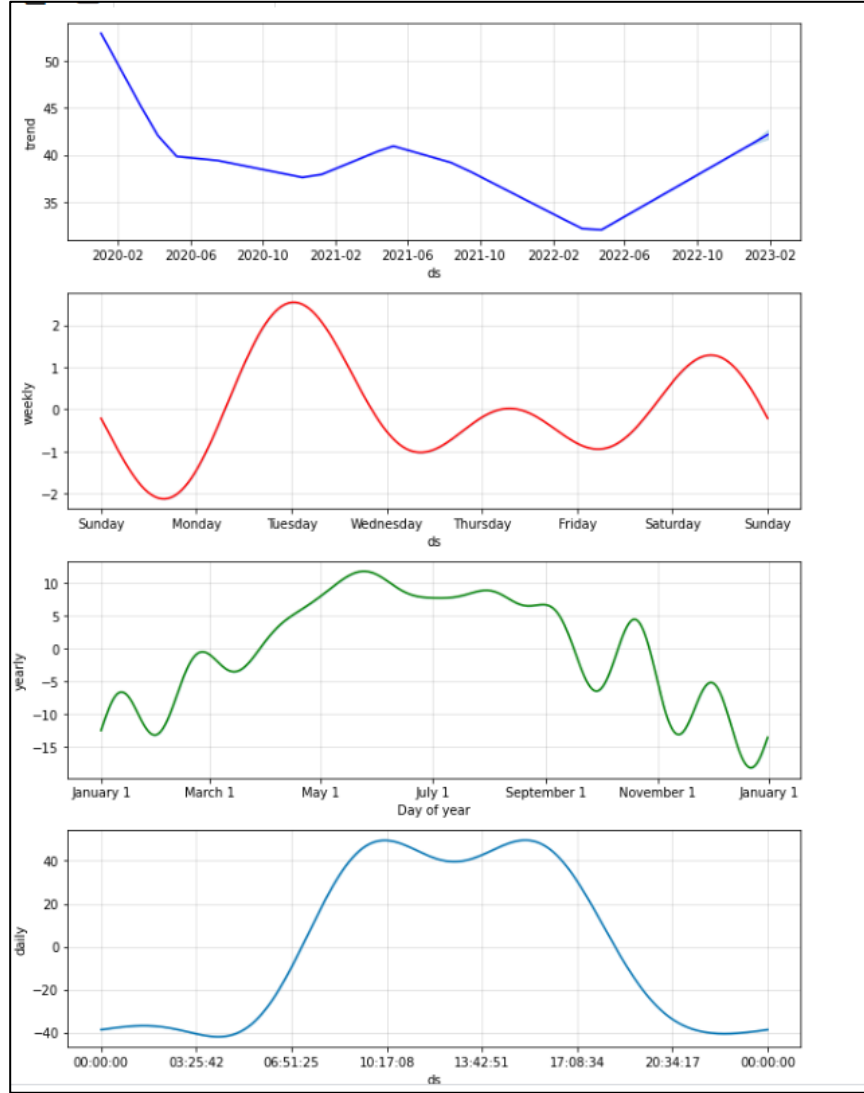


Рисунок Г.7 – Графіки річної, тижневої, сезонної, а також денної циклічності

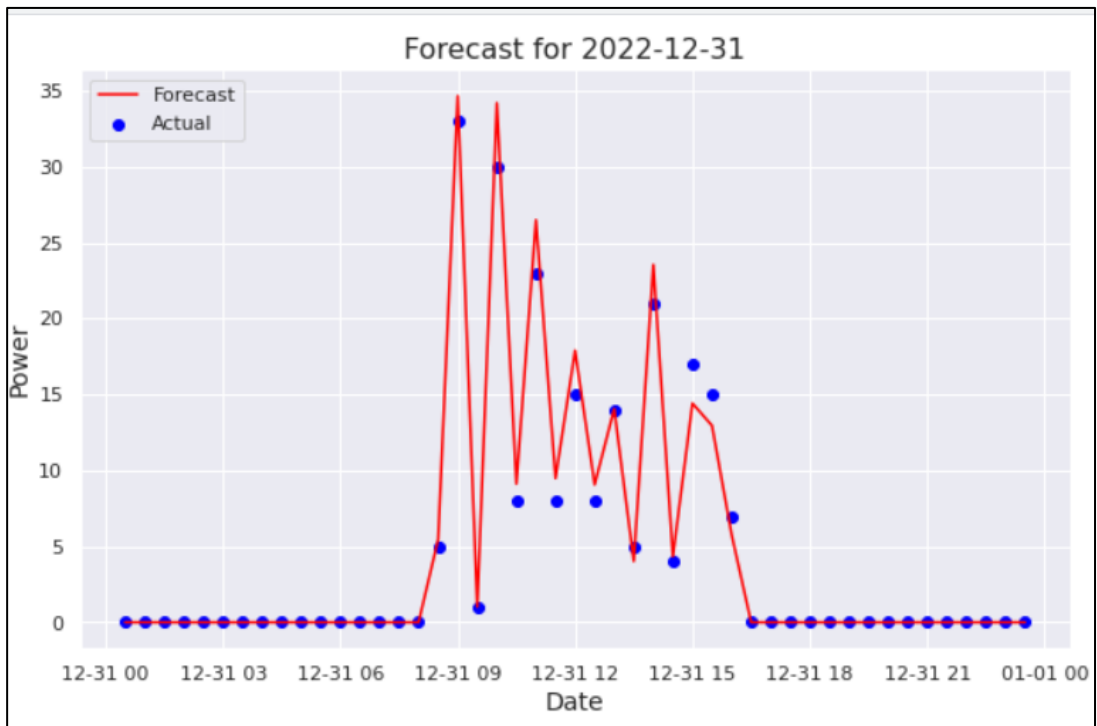


Рисунок Г.8 – Графік прогнозування генерації електроенергії на добу