

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**  
**НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ**  
Факультет кібербезпеки, комп'ютерної та програмної інженерії  
Кафедра прикладної математики

ДОПУСТИТИ ДО ЗАХИСТУ

Завідувач кафедри

\_\_\_\_\_ П. О. Приставка

“ \_\_\_\_\_ ” \_\_\_\_\_ 2020 р.

# **ДИПЛОМНА РОБОТА**

**(ПОЯСНЮВАЛЬНА ЗАПИСКА)**

**ВИПУСКНИКА ОСВІТНЬО–КВАЛІФІКАЦІЙНОГО РІВНЯ**  
**«МАГІСТР»**

**Тема:** Інформаційна технологія аналізу динаміки цін ринку  
електричної енергії країн Євросоюзу

**Виконавець:**

Глуцук. О. О.

**Керівник:**

д-р тех. наук, проф. Приставка П. О.

**Нормоконтролер:**

доцент Іллічева Л.М.

**Київ 2020**

# НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ

**Факультет** Кібербезпеки, комп'ютерної та програмної інженерії  
**Кафедра** Прикладної математики  
**Спеціальність** 113 «Прикладна математика»

ЗАТВЕРДЖУЮ

Завідувач кафедри

\_\_\_\_\_ П. О. Приставка

“ \_\_\_\_ ” \_\_\_\_\_ 2020 р.

## ЗАВДАННЯ

**на виконання дипломної роботи  
студентки Глущук Ольги Олегівни**

**1. Тема дипломної роботи:** Інформаційна технологія аналізу динаміки цін ринку електричної енергії країн Євросоюзу, затверджена наказом ректора від "04" листопада 2019 р. № 2476/ст.

**2. Термін виконання роботи:** з "15" жовтня 2020 р. по "09" лютого 2020 р.

**3. Вихідні дані до роботи:** дані з угорського оператора системи передачі електричної енергії, файл, у якому записані святкові та неробочі дні в Угорщині за досліджуваний проміжок часу, доступ до Інтернету, програмне середовище розробки RStudio.

**4. Зміст пояснювальної записки:**

- Огляд моделей прогнозування часових рядів.
- Опис вхідних даних. Дослідження взаємодії цільових показників і факторів.
- Розробка інформаційної технології короткострокового прогнозування ціни та споживання на електроенергію.
- Програмна реалізація ІТ. Тестування та аналіз результатів.

## 5. Календарний план-графік виконання дипломної роботи

| № п/п | Етапи виконання дипломної роботи  | Термін виконання | Відмітка про виконання |
|-------|---|------------------|------------------------|
| 1.    | Отримання індивідуального завдання  | 04.11.2019       | Виконано               |
| 2.    | Уточнення постановки задачі   | 07.11.2019       | Виконано               |
| 3.    | Аналіз ринку електричної енергії.<br>Проведення критичного аналізу літератури за темою прогнозування                      | 12.11.2019       | Виконано               |
| 4.    | Аналіз існуючих технологій прогнозування  | 15.11.2019       | Виконано               |
| 5.    | Розробка алгоритму та створення інформаційної технології з вирішення питання прогнозування споживання електричної енергії | 20.11.2019       | Виконано               |
| 6.    | Розробка алгоритму та створення інформаційної технології з вирішення питання прогнозування цін на РДН                     | 25.11.2019       | Виконано               |
| 7.    | Проведення тестування на реальних даних. Обчислення похибки.  | 14.12.2019       | Виконано               |
| 8.    | Оформлення пояснювальної записки  | 20.12.2019       | Виконано               |
| 9.    | Створення презентації   | 13.01.2020       | Виконано               |
| 10.   | Отримання рецензії від опонентів  | 23.01.2020       | Виконано               |
| 11.   | Захист дипломної роботи   | 05.02.2020       |                        |

6. Дата видачі завдання: "04" листопада 2019 р.

Керівник кваліфікаційної роботи: \_\_\_\_\_

П. О. Приставка

Завдання прийняв до виконання: \_\_\_\_\_

О. О. Глушук

## РЕФЕРАТ

Пояснювальна записка до дипломної роботи на тему: «Інформаційна технологія аналізу динаміки цін ринку електричної енергії країн Євросоюзу»: 90 сторінок, 48 рисунків, 11 таблиць, 60 використаних джерел.

ЕЛЕКТРИЧНА ЕНЕРГІЯ, СПОЖИВАННЯ ЕЛЕКТРОЕНЕРГІЇ, СТРУКТУРА ВИРОБНИЦТВА ЕЛЕКТРОЕНЕРГІЇ, РИНОК НА ДОБУ НАПЕРЕД, ЧАСОВИЙ РЯД, ПРОГНОЗУВАННЯ ЧАСОВОГО РЯДУ, ЛІНІЙНА РЕГРЕСІЯ, КУБІЧНА РЕГРЕСІЯ, КОРЕЛЯЦІЯ, АВТОРЕГРЕСІЯ, ПОХИБКА ПРОГНОЗУВАННЯ.

**Об'єктом дослідження** є процес підготовки та обробки даних ринку електричної енергії.

**Предметом дослідження** є методи і засоби аналізу та прогнозування часових рядів.

**Програмне середовище**, що використовувалося в процесі виконання дипломної роботи: RStudio.

**Мета:** розв'язання актуальної задачі зі створення інформаційної технології прогнозування цін ринку електричної енергії країн Євросоюзу на прикладі угорського ринку на добу наперед.

**Результатом** виконання дипломної роботи є, розроблена на мові R, інформаційна технологія для аналізу та прогнозування споживання та ціни на електричну енергію.

**Галузь застосування:** результати дипломної роботи можуть бути використані операторами системи передач для прогнозування споживання на добу наперед з метою покращення роботи генерації та енергосистеми в цілому; трейдерами для прогнозування ціни на РДН з метою отримання прибутку.

## ЗМІСТ

|  |    |
|--|----|
| РЕФЕРАТ .....  | 4  |
| ЗМІСТ .....  | 5  |
| ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ .....  | 7  |
| ВСТУП .....  | 8  |
| ПОСТАНОВКА ЗАДАЧІ .....  | 14 |
| РОЗДІЛ 1 .....   | 16 |
| ОСНОВНІ ТЕОРЕТИЧНІ ПОЛОЖЕННЯ .....   | 16 |
| 1.1.    Основні характеристики області енергетики країн Євросоюзу .....                | 16 |
| 1.2.    Аналіз існуючих моделей короткострокового прогнозування цін в енергетиці ..... | 24 |
| 1.3.    Огляд математичних моделей прогнозування часових рядів .....                   | 28 |
| 1.3.1.    Регресійні моделі .....  | 29 |
| 1.3.2.    Авторегресійні моделі .....  | 30 |
| 1.3.3.    Моделі експоненціального згладжування .....                                  | 33 |
| 1.3.4.    Моделі на основі нейронних мереж .....                                       | 34 |
| 1.3.5.    Моделі на базі ланцюгів Маркова .....  | 35 |
| 1.3.6.    Інші моделі та методи прогнозування .....                                    | 36 |
| 1.4.    Оцінка помилки прогнозування часового ряду .....                               | 37 |
| Висновки до розділу 1 .....  | 39 |
| РОЗДІЛ 2 .....   | 40 |
| ОПИС РОЗРОБЛЕНОЇ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ .....  | 40 |
| 2.1.    Опис та підготовка вхідних даних .....   | 41 |
| 2.2.    Модель прогнозування споживання електроенергії .....                           | 46 |
| 2.2.1.    Первинний аналіз вхідних даних .....   | 46 |
| 2.2.2.    Побудова загальної моделі прогнозування споживання .....                     | 50 |
| 2.2.3.    Формалізація моделі та уточнення параметрів .....                            | 52 |
| 2.2.4.    Приклад роботи. Аналіз результатів .....                                     | 61 |
| 2.3.    Модель прогнозування ціни на електроенергію на добу наперед ..                 | 66 |

|               |   |           |
|---------------|---|-----------|
| <b>2.3.1.</b> | <b>Первинний аналіз даних .....</b>                       | <b>66</b> |
| <b>2.3.2.</b> | <b>Побудова загальної моделі прогнозування ціни .....</b> | <b>69</b> |
| <b>2.3.3.</b> | <b>Формалізація моделі та уточнення параметрів.....</b>   | <b>71</b> |
| <b>2.3.4.</b> | <b>Приклад роботи. Аналіз результатів .....</b>           | <b>75</b> |
| <b>2.4.</b>   | <b>Алгоритм ідентифікації моделі.....</b>                 | <b>79</b> |
| <b>2.5.</b>   | <b>Опис мови та середовища програмування .....</b>        | <b>80</b> |
| <b>2.6.</b>   | <b>Висновки до розділу 2.....</b>                         | <b>82</b> |
|               | <b>ВИСНОВКИ .....</b>                                     | <b>84</b> |
|               | <b>СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....</b>                    | <b>85</b> |

## ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

|       |  |
|-------|--|
| РДН   | Ринку на добу наперед  |
| БР    | Балансуючий ринок  |
| ТГ    | Теплова генерація  |
| ВДЕ   | Відновлювані джерела енергії   |
| TSO   | Transmission system operators (оператор системи передачі)                            |
| ЧР    | Часовий ряд  |
| АР    | Авторегресія   |
| SSA   | Singular spectrum analysis (сингулярний спектральний аналіз)                         |
| ARIMA | Autoregressive integrated moving average (авторегресійна модель ковзного середнього) |
| ANN   | Artificial neural networks (штучні нейронні мережі)                                  |
| ІТ    | Інформаційна технологія  |
| ПЗ    | Програмне забезпечення   |
| ЕЕ    | Електрична енергія   |

## ВСТУП

Сектор енергетики являє собою стратегічну інфраструктуру національної економіки, від якої залежить розвиток країни в цілому. Безпечне, надійне і ефективне функціонування енергетичного сектора має велике значення для економіки. Крім того, енергетика надає комунальні послуги, що мають важливе соціальне значення.

Лібералізація ринку електроенергії в Євросоюзі призвела до появи оптового ринку електроенергії та потужності, де учасники ринку функціонують в конкурентних умовах та кожен день вирішують проблемами формування ринкової стратегії та планування майбутніх фінансових потоків.

Електроенергія є товаром, що демонструє набір властивостей, який не є характерними для інших ринків: постійний баланс між виробництвом і споживанням, навантаженням і виробленням, на які впливають зовнішні погодні умови, залежність споживання від години, дня тижня і пори року [1]. Через ці фактори динаміка цін на електроенергію демонструє поведінку, невластиву іншим ринкам, наприклад, раптові і несподівані піки цін або сезонність [2].

В результаті чого ринки електроенергії стали центральною точкою досліджень в енергетичному секторі, і точне прогнозування цін на електроенергію стало однією з найбільших проблем, з якими стикаються різні суб'єкти ринку. Звичайна мотивація цих зусиль – чисто економічна: у міру підвищення точності прогнозування учасники ринку отримують економічний прибуток. Крім того, ще один важливий факт, який слід враховувати, полягає в тому, що ринки електроенергії створені для підтримки стабільності мережі. Зокрема, коли ціни стають більш волатильними, баланс енергосистеми порушується, що може призвести до використання стратегічних резервів [3].

**Актуальність теми.** Рішення завдання прогнозування грає важливу роль в процесах як стратегічного планування, так і оперативного управління в сфері електроенергетики.



Актуальність застосування сучасних методів аналізу і прогнозування обумовлена необхідністю формування у суб'єктів ринку, органів виконавчої влади, експертного співтовариства і інших зацікавлених осіб достовірних і обґрунтованих знань про перспективну кон'юнктуру ринку з метою підвищення якості планування своєї роботи на енергетичних ринках.

Для учасників оптового ринку електроенергії і потужності (ОРЕМ) найбільш актуальні коротко- і середньострокові прогнози ціни на електроенергію на ринку на добу вперед (РДН). Точне прогнозування майбутніх цінових показників РДН дозволяє споживачам раціонально закуповувати необхідні обсяги електроенергії, генераціям – оптимізувати склад і пермеаметрів генеруючого обладнання та знизити до мінімуму збитки при виході на балансуючий ринок [4].

Для вирішення завдань прогнозування ціни РДН зазвичай використовують лінійні регресійні моделі, засновані на аналізі часових рядів. Основним недоліком лінійних моделей є те, що вони описують саме лінійну залежність між впливають факторами і результатом, в той час як прогнозування ринку електроенергії призводить до більш складних залежностей. Іншим суттєвим недоліком лінійних моделей є чутливість до шумів в даних.

Для прогнозування волатильного ринку електроенергії з нелінійними закономірностями оптимальним методом є машинне навчання. Але застосування алгоритмів машинного навчання до прогнозування ціни вимагає потужних обчислювальних ресурсів.

**Актуальність теми в Україні.** 13 квітня 2017 року Верховною Радою України прийнято Закон України «Про ринок електричної енергії» [5].

Згідно даного документу з 1 липня 2019 року почав діяти новий ринок електричної енергії, метою якого є запровадження конкурентних механізмів функціонування ринку електричної енергії, вільний вибір контрагентів та забезпечення права споживача вільно обирати постачальника електричної енергії. Законом передбачені різні механізми купівлі-продажу електричної енергії – двосторонні договори, ринок «на добу наперед» та внутрішньодобовий ринок.

Нова модель стимулює конкуренцію, тому вона вигідна для гнучких та прогресивних гравців — виробників, постачальників, споживачів.

Населення в Україні є дотаційною частиною споживачів та істотних змін не відчує. Іншим споживачам доведеться будувати точні прогнози споживання, постачальникам — індивідуально підходити до кожного клієнта, враховуючи сезонне та добове споживання, генеруючим компаніям — конкурувати між собою у виробництві [6].

Прийняття закону є необхідною передумовою структурних змін в електроенергетиці України, підґрунтям для модернізації галузі та інтеграції ринку електричної енергії України до регіональних енергетичних ринків, з наступним входженням до загальноєвропейського енергетичного ринку.

**Метою дипломної роботи** є розв’язання актуальної задачі зі створення інформаційної технології прогнозування цін ринку електричної енергії країн Євросоюзу на прикладі угорського ринку на добу наперед.

Для досягнення даної мети були визначені наступні **задачі**:

1. Дослідити існуючі методи та засоби короткострокового прогнозування часових рядів.
2. Розробити інформаційну технологію прогнозування споживання та ціни на електроенергію на добу наперед для угорського ринку.
3. На основі запропонованої технології розробити програмне забезпечення (ПЗ).
4. Провести експериментальні дослідження запропонованих методів на реальних даних.
5. Висновки та результати оформити у вигляді пояснювальної записки згідно з ДСТУ 3008-2015.

**Об'єктом дослідження** є процес підготовки та обробки даних ринку електричної енергії.

**Предметом дослідження** є методи і засоби короткострокового прогнозування часових рядів.

**Методи дослідження.** При вирішенні поставлених задач використані такі методи, як кореляційний та регресійний аналіз; для знаходження коефіцієнтів моделі був використаний метод найменший квадратів; для розкладу ЧР на складові було застосовано сингулярний розклад.

**Оцінка сучасного стану проблеми на основі вітчизняної та зарубіжної літератури.** Моделі з прогнозування цін зазвичай поділяється на п'ять областей: (1) теоретичні моделі ігор, що імітують роботу ринку, (2) фундаментальні методи, що використовують фізичні та економічні фактори, (3) моделі з використанням статистичних властивостей торгівлі електроенергією для оцінки ризику та похідних інструментів, (4) статистичні моделі, що містять часові ряди та економетричні моделі, та (5) методи штучного інтелекту [7]. Для прогнозування цін на день вперед статистичні методи та штучний інтелект показали найкращі результати, оскільки є можливість дослідити нелінійні залежності між багатьма змінними [7].

Типовими статистичними методами є: моделі AR та ARX [8], моделі ARIMA [9], [10], динамічна регресія [11], подвійна сезонна модель Гольц-Зима [12], модель TARX [13], непараметричні моделі [8] або моделі на основі GARCH [14]. Крім того, в межах одного класу методів також застосовуються різні гібридні моделі, наприклад моделі на основі вейвлетів [10], [15].

У розглянутій літературі [16-30] численні дослідження спрямовані на аналіз математичних моделей і методів середньострокового і довгострокового прогнозування ЧР в різних областях науки і техніки. В [31] представлені адитивні, мультиплікативні і змішані моделі прогнозування електроспоживання.

В [32] розглянуті різні комбіновані математичні моделі для прогнозування трендової, сезонної, тижневої, залишкової складових процесів споживання електроенергії та запропоновано декомпозиційний метод моделювання (ДММ). Серед пропонованих моделей прогнозування окремих складових процесу використовуються експоненціальне згладжування, модель Хольта-Вінтерса,

подвійного експоненціального згладжування, фільтр Калмана, ряди Фур'є, моделі спектрального аналізу, експоненціально зваженого ковзного середнього та інші.

ЧР споживання цільових продуктів в системах *енергетики* є істотно нелінійними із середньо- і довгостроковими залежностями, тому апроксимація таких складних ЧР тільки моделлю ARIMA не завжди є задовільною. Доведено, що для підвищення точності прогнозування нестационарних ВР необхідно ефективно вилучити трендову, гармонійні і шумову складові і далі підставляти їх в моделі прогнозування. Таким чином, методи декомпозиції ЧР є сприятливими для первинної обробки даних.

Сингулярний спектральний аналіз (SSA) є одним з найбільш ефективних методів для аналізу ЧР зі складними періодичними компонентами. Зародження SSA зазвичай асоціюється з публікацією роботи [33]. SSA став широко використовуватися в аналізі періодичних і трендових складових ЧР [34-37].

В [38] описані процедури прогнозування ціни на електроенергію на датському ринку. Прогнозування проводиться на 24 години наперед. В даній статті використовуються авторегресійна модель ковзного середнього (ARIMA), тригонометрична сезонна трансформація Бох-Сох із залишками ARMA в тренді та сезонних компонентах (TBATS) та штучні нейронні мережі (ANN). ARIMA та ANN використовуються із зовнішніми факторами. Загалом у даній роботі продемонстрована ґрунтовна методологія прогнозування, відображаються фактичні результати прогнозування та наведено вдосконалення різних методів прогнозування.

Врахування зростаючої інтеграції між ринками електроенергії в прогнозуванні цін та покращення прогнозованих показників описано у статті [39]. Приведено опис нейронної мережі, яка розглядає особливості пов'язаних ринків для підвищення точності прогнозування на місцевому ринку. Для вимірювання важливості цих особливостей запропоновано новий алгоритм вибору функцій, який, використовуючи байєсівську оптимізацію та функціональний аналіз дисперсії, оцінює вплив функцій на продуктивність алгоритму. В статті

пропонується модель, яка одночасно екстраполює ціни з двох ринків, що покращує точність. Також в даній роботі запропоновано алгоритм вибору зовнішніх факторів, який відкидає показники, що завдають незначний вплив. Дослідження проводили на базі бельгійського ринку електроенергії.

**Практичне значення одержаних результатів.** Результати дипломної роботи можуть бути використані:

1. При розробці автоматизованої системи прогнозування ціни на РДН, яка може бути використана трейдерами з метою отримання прибутку.
2. Для прийняття рішення операторами системи передач для прогнозування споживання на добу наперед з метою покращення роботи генерації та енергосистеми в цілому.

**Структура дипломної роботи.** Дипломна робота складається зі вступу, 2-х розділів, висновків, списку використаних джерел, що налічує 60 найменування. Повний обсяг роботи складає 90 ст. друкованого тексту.

## ПОСТАНОВКА ЗАДАЧІ

Нехай маємо:

1) цільові показники  $Y$ :

- споживання електроенергії  $Y_1(t)$  (МВт)
- ціна на електричну енергію  $Y_2(t)$  (грн/МВт)

2) множину факторів  $X$ , серед яких:

- година доби  $H(t)$  (1-24)
- тип дня  $W(t)$ : 1-7 (Понеділок-Неділя), 8- святкові дні, 10- відспрацювання, 11-канікули
- температура повітря  $T(t)$  (°C)
- сонячності  $S(t)$
- швидкість вітру  $Wnd(t)$  (м/с)
- перетоки електроенергії між сусідніми державами  $Exch(t)$  (МВт)

Всі дані задані на однаковому проміжку часу  $[1: T]$  в дискретні моменти часу  $t = 1, 2, \dots, T$ , де крок – одна година.

Зауважимо, що значення зовнішніх факторів в моменти часу  $(T + 1), \dots, (T + P)$  є доступними. Фактори з множини  $X$  можуть входити в модель повністю або частково.

Поставлені задачі:

1) Визначити взаємодію цільових показників та факторів, що формують ситуацію на ринку.

2) Формалізувати та описати моделі, що реалізують короткостроковий (на 24 години наперед ( $P = 24$ )) погодинний прогноз цільових показників.

3) Оцінити параметри розроблених моделей, що являють собою залежність цільових показників від множини факторів.

4) Розробити інформаційну технологію короткострокового прогнозування цін та споживання на електроенергію на РДН.

- 5) На основі запропонованої технології розробити ПЗ, що реалізує вирішення задачі прогнозування цільових показників.
- 6) Провести тестування ПЗ на реальних даних на прикладі угорського ринку електричної енергії
- 7) Оцінити похибку прогнозування. Проаналізувати вплив окремих факторів на похибку.
- 8) Порівняти результати прогнозування розробленого ПЗ з існуючими системами.

## РОЗДІЛ 1

### ОСНОВНІ ТЕОРЕТИЧНІ ПОЛОЖЕННЯ

Ринки електроенергії в Європі почали лібералізуватися на початку 1990х років. З того часу виробництво, постачання, передача та розподілення електроенергії стало конкурентною діяльністю. В даних умовах задача прогнозування ціни на електроенергію стає невід'ємним і повсякденним завданням для більшості учасників ринку. Особливо актуальними є прогнозні моделі, що враховують вплив зовнішніх факторів.

В даному розділі наведено огляд методів прогнозування часових рядів. У підрозділі 1.1 наведені характеристики електроенергії як товару, описані принципи роботи ринку електричної енергії та відображена структура споживання електроенергії в Європі. У підрозділі 1.2 проведено огляд існуючих моделей короткострокового прогнозування цін в енергетиці. У підрозділі 1.3 надана формальна постановка задачі прогнозування часових рядів та проведено огляд математичних моделей серед яких, зокрема, регресійні та авторегресійні моделі, моделі на основі експоненційного згладжування та класифікаційно-регресійних дерев, нейронні мережі, ланцюги Маркова. У підрозділі 1.4 наведені оцінки точності прогнозування, які використані в порівняльному аналізі якості роботи ІТ.

#### **1.1. Основні характеристики області енергетики країн Євросоюзу**

Електроенергія як ринковий товар принципово відрізняється від інших видів товарів двома особливостями. По-перше, в момент споживання електроенергії ми не знаємо точні витрати на її виробництво і не можемо регулювати власне споживання, виходячи з величини цих витрат. По-друге, логістика електроенергії підпорядковується законам електротехніки.

Основними учасниками торгівлі електроенергією є генеруючі компанії, які володіють електростанціями і виробляють електроенергію, назвемо їх генераторами, і компанії - споживачі електроенергії, які її купують [40].



*Електроенергія* – це енергетичний товар, що відрізняється від нафти, природного газу і вугілля, оскільки його не можна легко зберігати на місцях у великій кількості. Зберігання електроенергії в масштабі електромережі є бажаною характеристикою електромережі, однак вона не використовується широко, оскільки не є економічно здійсненою та абсолютно неконкурентоспроможною. Крім того, електроенергія в електромережі повинна бути ідеально збалансована в усі часи, щоб запобігти відключенням та іншим проблемам.

Ринки електроенергії в Європі почали лібералізуватися на початку 1990х років. З лібералізацією ринків виробництво, постачання, передача та розподілення електроенергії стало конкурентоспроможною діяльністю [41].

Важливою частиною лібералізації енергетичного ринку в *Угорщині* стало у 2010 році MAVIR створення угорським TSO дочірньої компанії, угорську біржову компанію Limited Limited. HUPX Ltd. [42], що є оператором організованого угорського ринку електроенергії, який займає лідируючі позиції в Центральній та Східній Європі. За допомогою свого регулювання та прийнятих торгових рамок він сприяє ліквідності угорського енергетичного ринку.

Основна діяльність HUPX – надання опорних цін і біржової торгівлі – що ефективно сприяє розвитку угорського ринку електроенергії .

За рівнем економічного розвитку Угорська Республіка відстає від передових країн, зокрема, ВВП на душу населення становить ~ 60-70% від середнього рівня держав, що входять в Організацію економічного співробітництва. Споживання електрики на душу населення становить ~ 4,2 тис. КВт\*год на рік (2015 рік) – що є середнім показником для промислово розвинених держав [43].

Угорщина забезпечує себе енергоносіями лише на 45-50% (не рахуючи імпорту 100% ядерного палива) і отримує з-за кордону до 30% і більше електрики. У нинішньому десятилітті залежність від імпорту електроенергії зросла на тлі зниження вироблення електростанцій на 1/4. Незважаючи на наявність достатніх енергетичних потужностей (рис. 1.1 ) більшість енергоджерел в країні зупинені.

Пояснюється це тим, що собівартість вироблення ЕЕ на даних станціях перевищує номінальну вартість, за яку ЕЕ купується на загальноєвропейському ринку.

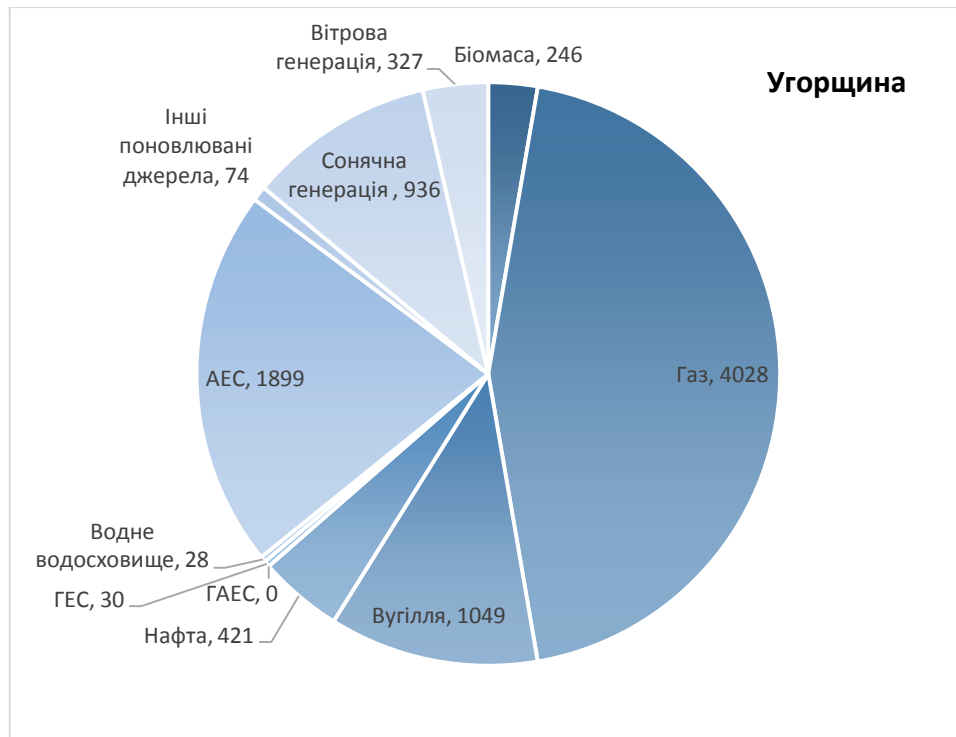


Рис. 1.1. Структура виробництва електроенергії за типом генерації в Угорщині

Політика енергоефективності *Чехії* включає довгострокову мету щодо зниження енергоємності та рівня споживання сирих матеріалів і збільшення частки ВДЕ в загальному енергоспоживанні країни, а також приведення цих показників у відповідність з показниками високорозвинених країн.

Згідно з даними Європейської комісії, в 2018 році в Чехії було вироблено 31,17 млн. т.н.е. енергії. При цьому близько 70% виробленої в Чехії енергії припадає на тверде викопне паливо (23% на кам'яне вугілля), 22,6% - на атомну енергетику, 0,7% - на нафту і СПГ, 0,5% - на природний газ (рис 1.2). У 2018 році частка ВДЕ в загальному виробництві енергії в Чехії становила близько 8%.

Чехія імпортує більшу частину таких енергоресурсів, як нафта і газ. Велика частина імпорту припадає на Росію. За даними Міністерства промисловості і торгівлі Чехії, в 2018 році частка біоенергетики в загальному споживанні первинних джерел енергії становила близько 4,5% [44].

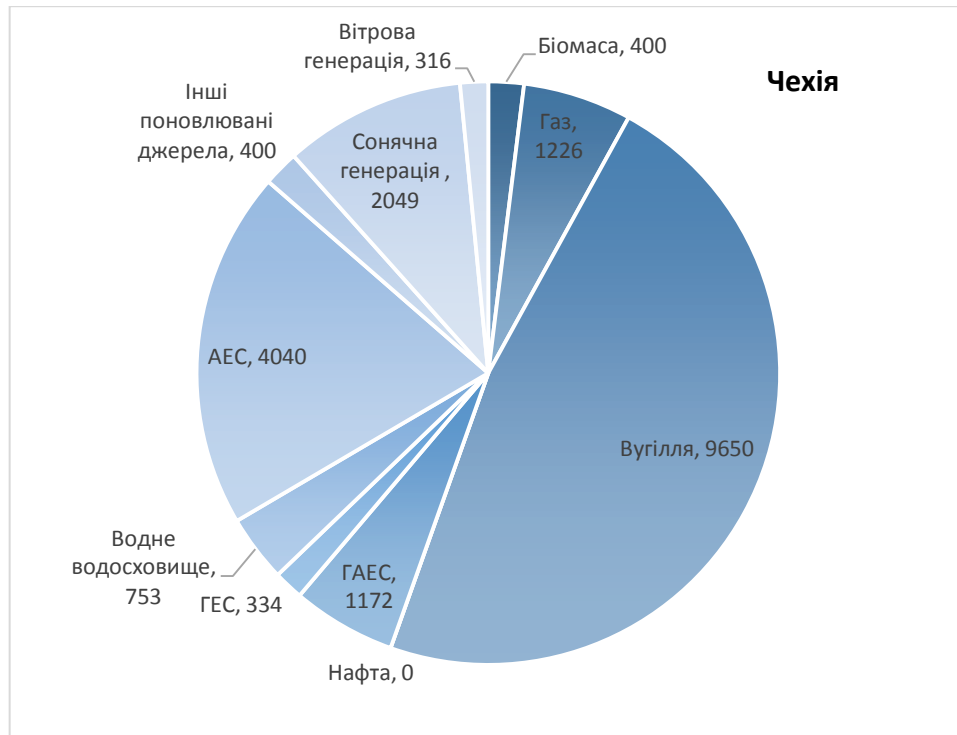


Рис. 1.2. Структура виробництва електроенергії за типом генерації в Чехії

Влада **Румунії** впроваджує систему підтримки відновлюваних джерел енергії, засновану на зелених сертифікатах. Тим не менш на даний момент найбільшу долю ЕЕ виробляють ГЕС (36 %) та ТЕС (28 %) (рис 1.3).

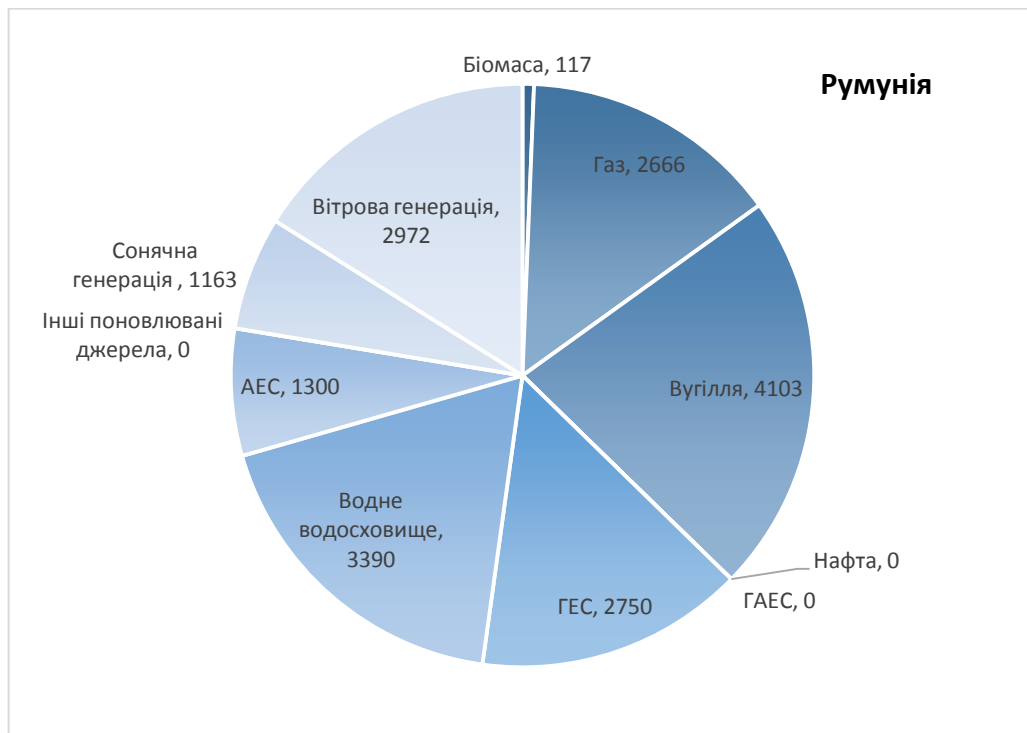


Рис. 1.3. Структура виробництва електроенергії за типом генерації в Румунії

Найбільша економіка Європи – **Німеччина** – визнаний лідер в сучасній енергетиці. Прийнявши державну програму, названу «Енергоповорот (Energiewende)», з вкрай амбітними цілями по декарбонізації своєї економіки, Німеччина вельми послідовно її виконує. Вироблення нових ВДЕ перевищила третину в загальній генерації країни (рис 1.4).

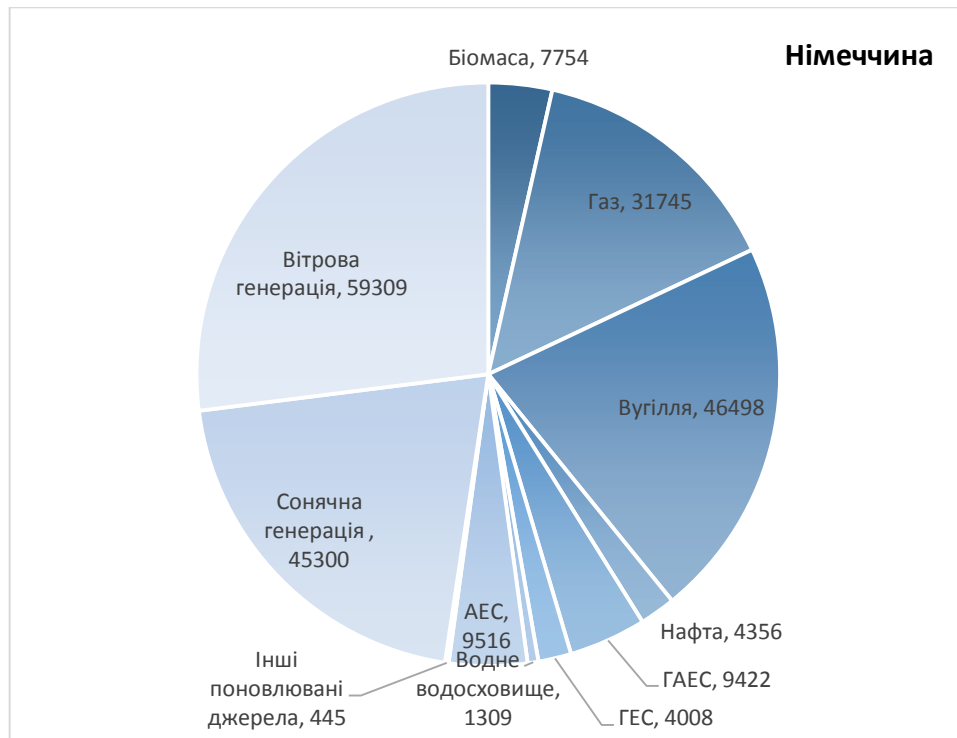


Рис. 1.4. Структура виробництва електроенергії за типом генерації в Німеччині

Основними **цілями енергетичної політики ЄС** є безпека поставок, конкурентоспроможність та сталість. Під безпекою поставок розуміються надійні поставки енергетичних ресурсів, під конкурентоспроможністю — наявність цих ресурсів за конкурентоздатною ціною, а під сталістю — мінімально можливий негативний вплив енергетичного сектору (як виробництва, так і споживання) на навколишнє середовище. Досягнення відповідного балансу між цими трьома цілями має вирішальне значення для економічної конкурентоспроможності ЄС, його політичної самостійності та активізації боротьби зі зміною клімату.

Питання енергетичної безпеки вперше з'явилося в порядку денному Європейського Союзу після того як країни Центрально-Східної Європи вступили до ЄС у 2004 р. Особливе значення енергетичної безпеки та, зокрема,

диверсифікації поставок енергоресурсів було чітко продемонстровано під час газових криз 2006 та 2009 рр., коли транзит територією України було раптово припинено, унаслідок чого велика частина Центрально-Східної Європи залишилася без достатньої кількості енергоресурсів у зимовий період, з відповідними гуманітарними та економічними наслідками.

За окремих вірогідних ситуацій енергоресурси можуть бути використані як геополітичні інструменти, спрямовані на досягнення стратегічних і політичних цілей. З урахуванням можливих ризиків для Європи важливо знайти спосіб пом'якшити наслідки поточного вразливого становища, в першу чергу шляхом диверсифікації джерел та маршрутів поставок, збільшення обсягів внутрішнього виробництва енергії, оптимізації структури споживання енергії, забезпечивши при цьому конкурентоспроможний рівень цін.

За даними Євростату, у 2016 р. в ЄС рівень залежності від імпорту становив 54% споживаної енергії за близько 400 млрд євро на рік (близько 1 млрд євро в день), що означає, що більш ніж половина енергетичних потреб ЄС була забезпечена чистим імпортом. ЄС імпортує майже 90% сирої нафти, 66% природного газу, 42% вугілля та інших видів твердого палива, а також 40% урану й інших видів ядерного палива.

Найвищий ступінь залежності від імпорту спостерігається у двох стратегічно важливих секторах — нафтовому та газовому. Ця залежність по країнам коливається від понад 90% на Мальті, Кіпрі та у Люксембурзі та до нижче 20% в Естонії та Данії. Залежність від імпорту енергії зросла з 2000 р., коли вона становила лише 47% [45].

### **Об'єднана енергетична система України на шляху до євроінтеграції.**

Енергетика є базовою галуззю національної економіки, ефективне функціонування якої є необхідною умовою сталого економічного розвитку, забезпечення енергетичної безпеки та енергетичної незалежності.

Об'єднана енергетична система України (ОЕС України) за встановленою потужністю електростанцій є однією з найбільших енергосистем Європи.

Загальна встановлена потужність ОЕС України станом на 01.01.2018 р. становила 51,8 ГВт (рис 1.5), з яких потужність ТЕС, ТЕЦ та блок-станцій становила 59,0%; АЕС – 26,7%; ГЕС – 9,1%; ГАЕС – 2,9%; СЕС – 1,5%; ВЕС – 0,6%; та БіоЕС – 0,2%.

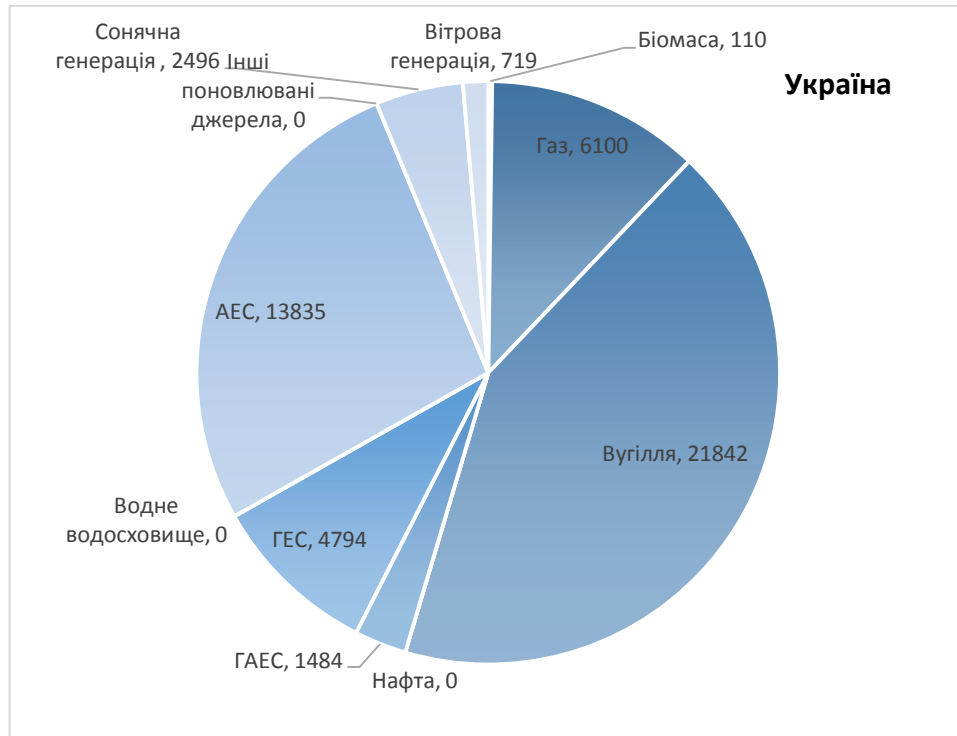


Рис. 1.5. Структура виробництва електроенергії за типом генерації в Україні

Основна частина ОЕС України працює у паралельному режимі з енергетичними об'єднаннями Республіки Білорусь, Республіки Молдови, Російської Федерації (ОЕС Центра, ОЕС Півдня).

Південно-західна частина енергосистеми Західного регіону – «Острів Бурштинської ТЕС», працює у складі об'єднання енергетичних систем країн Континентальної Європи – ENTSO-E з 01.07.2003 р.

Величина потужності для забезпечення максимально допустимого експорту електроенергії міждержавними лініями електропередачі у наступних напрямках становить з «Острова Бурштинської ТЕС» до Угорщини, Словаччини, Румунії – до 650 МВт. Максимальна величина перетоку між ОЕС України та ЕС Молдови – до 700 МВт з можливістю обмеження до нуля.

Кінцеве споживання енергії в ЄС склало дві третини (66,6%) валового внутрішнього споживання. Енергоємність є мірою енергоефективності економіки. Найменш енергоємними економіками в ЄС в 2018 році були Ірландія, Данія, Люксембург, Мальта і Великобританія; вони використовували найменшу кількість енергії щодо їх загального економічного розміру.

В Європі кількість спожитої електричної енергії можна поділити на чотири сектори:

- Житлове (опалення, освітлення та прилади);
- Комерційне (освітлення, опалення та охолодження комерційних будівель та надання водопровідних та каналізаційних послуг);
- Промислове (сільське господарство, видобуток, виробництво та будівництво);
- Транспортне (пасажирські, вантажні та трубопровідні).

Дані відношення наведені на рис 1.6.

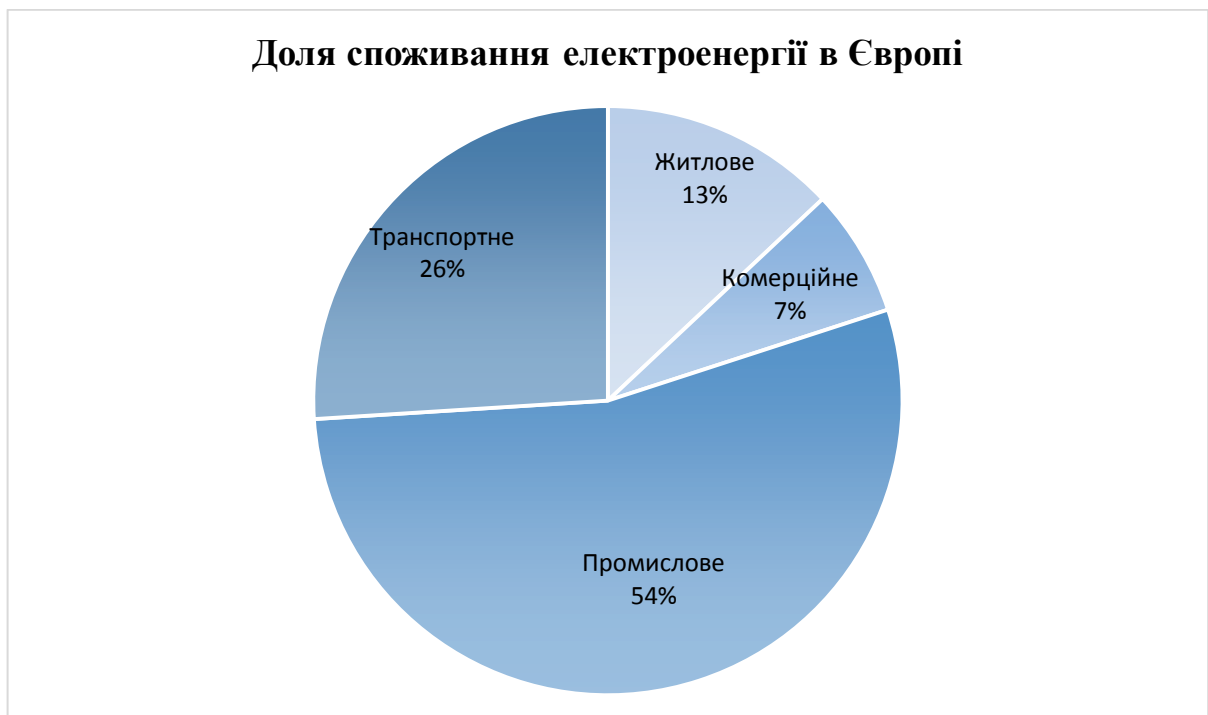


Рис. 1.6. Структура споживання електроенергії в країнах Європи

## **1.2. Аналіз існуючих моделей короткострокового прогнозування цін в енергетиці**

Напрямок прогнозування цін на електроенергію відносно молодий, і свою актуальність набув з моменту масштабної лібералізації галузі та розвитку конкурентних відносин.

Керуючись потребами ринку, дослідниками було запропоновано низку підходів щодо створення прогнозних моделей, які, в свою чергу, можуть бути розділені на кілька груп.

Перша група прогнозних моделей – моделі на основі теорії ігор. Для моделювання ринкової кон'юнктури в енергетиці використовуються моделі рівноваги Неша, модель Курно, модель Бертрана та інші. Більш детально з моделями прогнозу на основі теорії ігор можна ознайомитися в роботі [46].

До другої групи моделей відносять, так звані, імітаційні або фундаментальні моделі. Відмінною особливістю моделей даної групи є моделювання цінкових стратегій учасників ринку. Прогнозне значення ціни на електроенергію визначається за результатами аукціону для даної зони енергосистеми.

Даний принцип ціноутворення передбачає різний маржинальний дохід електричних станцій оптимізаційної задачі, що враховує цінкові заявки споживачів (покупців) і виробників (продавців) електроенергії, а також економічні та технологічні обмеження енергосистеми. З урахуванням того, що фундаментальні моделі вимагають великого обсягу вихідних даних, їх використання для короткострокового прогнозування обмежена. Детальний розгляд фундаментальних моделей прогнозування ринкової кон'юнктури представлено в роботах [47,48].

Третя група прогнозних моделей ґрунтується на аналізі часових рядів, при якому використовується сукупність статистичних методів, призначених для виявлення структури часових рядів, вивченні історичної динаміки досліджуваних показників і для прогнозування (екстраполяції). У даній групі прогнозних моделей виділяють два основних напрямки: традиційні статистичні моделі та моделі штучного інтелекту.



До традиційних статистичних моделей відносяться досить поширені і широко використовувані параметричні моделі регресії, авторегресії (AR), ковзного середнього (MA), моделі авторегресії – ковзного середнього (ARMA), інтегровані моделі авторегресії - ковзного середнього (ARIMA), розширеної версії останньої моделі (ARIMAX), авторегресійної умовної гетероскедастичності (GARCH), а також деякі інші модифікації зазначених моделей. Даний клас моделей відрізняється своєю простотою: параметричні моделі, за винятком моделі ARIMAX, моделюють динаміку досліджуваного показника тільки на основі його історичних значень і не враховують впливу зовнішніх чинників. Виходячи з даної характеристики параметричних моделей, їх використання (за винятком моделі ARIMAX) для оцінки перспективних значень цін на електричну енергію може істотно знизити якість прогнозу, з огляду на істотну значущість зовнішніх факторів на рівень цін на електроенергію.

Досягнення адекватності моделей ARIMAX, в свою чергу, накладають обмеження на кількість зовнішніх факторів, що також може бути «бар'єром» для використання даних моделей для прогнозування цін в електроенергетиці. Застосування параметричних моделей для прогнозування ринкової кон'юнктури ринку електроенергії детально розглянуто в роботах [49].

Представниками моделей штучного інтелекту є прогнозні моделі на основі нейронних мереж. На відміну від традиційних статистичних моделей, які, по суті, є лінійними, моделі штучного інтелекту дозволяють врахувати складний нелінійний взаємозв'язок між залежною і незалежними змінними. Останнє пов'язано з архітектурою таких моделей, які повторюють пристрій головного мозку людини. Використання даних моделей дозволяє асоціювати досліджуваний показник (ціна на електроенергію) з набором різних факторів (не тільки історичних значень) і екстраполювати.

З початку 1990-х років процес дерегулювання і введення конкурентних ринків електроенергії змінив ландшафт традиційно монополістичних і контрольованих урядом секторів енергетики. У Європі, Північній Америці та

Австралії електроенергія в даний час продається за ринковими правилами з використанням спотових і похідних контрактів. Проте, електроенергія є дуже особливим товаром: вона економічно не вигідна, а стабільність енергосистеми вимагає постійного балансу між виробництвом і споживанням. У той же час попит на електроенергію залежить від погоди (температура, швидкість вітру, опади і т. Д.) І інтенсивності ділових і повсякденних дій (в години пік і в непіковий час, дні тижня, вихідні, свята і т. Д.) . Ці унікальні характеристики призводять до динаміки цін, що не спостерігається ні на жодному іншому ринку, демонструючи щоденну, щотижневу і часто щорічну сезонність, а також різкі, короткочасні і в цілому непередбачені стрибки цін.

За останні 15 років були випробувані різні методи прогнозування і ідеї з різним ступенем успіху. Їх можна умовно розділити на групи.

Багатоагентні моделюють роботу системи різнорідних агентів (одиниць, компаній), що взаємодіють один з одним, і формують цінової процес шляхом зіставлення попиту і пропозиції на ринку. Цей клас включає моделі, засновані на витратах, рівноважні або теоретико-ігрові підходи

Багатоагентні моделі зазвичай зосереджені на якісних питаннях, а не на кількісних результатах. Вони можуть дати уявлення про те, чи будуть ціни перевищувати граничні витрати і як це може вплинути на результати гравців. Однак вони створюють проблеми, якщо необхідно зробити більш кількісні висновки, особливо якщо ціни на електроенергію необхідно прогнозувати з високим ступенем точності.

Фундаментальні (структурні) методи намагаються охопити основні фізичні і економічні відносини, які присутні у виробництві і торгівлі електроенергією. Функціональні зв'язки між основними факторами (навантаженнями, погодними умовами, параметрами системи) досліджуються, а основні вхідні дані моделюються і прогнозуються незалежно, часто за допомогою статистичних методів.

При практичній реалізації фундаментальних моделей виникають дві основні проблеми: доступність даних і включення стохастичних коливань фундаментальних факторів.

Моделі в зменшеній формі (кількісні, стохастичні) характеризують статистичні властивості цін на електроенергію в часі з кінцевою метою оцінки похідних інструментів і управління ризиками. Їх основна мета полягає не в тому, щоб надавати точні погодинні прогнози цін, а в тому, щоб відтворити основні характеристики щоденних цін на електроенергію, такі як граничні розподіли в майбутні моменти часу, динаміка цін і кореляції між цінами на сировину.

Статистичні (економетричні, технічні аналізи) методи прогнозують поточну ціну, використовуючи математичну комбінацію попередніх цін і поточних значень зовнішніх факторів, зазвичай показників споживання та виробництва, або погодних змін. Дві найбільш важливі категорії – адитивна і мультиплікативна моделі. Перші більш популярні, але обидва тісно пов'язані – мультиплікативна модель для цін може бути перетворена в адитивну модель для логарифмічних цін. Статистичні моделі привабливі тим, що до їх компонентів може бути додана деяка фізична інтерпретація, що дозволяє інженерам і системним операторам зрозуміти їх поведінку. Їх часто критикують за їх обмежену здатність моделювати (зазвичай) нелінійне поведінку цін на електроенергію і пов'язані з цим фундаментальні змінні. Однак в практичних додатках характеристики статистичних моделей не гірше, ніж у нелінійних методів штучного інтелекту.

Ключовим моментом в моделюванні і прогнозуванні цін на електроенергію є врахування сезонності. Ціна на електроенергію демонструє сезонність на трьох рівнях: щоденний і щотижневий, а в деякій мірі – річний. При короткостроковому прогнозуванні річна або довгострокова сезонність зазвичай ігнорується, але щоденні і щотижневі моделі (включаючи окрему обробку свят) мають першорядне значення.

### 1.3. Огляд математичних моделей прогнозування часових рядів

#### Формальна постановка задачі прогнозування часових рядів

*Прогнозування без урахування зовнішніх чинників.* Нехай значення часового ряду доступні в дискретні моменти часу  $t = 1, 2, \dots, T$ . Позначимо часовий ряд

$$Z(t) = Z(1), Z(2), \dots, Z(T)$$

У момент часу  $T$  необхідно визначити значення процесу  $Z(t)$  в моменти часу  $T + 1, \dots, T + P$ . Момент часу  $T$  називається моментом прогнозу, а величина  $P$  - часом випередження.

Для обчислення значень часового ряду в майбутні моменти часу потрібно визначити функціональну залежність, яка відобразить зв'язок між минулими і майбутніми значеннями цього ряду

$$Z(t) = F(Z(t - 1), Z(t - 2), Z(t - 3), \dots) + \varepsilon_t \quad (1.1)$$

Залежність (1.1) називається моделлю прогнозування, де  $\varepsilon_t$  - випадкова складова.

Потрібно створити таку модель прогнозування, для якої середнє абсолютне відхилення реального значення від прогнозованого прямує до мінімального для заданого  $P$  (1.2)

$$E = \frac{1}{P} \sum_{t=T+1}^{T+P} |\varepsilon_t| \rightarrow \min \quad (1.2)$$

Прогнозні (розрахункові)  $\hat{Z}(t)$  значення часового ряду  $Z(t)$  можна записати у вигляді

$$\hat{Z}(t) = F(Z(t - 1), Z(t - 2), Z(t - 3), \dots).$$

*Прогнозування з урахуванням зовнішніх факторів.* Нехай значення вихідного часового ряду  $Z(t)$  доступні в дискретні моменти часу  $t = 1, 2, \dots, T$ . Передбачається, що на значення  $Z(t)$  впливає набір зовнішніх факторів. Нехай перший зовнішній фактор  $X_1(t_1)$  доступний в дискретні моменти часу  $t_1 = 1, 2, \dots, T_1$ , другий зовнішній фактор  $X_2(t_2)$  доступний в моменти часу  $t_2 =$

$1, 2, \dots, T_2$  і т. д. У разі, якщо дискретність вихідного часового ряду і зовнішніх чинників, а також значення  $T, T_1, \dots, T_S$  різні, то часові ряди зовнішніх факторів  $X_1(t_1), \dots, X_S(t_S)$  необхідно привести до єдиної шкалою часу  $t$ .

У момент прогнозу  $T$  необхідно визначити майбутні значення вихідного процесу  $Z(t)$  в моменти часу  $T + 1, \dots, T + P$ , враховуючи вплив зовнішніх факторів  $X_1(t), \dots, X_S(t)$ . При цьому вважаємо, що значення зовнішніх факторів в моменти часу  $X_1(T + 1), \dots, X_1(T + P), \dots, X_S(T + 1), \dots, X_S(T + P)$  є доступними.

Для обчислення майбутніх значень процесу  $Z(t)$  в зазначені моменти часу потрібно визначити функціональну залежність, яка відобразить зв'язок між минулими значеннями  $Z(t)$  і майбутніми, а також приймаючи до уваги вплив зовнішніх факторів  $X_1(t), \dots, X_S(t)$  на вихідний часовий ряд

$$Z(t) = F(Z(t - 1), Z(t - 2), \dots, X_1(t - 1), \dots, X_S(t), X_S(t - 1), \dots) + \varepsilon_t \quad (1.3)$$

Залежність (1.3) називається *моделлю прогнозування з урахуванням зовнішніх факторів*  $X_1(t), \dots, X_S(t)$ .

### 1.3.1. Регресійні моделі

Існує багато задач, які потребують вивчення залежності між двома і більше змінними. Для вирішення таких завдань використовується регресійний аналіз.

Метою регресійного аналізу є визначення залежності між вихідної змінної і зовнішніми факторами (регресорами). При цьому коефіцієнти регресії можуть визначатися методом найменших квадратів або методом максимальної правдоподібності [50].

**Лінійна регресійна модель.** Найпростішим варіантом регресійній моделі є лінійна регресія. В основу моделі покладено припущення, що існує дискретний зовнішній фактор  $X(t)$ , який впливає на досліджуваний процес  $Z(t)$ , при цьому зв'язок між процесом і зовнішнім фактором лінійна. Модель прогнозування на підставі лінійної регресії описується рівнянням (1.4):

$$Z(t) = \alpha_0 + \alpha_1 X(t) + \varepsilon_t \quad (1.4),$$

де  $\alpha_0$  і  $\alpha_1$  – коефіцієнти регресії;  $\varepsilon_t$  – помилка моделі. Для отримання прогнозних значень  $Z(t)$  в момент часу  $t$  необхідно мати значення  $X(t)$  в той же момент часу  $t$ .

**Множинна регресійна модель.** На практиці на процес  $Z(t)$  впливають цілий ряд дискретних зовнішніх чинників  $X_1(t), \dots, X_S(t)$ . Тоді модель прогнозування має вид (1.5):

$$Z(t) = \alpha_0 + c + \alpha_2 X_2(t) + \alpha_S X_S(t) + \varepsilon_t \quad (1.5)$$

Недоліком даної моделі є те, що для обчислення майбутнього значення процесу  $Z(t)$  необхідно знати майбутні значення всіх факторів  $X_1(t), \dots, X_S(t)$ , що є важким завданням на практиці.

В основу **нелінійної регресійної моделі** покладено припущення про те, що існує відома функція, що описує залежність між вихідним процесом  $Z(t)$  і зовнішнім фактором  $X(t)$

$$Z(t) = F(X(t), A).$$

В рамках побудови моделі прогнозування необхідно визначити параметри функції  $A$ . Наприклад, можна припустити, що

$$Z(t) = \alpha_1 \cos(X(t)) + \alpha_0$$

Для побудови моделі досить визначити параметри  $A = [\alpha_1, \alpha_0]$ . Однак на практиці рідко зустрічаються процеси, для яких вид функціональної залежності між процесом  $Z(t)$  і зовнішнім фактором  $X(t)$  заздалегідь відомий. У зв'язку з цим нелінійні регресійні моделі застосовуються рідко.

### 1.3.2. Авторегресійні моделі

В основу авторегресійних моделей закладено припущення про те, що значення процесу  $Z(t)$  лінійно залежить від деякої кількості попередніх значень того ж процесу  $Z(t - 1), \dots, Z(t - p)$ .

**Авторегресійна модель ковзного середнього.** В області аналізу часових рядів модель авторегресії (autoregressive, AR) і модель змінного середнього (moving average, MA) є однією з найбільш використовуваних.

У цій моделі поточне значення процесу виражається як кінцева лінійна сукупність попередніх значень процесу і імпульсу, який називається «білим шумом»,

$$Z(t) = C + \varphi_1 Z(t-1) + \varphi_2 Z(t-2) + \dots + \varphi_p Z(t-p) + \varepsilon_t \quad (1.6)$$

Формула (1.6) описує процес авторегресії порядку  $p$ , який в літературі часто позначається  $AR(p)$ , тут  $C$  – константа,  $\varphi_1, \dots, \varphi_p$  - коефіцієнти,  $\varepsilon_t$  - помилка моделі. Для визначення  $\varphi_i$  і  $C$  використовують метод найменших квадратів або метод максимальної правдоподібності.

Інший тип моделі має велике значення в описі часових рядів і часто використовується спільно з авторегресії, називається моделлю ковзного середнього порядку  $q$  і описується рівнянням

$$Z(t) = \frac{1}{q} (Z(t-1) + Z(t-2) + \dots + Z(t-q)) + \varepsilon_t \quad (1.7)$$

У літературі процес (1.7) часто позначається  $MA(q)$ ; тут  $q$  - порядок ковзного середнього,  $\varepsilon_t$  – помилка прогнозування. Модель ковзного середнього є по суті фільтром низьких частот. Потрібно відзначити, що існують прості, зважені, кумулятивні, експоненціальні моделі змінного середнього.

Для досягнення більшої гнучкості в підгонці моделі часто доцільно об'єднати в одну моделі авторегресії і ковзного середнього. Загальна модель позначається  $ARMA(p, q)$  і поєднує в собі фільтр у вигляді ковзного середнього порядку  $q$  і авторегресії відфільтрованих значень процесу порядку  $p$ .

Якщо в якості вхідних даних використовуються не самі значення часового ряду, а їх різницю  $d$  – того порядку (на практиці  $d$  необхідно визначати, проте в більшості випадків  $d \leq 2$ ), то модель носить назву авторегресії проінтегрованого ковзного середнього. В літературі дану модель називають  $ARIMA(p, d, q)$  (autoregression integrated moving average).

Розвитком моделі ARIMA ( $p, d, q$ ) є модель ARIMAX ( $p, d, q$ ), яка описується рівнянням

$$Z(t) = AR(p) + \alpha_1 X_1(t) + \dots + \alpha_s X_s(t)$$

Тут  $\alpha_1, \dots, \alpha_s$  – коефіцієнти зовнішніх чинників  $X_1(t), \dots, X_s(t)$ . У даній моделі найчастіше процес  $Z(t)$  є результатом моделі MA( $q$ ), тобто відфільтрованими значеннями вихідного процесу. Далі для прогнозування  $Z(t)$  використовується модель авторегресії, в якій введені додаткові регресори зовнішніх чинників  $X_1(t), \dots, X_s(t)$ .

*Авторегресійна модель з умовною гетероскедастністю* (autoregressive conditional heteroskedasticity, GARCH) була розроблена в 1986 році Тімом Петером Борреслевом і є моделлю залишків для моделі  $AR(p)$ . На першому етапі для вихідного часового ряду визначається модель  $AR(p)$  (1.3). Далі передбачається, що помилка моделі (1.3)  $\varepsilon_t$  має дві складові

$$\varepsilon_t = \sigma_t * \zeta_t$$

де  $\sigma_t$  – залежне від часу стандартне відхилення;  $\zeta_t$  - випадкова величина, що має нормальний розподіл, середнє значення, що рівне 0, і стандартне відхилення, яке дорівнює 1. При цьому залежне від часу стандартне відхилення описується рівнянням

$$\sigma_t^2 = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \dots + \beta_q \varepsilon_{t-q}^2 + \gamma_1 \sigma_{t-1}^2 + \dots + \gamma_q \sigma_{t-q}^2 \quad (1.8)$$

Тут  $\beta_0, \dots, \beta_q$  і  $\gamma_0, \dots, \gamma_q$  - коефіцієнти. Рівняння (1.8) називається моделлю GARCH ( $p, q$ ) і має два параметри:  $p$  характеризує порядок авторегресії квадратів залишків;  $q$  - кількість попередніх оцінок залишків.

Найбільш часте застосування дана модель отримала в фінансовому секторі, де за допомогою неї моделюється волатильність. На сьогоднішній день існує ряд модифікацій моделі під назвами NGARCH, IGARCH, EGARCH, GARCH-M і інші.

*Авторегресійна модель з розподіленим лагом* (autoregressive distributed lag models, ARDLM). Часто при моделюванні процесів на досліджувану змінну впливають не тільки поточні значення процесу, а й його лаги, тобто значення



часового ряду, попередні досліджуваному моменту часу. Модель авторегресії розподіленого лага описується рівнянням

$$Z(t) = \varphi_0 + \varphi_1 Z(t - l - 1) + \dots + \varphi_p Z(t - l - p) + \varepsilon_t, \quad (1.9)$$

Тут  $\varphi_0, \dots, \varphi_p$  – коефіцієнти,  $l$  – величина лага. Модель (1.9) називається ARDLM ( $p, l$ ) і найчастіше застосовується для моделювання економічних процесів.

### 1.3.3. Моделі експоненціального згладжування

Моделі експоненціального згладжування розроблені в середині XX століття і до сьогоднішнього дня є широко поширеними в силу їх простоти і наочності.

Модель експоненціального згладжування (exponential smoothing, ES) застосовується для моделювання фінансових і економічних процесів. В основу експоненціального згладжування закладена ідея постійного перегляду прогнозних значень в міру надходження фактичних. Модель ES (1.10) задає експоненціально спадаючі ваги спостереженнями в міру їх старіння. Таким чином, останні доступні спостереження мають більший вплив на прогнозне значення, ніж давніші спостереження [51].

Функція моделі ES має вигляд

$$\begin{aligned} Z(t) &= S(t) + \varepsilon_t, \\ S(t) &= \alpha * Z(t - 1) + (1 - \alpha) * S(t - 1), \end{aligned} \quad (1.10),$$

де  $\alpha$  – коефіцієнт згладжування,  $0 < \alpha < 1$ ; початкові умови визначаються як  $S(1) = Z(0)$ .

У даній моделі кожне наступне згладжене значення  $S(t)$  є зваженим середнім між попереднім значенням часового ряду  $Z(t)$  і попереднього згладженого значення  $S(t - 1)$ .

**Модель Хольта** або подвійне експоненціальне згладжування застосовується для моделювання процесів, що мають тренд. В цьому випадку в моделі необхідно розглядати дві складові: рівень і тренд. Рівень і тренд згладжуються окремо

$$(t) = S(t) + \varepsilon_t,$$

$$(t) = \alpha * Z(t - 1) + (1 - \alpha) * S(t - 1) - B(t - 1),$$

$$B(t) = \gamma * (S(t - 1) - S(t - 2)) + (1 + \gamma) * B(t - 1).$$

де  $\alpha$  – коефіцієнт згладжування рівня, як і в моделі (1.8),  $\gamma$  - коефіцієнт згладжування тренда.

### 1.3.4. Моделі на основі нейронних мереж

В даний час найпопулярнішою серед структурних моделей є модель на основі штучних нейронних мереж (artificial neural network, ANN). Нейронні мережі складаються з нейронів (рис. 1.7).

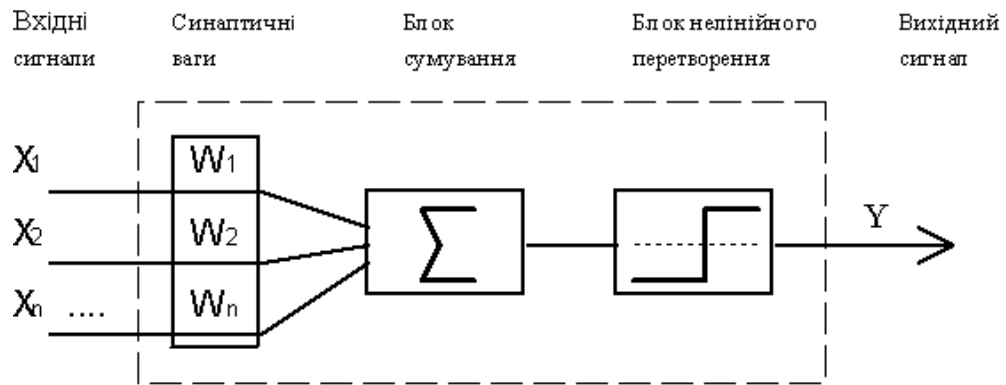


Рис. 1.7. Нелінійна модель нейрона

Модель нейрона можна описати парою рівнянь

$$U(t) = \sum_{i=1}^m \omega_i * Z(t - i) + b,$$

$$Z(t) = \varphi(U(t)),$$

де  $Z(t - 1), \dots, Z(t - m)$  – вхідні сигнали;  $\omega_1 \dots \omega_m$  - синаптичні ваги нейрона;  $b$  - поріг;  $\varphi(U(t))$ - функція активації.

Функція активації буває трьох основних типів:

- Функція одиничного стрибка;
- Кусково-лінійна функція;
- Сигмоїдальна функція.

Спосіб зв'язку нейронів визначає архітектуру нейронної мережі. Залежно від способу зв'язку нейронів мережі діляться на

- Одношарові мережі прямого поширення,
- Багатошарові мережі прямого поширення,
- Рекурентні мережі

Таким чином, за допомогою нейронних мереж можливо моделювання нелінійної залежності майбутнього значення часового ряду від його фактичних значень і від зовнішніх чинників. Нелінійна залежність визначається структурою мережі і функцією активації.

### 1.3.5. Моделі на базі ланцюгів Маркова

Моделі прогнозування на основі ланцюгів Маркова (Markov chain model) припускають, що майбутній стан процесу залежить тільки від його поточного стану і не залежить від попередніх. У зв'язку з цим процеси, що моделюються ланцюгами Маркова, повинні відноситись до процесів з короткою пам'яттю.

Приклад ланцюга Маркова для процесу, що має три стану, представлений на рис. 1.8.

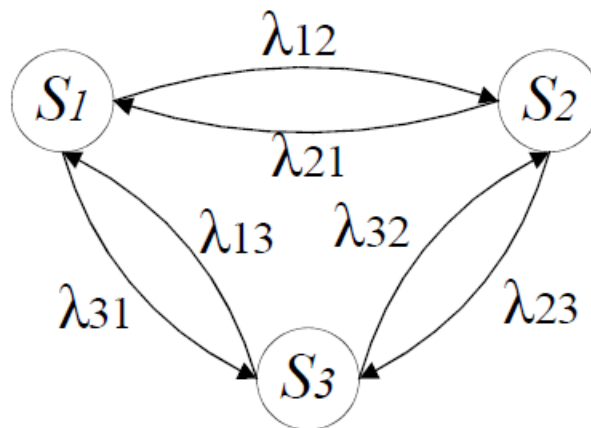


Рис. 1.8. Ланцюг Маркова з трьома станами

Тут  $S_1, \dots, S_3$  – стан процесу  $Z(t)$ ;  $\lambda_{12}$ - ймовірність переходу зі стану  $S_1$  в стан  $S_2$ ,  $\lambda_{23}$ - ймовірність переходу зі стану  $S_2$  в стан  $S_3$  і т. д. При побудові ланцюга Маркова визначається безліч станів і ймовірності переходів. Є поточний стан процесу  $S_i$ , то в якості майбутнього стану процесу вибирається такий стан  $S_j$ , ймовірність переходу в яке (значення  $\lambda_{ij}$ ) максимальна.

Таким чином, структура ланцюга Маркова і ймовірності переходу станів визначають залежність між майбутнім значенням процесу і його поточним значенням.

### **1.3.6. Інші моделі та методи прогнозування**

Крім класів моделей прогнозування, розглянутих вище, існують менш поширені моделі і методи прогнозування. Головним недоліком моделей і методів, згаданих у цьому розділі, є недостатня методологічна база, тобто недостатньо докладний опис можливостей як моделей, так і шляхів визначення їх параметрів. Крім того, у відкритому доступі можна знайти лише невелику кількість статей, присвячених застосуванню цих методів [52].

*Метод опорних векторів* (support vector machine, SVM) застосовується, наприклад, для прогнозування руху ринків і цін на електроенергію. В основу методу покладена класифікація, вироблена за рахунок переведення вихідних часових рядів, представлених в вигляді векторів, в простір більш високої розмірності і пошуку розділяючої гіперплощини з максимальним зазором в цьому просторі. Завдання прогнозування вирішується таким чином, що на етапі навчання класифікатора виявляються незалежні змінні (зовнішні фактори), майбутні значення яких визначають в якій з визначених раніше підкласів потрапить прогноз  $Z(t)$ .

*Структурні моделі CART* (classification and regression trees, CART) розроблені для моделювання процесів, на які надають вплив, як безперервні зовнішні фактори, так і категоріальні. Якщо зовнішні фактори, що впливають на процес  $Z(t)$ , неперервні, то використовуються регресійні дерева; якщо чинники категоріальні, то – класифікаційні дерева. У разі, якщо необхідно враховувати фактори обох типів, то використовуються змішані класифікаційно-регресійні дерева.

**Модель на основі функцій передачі** (transfer function, TF) застосовується для прогнозування процесу  $Z(t)$  з урахуванням зовнішнього фактора  $X(t)$ . Рівняння, що відображає залежність майбутнього значення має вигляд

$$Z(t) = v(B)X(t) + \eta(t),$$

де  $B$  – оператор зсуву  $BZ(t) = Z(t - 1), \dots, B^k Z(t) = Z(t - k)$ . Часовий ряд  $\eta(t)$  характеризує зовнішнє збурення. При цьому функція  $v(B)$  має вигляд

$$v(B) = v_0 + v_1 B + v_2 B^k + \dots \quad (1.11)$$

Коефіцієнти функції (1.11)  $v_i$  описують динамічні відносини між процесами  $Z(t)$  і  $X(t)$ .

#### 1.4. Оцінка помилки прогнозування часового ряду

Для оцінки якості обраної моделі, необхідно проаналізувати систему показників, що характеризують як адекватність моделі, так і її точність.

Точність прогнозу визначає величина помилки (похибки) прогнозу. Помилка прогнозу  $\varepsilon_t$  – значення, що характеризує розбіжність між фактичним і прогнозним значеннями показника:

$$\varepsilon_t = \hat{Z}(t) - Z(t),$$

де  $\hat{Z}(t)$  – прогнозне значення

$Z(t)$  – фактичне значення

Тоді формули для оцінок помилки прогнозування часових рядів для  $N$  відліків можна записати в наступному вигляді:

MAPE – середня абсолютна помилка у відсотках:

$$MAPE = \frac{1}{N} \sum_{t=1}^N \frac{|Z(t) - \hat{Z}(t)|}{Z(t)} * 100\% \quad (1.12)$$

Дана оцінка (1.12) застосовується для часових рядів, фактичні значення яких значно більше 1. Наприклад, оцінки помилки прогнозування енергоспоживання майже у всіх статтях наводяться як значення MAPE.

Для рядів, що містять значення близькі до нуля, застосовують наступну оцінку помилки прогнозу.

MAE – середня абсолютна помилка:

$$MAE = \frac{1}{N} \sum_{t=1}^N |Z(t) - \hat{Z}(t)| \quad (1.13)$$

Для оцінки помилки прогнозування цін на РДН коректніше використовувати MAE (1.13).

При цьому, переглядаючи статті, можна скласти загальне враження про помилки прогнозування енергоспоживання, для якого MAPE зазвичай коливається від 1 до 5%; або помилки прогнозування цін на електроенергію, для якого MAPE коливається від 5 до 15% залежно від періоду і ринку.

Крім зазначених іноді використовують інші оцінки помилки, наведені нижче.

MSE – середньоквадратична помилка:

$$MSE = \frac{1}{N} \sum_{t=1}^N (Z(t) - \hat{Z}(t))^2 \quad (1.14)$$

RMSE – квадратний корінь з середньоквадратичної помилки:

$$RMSE = \sqrt{MSE} \quad (1.15)$$

ME – середня помилка:

$$ME = \frac{1}{N} \sum_{t=1}^N (Z(t) - \hat{Z}(t))$$

SD – стандартне відхилення:

$$SD = \frac{1}{N} \sum_{t=1}^N (\hat{Z}(t) - ME)^2,$$

де ME – середня помилка, визначена за формулою вище.

## **Висновки до розділу 1**

1. Для учасників оптового ринку електроенергії і потужності найбільш актуальні коротко- і середньострокові прогнози ціни на електроенергію на ринку на добу вперед. Точне прогнозування майбутніх цінових показників РДН дозволяє споживачам раціонально закуповувати необхідні обсяги електроенергії, генераціям – оптимізувати склад і периметрів генеруючого обладнання.

2. Для вирішення завдань прогнозування ціни РДН зазвичай використовують лінійні регресійні моделі, засновані на аналізі часових рядів. Основним недоліком лінійних моделей є те, що вони описують саме лінійну залежність між впливають факторами і результатом, в той час як прогнозування ринку електроенергії призводить до більш складних залежностей. Іншим суттєвим недоліком лінійних моделей є чутливість до шумів в даних.

3. Література з прогнозування цін зазвичай поділяється на п'ять областей: теоретичні моделі ігор, фундаментальні методи, що використовують фізичні та економічні фактори, моделі з використанням статистичних властивостей торгівлі електроенергією для оцінки ризику та похідних інструментів, статистичні моделі, що містять часові ряди та економетричні моделі, та методи штучного інтелекту. Для прогнозування цін на день вперед статистичні методи та штучний інтелект показали найкращі результати.

4. Серед статистичних методів часто використовують використовується регресійний аналіз. Метою регресійного аналізу є визначення залежності між вихідної змінної і зовнішніми факторами. При цьому коефіцієнти регресії можуть визначатися методом найменших квадратів або методом максимальної правдоподібності.

5. В даний час найпопулярнішою серед структурних моделей є модель на основі штучних нейронних мереж (ANN), за допомогою яких можливе моделювання нелінійної залежності майбутнього значення часового ряду від його фактичних значень і від зовнішніх чинників. Нелінійна залежність визначається структурою мережі і функцією активації.

## РОЗДІЛ 2

### ОПИС РОЗРОБЛЕНОЇ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ

В даному розділі описано розв’язання двох задач: прогнозування споживання електроенергії на добу наперед та прогнозування ціни на електроенергію. Для аналізу обрано угорську енергосистему в силу можливої та перспективної міждержавної передачі електричної енергії в Україну з метою отримання прибутку та підтримання стабільної роботи мережі.

Розділ складається з 5 підрозділів. У підрозділі 2.1 наведені джерела отримання даних, серед яких основним є MAVIR – угорський оператор системи передач. Вказано структуру та обсяг вхідних показників. У підрозділі 2.2 описано процес розробки моделі прогнозування споживання, викладена математична постановка задачі. У підрозділі 2.2.1 проведено первинний статистичний аналіз споживання електричної енергії в Угорщині за останні три роки. У підрозділі 2.2.2 представлено загальне бачення моделі прогнозування споживання на добу наперед. У підрозділі 2.2.3 викладено формалізацію моделі та уточнення параметрів за допомогою кореляційного, регресійного та авторегресійного аналізу. У підрозділі 2.2.4 містяться результати роботи розробленого алгоритму у вигляді графіків та порівняльних таблиць.

У підрозділі 2.3 описано процес розробки моделі прогнозування ціни на електроенергію на добу наперед на угорському ринку. У підрозділі 2.3.1 проведено первинний статистичний аналіз ціни на електричну енергію. У підрозділі 2.3.2 встановлено залежності між вхідними даними, визначені зовнішні фактори, що включені в модель. Підрозділ 2.3.3 складається з таблиць, що містять уточнені коефіцієнти моделі. У підрозділі 2.3.4 наведено результати прогнозування ціни на добу наперед та проведено оцінку точності роботи моделі.

У підрозділі 2.4 наведено схему алгоритму ідентифікації моделей. У підрозділі 2.5 описано мову та середовище програмування. У підрозділі 2.6 підведені підсумки та наведені висновки за розділом.



## 2.1. Опис та підготовка вхідних даних

Ринки електроенергії Європи є лібералізованими. Електроенергія як товар торгується на біржі, одною із яких є НУРХ.

На НУРХ беруть участь 4 країни: Угорщина, Чехія, Словаччина та Румунія. Як видно за РИС динаміка цін в даних країнах однакова.

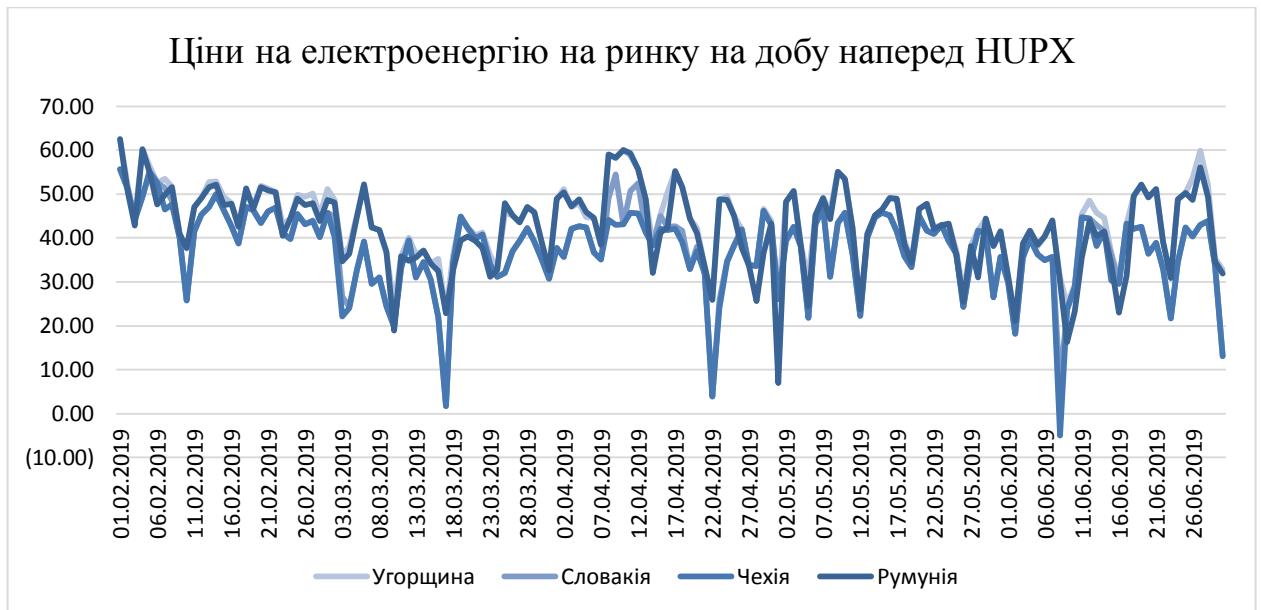


Рис 2.1. Динаміка цін на суміжних ринках електроенергії за 4 місяці

Для прикладу обрано угорську енергосистему, оскільки в силу географічного положення в даній країні є мережеве з'єднання з українською енергосистемою через Бурштинський острів.

Імпорт електроенергії через Бурштинський острів є наступним кроком в інтеграції України в європейські енергетичні ринки та гарантією справедливих цін для українського споживача.

Розширення ENTSO-E, європейської мережі операторів систем передач, шляхом підключення української енергосистеми, забезпечить інтеграцію українського ринку електроенергії в внутрішній енергетичний ринок ЄС. В даний час ENTSO-E являє собою 43 оператора систем передачі електроенергії з 36 країн Європи.

В цілому, при розгляді літератури з прогнозування на день наперед, вхідні дані були запропоновані в якості значущих змінних, таких як, температура, ціни на газ і вугілля, навантаження у мережі, доступна потужність, хмарність та інші [1].

Збір даних проводився з таких джерел:

- 1) MAVIR – угорський оператор системи передач [53].
- 2) ENTSOE – європейська мережа системних операторів передачі електроенергії. Забезпечення взаємодії системних операторів загальноєвропейського і регіонального рівня [54].
- 3) Wattsight – організація, що надає прогнози основних показників і ціни, адаптовані до потреб ринку електроенергії в короткостроковій, середньостроковій і довгостроковій перспективі [55].
- 4) HUPX – платформа, що містить біржові дані та ціни [56].
- 5) Pogodaiklimat – платформа, що містить погодні показники (температуру, вологість, хмарність, опади та інше). В наявності історичні та прогнозні дані [57].
- 6) Timeanddate – календар державних свят [58].
- 7) Open Power System Data – містить первинні джерела даних, які використовуються для моделювання енергосистеми Європи [59].

Зокрема, для розгляду та ідентифікацій моделі обрано дані у період з 30.10.2017 по 10.01.2020.

Важливо відзначити, що вхідні дані в основному не обробляються. Викиди цін не усуваються, оскільки необхідно виявляти піки. Єдине перетворення даних – це інтерполяція даних із-за переходу на літній час.

Історичні дані представляють собою упорядкований набір вузлових параметрів ринку, зібраних за останні 3 роки, в погодинної прив'язці.

Відлік часу  $t = 0$  починається з 30.10.2017 00:00:00, що є понеділком.

Так як всі властивості часового ряду залежать від години доби та дня тижня, то введемо такі позначення для подальшого використання.

Година доби вираховується за формулами:

$$\begin{aligned} \text{mod} \left[ \frac{t}{24} \right] &= 0 \text{ – перша година доби} \\ \text{mod} \left[ \frac{t}{24} \right] &= 1 \text{ – друга година доби} \\ \text{mod} \left[ \frac{t}{24} \right] &= 2 \text{ – третя година доби} \\ \text{mod} \left[ \frac{t}{24} \right] &= 3 \text{ – четверта година доби} \\ \text{mod} \left[ \frac{t}{24} \right] &= 4 \text{ – п'ята година доби} \\ \text{mod} \left[ \frac{t}{24} \right] &= 5 \text{ – шоста година доби} \\ &\dots \\ \text{mod} \left[ \frac{t}{24} \right] &= 22 \text{ – двадцять третя година доби} \\ \text{mod} \left[ \frac{t}{24} \right] &= 23 \text{ – двадцять четверта година доби} \end{aligned} \tag{2.1}$$

де  $\text{mod}[\ ]$  – залишок від ділення

День тижня вираховується за формулами:

$$\begin{aligned} \text{mod} \left[ \frac{t}{168} \right] &= 0 \text{ – Понеділок} \\ \text{mod} \left[ \frac{t}{168} \right] &= 1 \text{ – Вівторок} \\ \text{mod} \left[ \frac{t}{168} \right] &= 2 \text{ – Середа} \\ \text{mod} \left[ \frac{t}{168} \right] &= 3 \text{ – Четвер} \\ \text{mod} \left[ \frac{t}{168} \right] &= 4 \text{ – П'ятниця} \\ \text{mod} \left[ \frac{t}{168} \right] &= 5 \text{ – Субота} \\ \text{mod} \left[ \frac{t}{168} \right] &= 6 \text{ – Неділя} \end{aligned} \tag{2.2}$$

Або за універсальними формулами:

$$H(t) = \text{mod} \left[ \frac{t}{24} \right]$$

$$W(t) = \text{mod} \left[ \frac{t}{168} \right] + 1$$

Святкові дні подаються на вхід окремим файлом, де

$$\text{mod} \left[ \frac{t}{168} \right] + 1 = 8 - \text{святкові дні}$$

$$\text{mod} \left[ \frac{t}{168} \right] + 1 = 10 - \text{робочі вихідні (відпрацювання)}$$

$$\text{mod} \left[ \frac{t}{168} \right] + 1 = 11 - \text{канікули}$$

Споживання електроенергії пов'язане з діяльністю людей (виробництвом) і пропорційне числу споживачів електроенергії.

Метою етапу розробки загальної моделі прогнозування є прийняття рішення про формування безлічі незалежних змінних для включення в модель. Для розглянутого класу об'єктів можна виділити наступний набір найбільш значущих параметрів: 1) споживання електроенергії одну, дві, три і чотири тижні тому (тобто вівторки точніше прогнозувати по попереднім вівторках, ніж по понеділках); 2) середня температура кількох часових відліків назад від прогнозованої точки, оскільки включення систем клімат-контролю залежить від поточних температурних умов; 3) тип дня : робочі, вихідні та святкові. Приклад святкових днів наведено у таблиці 2.1.

Таблиця 2.1

Приклад вхідних даних з особливими днями

| Дата       | День тижня | Назва свята     | Тип дня        | Позначення дня тижня |
|------------|------------|-----------------|----------------|----------------------|
| 01.11.2017 | Середа     | All Saints' Day | Святковий день | 8                    |
| 25.12.2017 | Понеділок  | Christmas Day   | Святковий день | 8                    |
| 26.12.2017 | Вівторок   | Boxing Day      | Святковий день | 8                    |

## Продовження таблиці 2.1

|            |           |  |                |    |
|------------|-----------|--|----------------|----|
| 01.01.2018 | Понеділок | New Year's Day                               | Святковий день | 8  |
| 10.03.2018 | Субота    | Extra Work Day                               | Відпрацювання  | 10 |
| 15.03.2018 | Четвер    | 1848 Revolution Memorial Day                 | Святковий день | 8  |
| 16.03.2018 | П'ятниця  | 1848 Revolution Memorial Day (Extra holiday) | Святковий день | 8  |
| 30.03.2018 | П'ятниця  | Good Friday                                  | Святковий день | 8  |
| 01.04.2018 | Неділя    | Easter Sunday                                | Святковий день | 8  |
| 02.04.2018 | Понеділок | Easter Monday                                | Святковий день | 8  |
| 21.04.2018 | Субота    | Extra Work Day                               | Відпрацювання  | 10 |
| 30.04.2018 | Понеділок | Labor Day Holiday                            | Святковий день | 8  |
| 01.05.2018 | Вівторок  | Labor Day / May Day                          | Святковий день | 8  |
| 21.05.2018 | Понеділок | Whit Monday                                  | Святковий день | 8  |
| 20.08.2018 | Понеділок | Hungary National Day                         | Святковий день | 8  |
| 13.10.2018 | Субота    | Extra Work Day                               | Відпрацювання  | 10 |
| 22.10.2018 | Понеділок | 1956 Revolution Memorial Day Holiday         | Святковий день | 8  |

Кожен алгоритм передбачає існування вхідних даних та в результаті роботи призводить до отримання певного результату. Робота кожного алгоритму відбувається шляхом виконання послідовності деяких елементарних дій.

Необхідною умовою, яка задовольняє алгоритм, є детермінованість, або визначеність. Це означає, що виконання команд алгоритму відбувається у єдиний спосіб та призводить до однакового результату для однакових вхідних даних.

Для коректної роботи розробленої інформаційної технології на вхід необхідно подати дані у вигляді, що наведено на рис 2.2.

| Date_time               | Time | Weekday | Month | Price | Forecast | Consumption | Temper | Wind     | Solar    | Nuclear  | Cloudiness | Heating | Windchill | Cooling | esidual  | Loarmal  | Biome    | ermal    | Ligni    | Termal | Gas | Hydro |
|-------------------------|------|---------|-------|-------|----------|-------------|--------|----------|----------|----------|------------|---------|-----------|---------|----------|----------|----------|----------|----------|--------|-----|-------|
| 2017-10-27 00:00:00 UTC | 0    | 5       | 10    | 47,26 | 39,3705  | 4355,931375 | 10,7   | 76,37119 | 0        | 1873,629 | 0,84       | 0,19    | 0         | 0       | 4308,284 | 203,5833 | 523,1082 | 585,4779 | 20,66403 |        |     |       |
| 2017-10-27 01:00:00 UTC | 1    | 5       | 10    | 33,06 | 31,28193 | 4165,211035 | 10,5   | 51,97044 | 0        | 1873,379 | 0,87       | 0,2     | 0         | 0       | 4141,854 | 201,8938 | 521,0727 | 522,92   | 20,65977 |        |     |       |
| 2017-10-27 02:00:00 UTC | 2    | 5       | 10    | 28,71 | 27,52458 | 3970,341196 | 10,4   | 30,18168 | 0        | 1873,379 | 0,89       | 0,2     | 0         | 0       | 3968,662 | 193,4463 | 546,2613 | 520,5139 | 20,65551 |        |     |       |
| 2017-10-27 03:00:00 UTC | 3    | 5       | 10    | 29,12 | 27,6101  | 3902,738106 | 10,3   | 33,29797 | 0        | 1873,879 | 0,91       | 0,21    | 0         | 0       | 3897,832 | 139,8051 | 548,0423 | 511,4911 | 20,65125 |        |     |       |
| 2017-10-27 04:00:00 UTC | 4    | 5       | 10    | 33,06 | 30,60998 | 4009,722267 | 10,1   | 40,80654 | 0        | 1870,879 | 0,92       | 0,21    | 0         | 0       | 3997,197 | 124,5997 | 538,374  | 521,9174 | 20,647   |        |     |       |
| 2017-10-27 05:00:00 UTC | 5    | 5       | 10    | 47,95 | 41,18106 | 4340,782677 | 9,7    | 48,25721 | 0        | 1870,129 | 0,92       | 0,22    | 0         | 0       | 4320,696 | 137,2709 | 520,055  | 630,5917 | 20,64274 |        |     |       |
| 2017-10-27 06:00:00 UTC | 6    | 5       | 10    | 70,04 | 55,77362 | 5122,294338 | 9,3    | 12,27565 | 0        | 1869,629 | 0,91       | 0,24    | 0         | 0       | 5138,079 | 171,483  | 538,374  | 666,4823 | 20,63848 |        |     |       |
| 2017-10-27 07:00:00 UTC | 7    | 5       | 10    | 81,97 | 68,69084 | 5456,340248 | 9,4    | 18,21345 | 0        | 1869,879 | 0,89       | 0,23    | 0         | 0       | 5466,076 | 192,6016 | 561,7815 | 715,4058 | 28,65864 |        |     |       |
| 2017-10-27 08:00:00 UTC | 8    | 5       | 10    | 87,09 | 72,18169 | 5443,741908 | 10,1   | 20,59067 | 9,588013 | 1870,129 | 0,89       | 0,21    | 0         | 0       | 5441,402 | 199,3595 | 561,5271 | 770,7455 | 30,94495 |        |     |       |
| 2017-10-27 09:00:00 UTC | 9    | 5       | 10    | 80,39 | 68,89985 | 5494,272819 | 11,3   | 33,2527  | 26,26087 | 1868,879 | 0,92       | 0,17    | 0         | 0       | 5462,488 | 203,1609 | 561,7815 | 836,5115 | 30,93856 |        |     |       |
| 2017-10-27 10:00:00 UTC | 10   | 5       | 10    | 77,63 | 66,49382 | 5525,848229 | 12,6   | 47,93927 | 45,6457  | 1870,129 | 0,94       | 0,13    | 0         | 0       | 5459,881 | 198,0924 | 563,5626 | 864,3819 | 26,34963 |        |     |       |
| 2017-10-27 11:00:00 UTC | 11   | 5       | 10    | 81,75 | 63,94809 | 5573,66014  | 13,6   | 64,83777 | 44,13996 | 1871,379 | 0,96       | 0,11    | 0         | 0       | 5492,19  | 204,428  | 559,7461 | 906,4882 | 26,34419 |        |     |       |
| 2017-10-27 12:00:00 UTC | 12   | 5       | 10    | 86,02 | 63,3678  | 5612,6743   | 14,1   | 191,4254 | 38,26019 | 1874,129 | 0,97       | 0,1     | 0,01      | 0       | 5410,385 | 204,8504 | 560,7638 | 874,2067 | 29,77424 |        |     |       |
| 2017-10-27 13:00:00 UTC | 13   | 5       | 10    | 73,07 | 61,09059 | 5604,377961 | 14     | 267,9482 | 25,47496 | 1874,379 | 0,99       | 0,1     | 0,02      | 0       | 5338,241 | 204,428  | 557,2018 | 839,7196 | 30,91302 |        |     |       |
| 2017-10-27 14:00:00 UTC | 14   | 5       | 10    | 63,95 | 60,5537  | 5577,149871 | 13,6   | 237,6823 | 24,85034 | 1875,379 | 1          | 0,11    | 0,02      | 0       | 5341,793 | 204,428  | 551,6043 | 787,588  | 30,90663 |        |     |       |
| 2017-10-27 15:00:00 UTC | 15   | 5       | 10    | 63,96 | 63,58615 | 5598,020282 | 12,8   | 270,1096 | 14,75635 | 1873,879 | 0,99       | 0,13    | 0,05      | 0       | 5340,219 | 202,7385 | 558,9828 | 816,6614 | 30,90024 |        |     |       |
| 2017-10-27 16:00:00 UTC | 16   | 5       | 10    | 62,26 | 69,73408 | 5582,140442 | 11,5   | 289,1179 | 3,994169 | 1873,629 | 0,98       | 0,16    | 0,1       | 0       | 5315,983 | 196,8253 | 541,4271 | 830,4964 | 30,89386 |        |     |       |
| 2017-10-27 17:00:00 UTC | 17   | 5       | 10    | 72,01 | 73,76638 | 5595,080853 | 10,4   | 301,96   | 0,757949 | 1872,379 | 0,88       | 0,19    | 0,12      | 0       | 5319,206 | 193,024  | 556,6929 | 803,428  | 30,88747 |        |     |       |
| 2017-10-27 18:00:00 UTC | 18   | 5       | 10    | 87,95 | 82,73042 | 5674,842513 | 9,9    | 266,0163 | 0        | 1872,879 | 0,7        | 0,21    | 0,1       | 0       | 5435,559 | 193,024  | 553,6398 | 718,6139 | 30,88108 |        |     |       |
| 2017-10-27 19:00:00 UTC | 19   | 5       | 10    | 80    | 88,8351  | 5528,246424 | 9,5    | 277,3866 | 0        | 1873,129 | 0,51       | 0,22    | 0,1       | 0       | 5277,482 | 195,9806 | 533,5398 | 670,4925 | 30,8747  |        |     |       |
| 2017-10-27 20:00:00 UTC | 20   | 5       | 10    | 70,08 | 68,70947 | 5290,683834 | 9,2    | 288,9948 | 0        | 1872,379 | 0,4        | 0,23    | 0,1       | 0       | 5028,201 | 200,6267 | 529,7234 | 698,9643 | 30,86831 |        |     |       |
| 2017-10-27 21:00:00 UTC | 21   | 5       | 10    | 59,64 | 57,32027 | 4970,739995 | 8,8    | 290,8035 | 0        | 1872,629 | 0,37       | 0,25    | 0,09      | 0       | 4706,338 | 201,8938 | 536,0841 | 671,094  | 30,86192 |        |     |       |
| 2017-10-27 22:00:00 UTC | 22   | 5       | 10    | 54,93 | 51,41813 | 4733,423905 | 8,6    | 306,9155 | 0        | 1873,379 | 0,34       | 0,25    | 0,08      | 0       | 4452,799 | 202,7385 | 535,5752 | 606,13   | 30,85554 |        |     |       |
| 2017-10-27 23:00:00 UTC | 23   | 5       | 10    | 43,05 | 45,63737 | 4591,119316 | 8,5    | 304,8815 | 0        | 1873,129 | 0,29       | 0,26    | 0,08      | 0       | 4312,418 | 204,0056 | 514,712  | 574,6505 | 27,42147 |        |     |       |
| 2017-10-28 00:00:00 UTC | 0    | 6       | 10    | 41,31 | 33,59157 | 4341,355726 | 8,4    | 306,3944 | 0        | 1872,629 | 0,23       | 0,27    | 0,08      | 0       | 4061,031 | 202,3161 | 499,7006 | 485,6258 | 25,13114 |        |     |       |
| 2017-10-28 01:00:00 UTC | 1    | 6       | 10    | 30,89 | 32,56523 | 4105,268887 | 8,1    | 294,3714 | 0        | 1872,629 | 0,16       | 0,28    | 0,09      | 0       | 3836,856 | 201,049  | 505,5525 | 467,5803 | 25,12594 |        |     |       |

Рис. 2.2. Вигляд вхідних даних

## 2.2. Модель прогнозування споживання електроенергії

### 2.2.1. Первинний аналіз вхідних даних

Для більш адекватного розуміння моделі проводимо попередній аналіз даних. В даній частині роботи використані такі *інструменти*, як кореляційні графіки, функції автокореляції, декомпозиція, залишковий та регресійний аналіз.

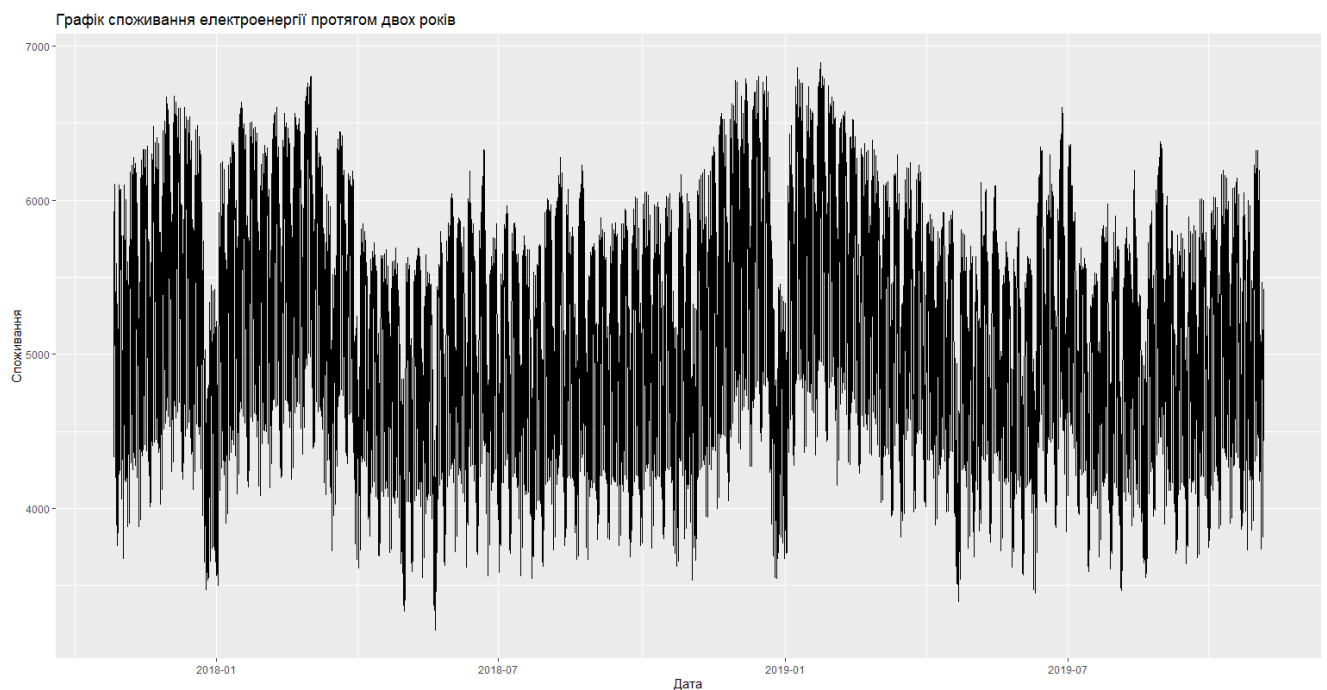


Рис. 2.3. Часовий ряд: споживання електроенергії за два останні роки

Вище наведено графік (рис. 2.3) споживання електричної енергії в Угорщині за останні два роки.

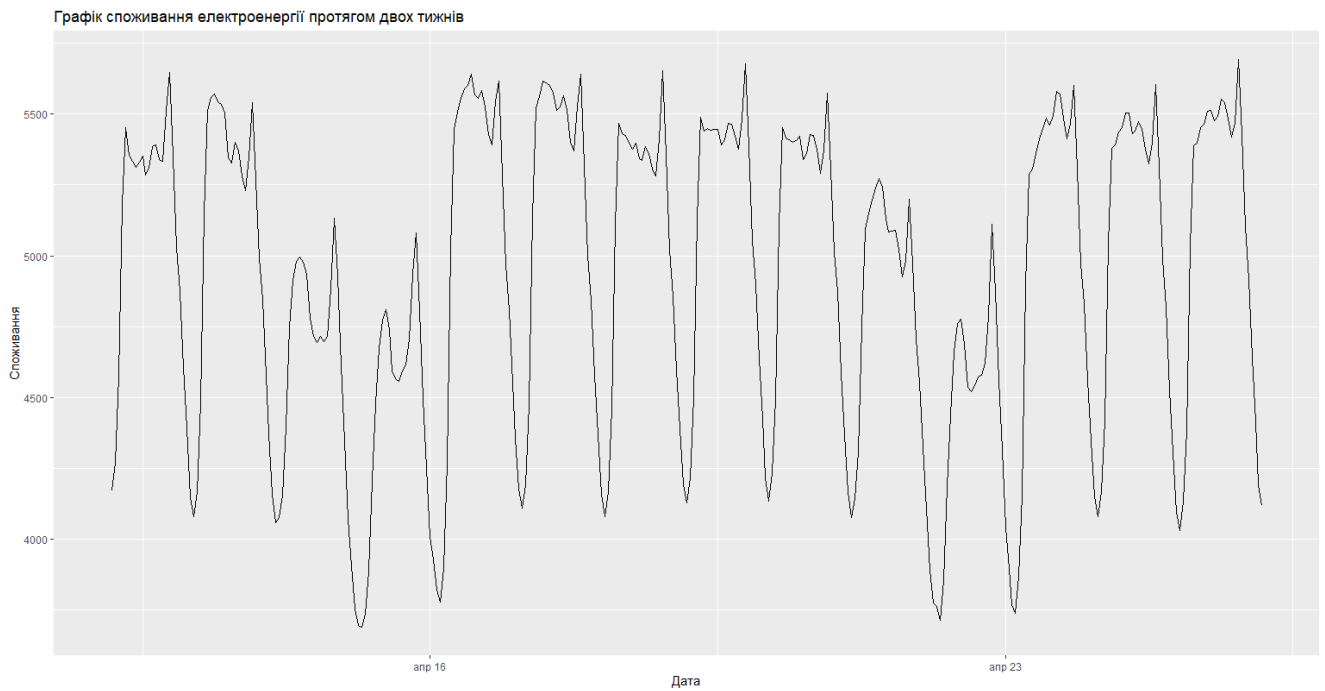


Рис. 2.4. Часовий ряд: споживання електроенергії за два тижні

Спостерігається (рис. 2.4), що волатильність (або зміна споживання від одного дня до наступного) досить висока.

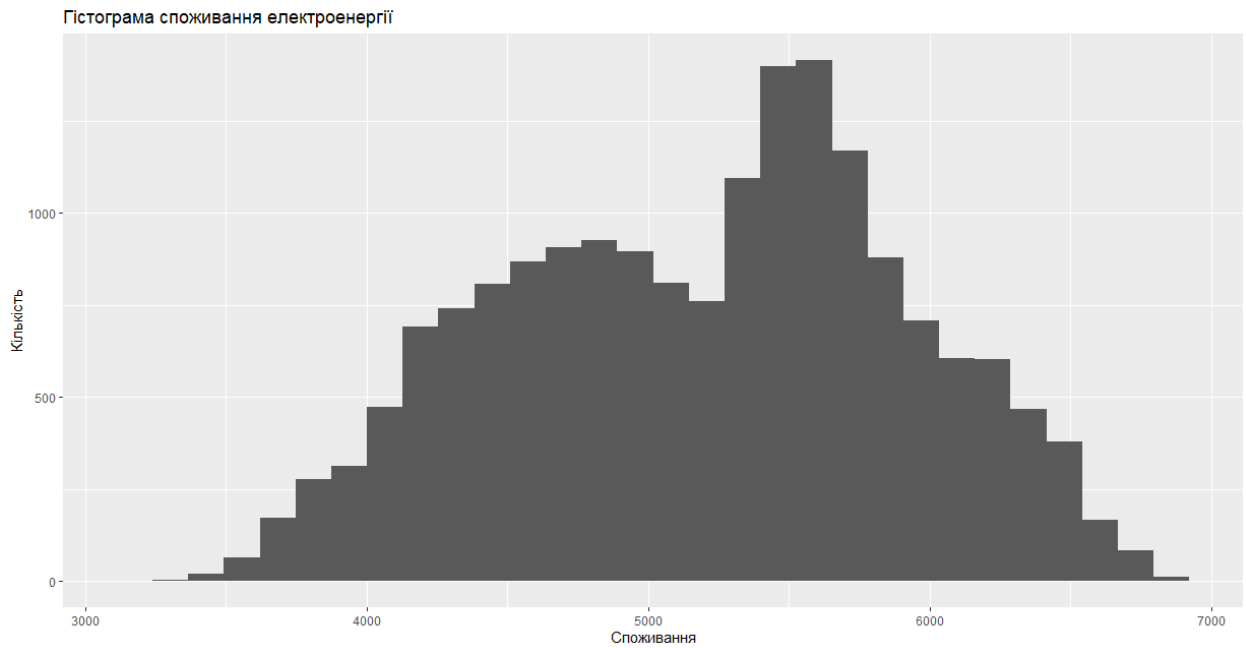


Рис. 2.5. Гістограма споживання електроенергії

Виходячи з аналізу гістограми наведеної на рис 2.5, можна припустити, що дані є неоднорідними. Гістограма частот є багатомодальною, що свідчить про різну природу даних. Тому доцільно розглянути розбиття даних.

В січні спостерігається помітне зменшення споживання. Дивлячись на дати, можна припустити, що цей період відповідає новорічним святкам, що є причиною зменшення попиту. Також можна спостерігати тенденцію до зростання з жовтня по лютий, і цей ефект може бути пов'язаний з більш холодною погодою, що призводить до збільшення споживання через потребу в опаленні. Щорічний сезонний ефект можна побачити більш чітко, глянувши на місячні значення наведені на рис 2.6.

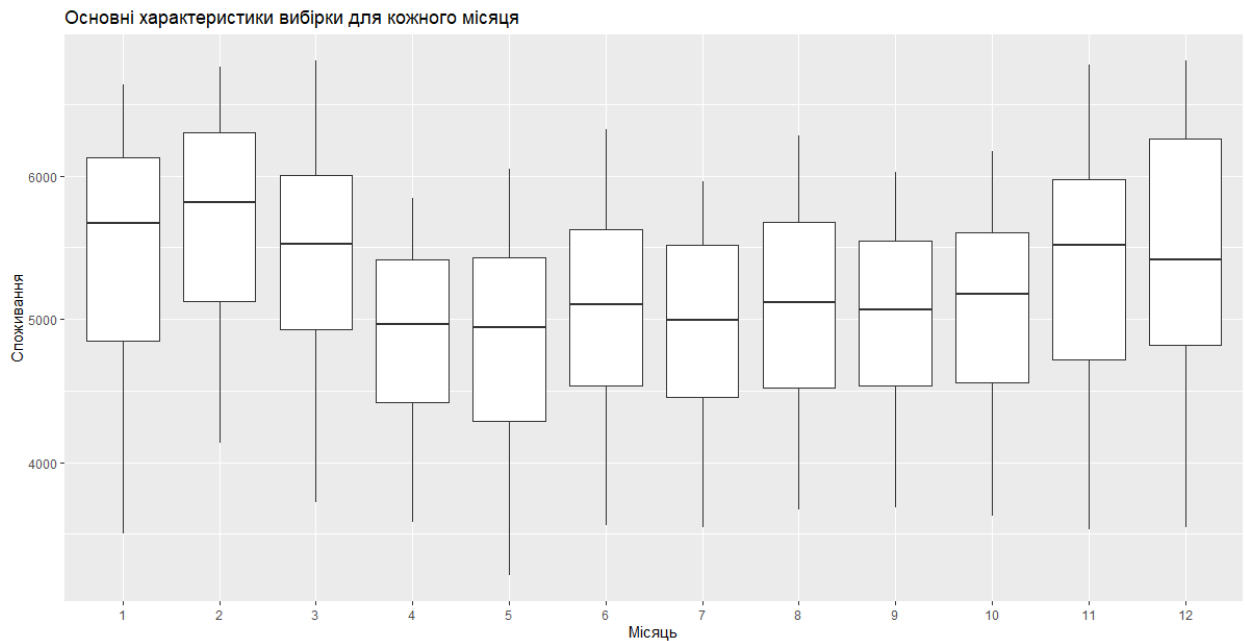


Рис. 2.6. Споживання електроенергії в місячному розбитті

Якщо подивитися на рис 2.7, на якому показані середньодобові значення та стандартні відхилення для кожного типу тижня, то видно зниження споживання по суботах, неділях та у святкові дні, що можливо, викликане зменшенням промислового споживання.



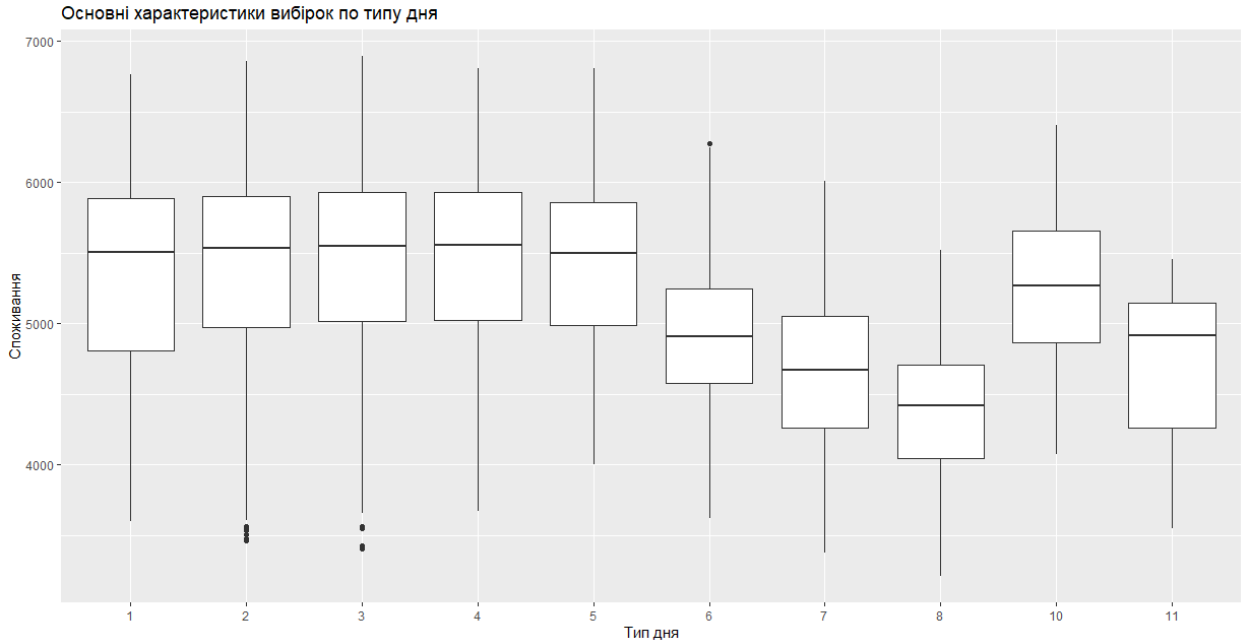


Рис. 2.7. Споживання електроенергії з залежності від типу дня

Щоденна сезонність відображена на рис 2.8.

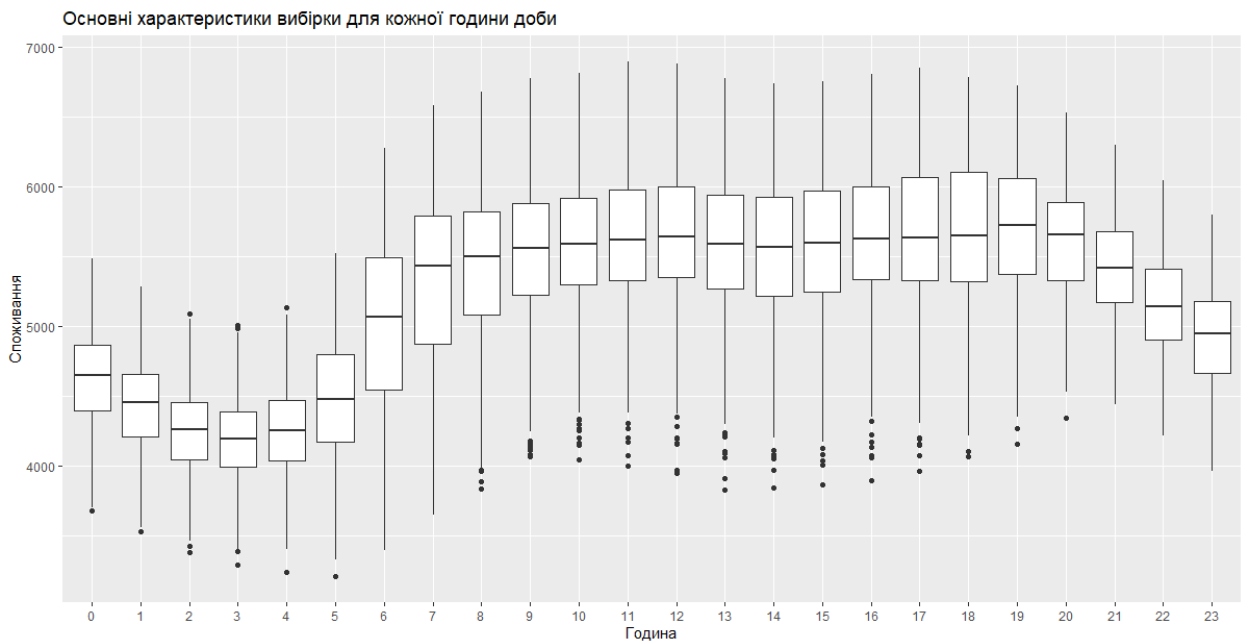


Рис. 2.8. Споживання електроенергії з залежності від години доби

Отже, ЧР споживання має декілька сезонних складових, які мають бути враховані в моделі.

Для подальшого аналізу даних і підтвердження щоденної і щотижневої сезонності, що наведена вище, можуть бути оцінені функції автокореляції (рис 2.9).

Автокореляція (ACF) – це кореляція функції з самою собою зміщеною на певну величину незалежної змінної.



Рис. 2.9. Графік автокореляційної функції ЧР споживання електроенергії

На рис 2.9 можна побачити лаг 24 і 168, що вказує на щоденну та тижневу сезонності, як очікувалося.

### 2.2.2. Побудова загальної моделі прогнозування споживання

Існує велика кількість різних способів розкладання даних, основна ідея яких полягає в тому, щоб розділити дані на їх складові, які зазвичай представляють собою трендову, сезонну і залишкову частини. При застосуванні адитивного розкладання передбачається, що дані складаються з суми його компонентів. З іншого боку, мультиплікативне розкладання передбачає, що дані представляють собою добуток компонентів, і в основному використовується, коли зміна даних збільшується або зменшується з часом.

Дані, які будуть використовуватися для прогнозування в даній роботі, не показують лінійної залежності в їх зміні в часі, тому пропустимо, що модель прогнозування споживання є адитивною, виду:

$$Y_1(t) = \sum_{i=0}^m F(X_i(t)) + \varepsilon(t), \quad (2.3)$$

де  $F(X_i(t))$  – функція залежності споживання від зовнішнього фактору  $X_i(t)$ ,  $\varepsilon(t)$  – похибка.

Припустимо, що  $\mathcal{E} = \{\varepsilon(t); t \in [1: T]\}$  має нормальний розподіл.

Подальша робота з ідентифікації моделі буде проходити наступним чином: знаходження фактору, що найбільше впливає на цільовий показник; побудова регресії; вилучення з ЧР отриманої складової; оцінка похибки. Дані етапи застосовуються до кожного значимого зовнішнього фактору.

Споживання в усіх сферах діяльності залежить від зміни температури, типу дня та періоду доби (день, ніч), що в свою чергу характеризується освітленістю. Нижче наведено графіки (рис 2.10 деяких перерахованих факторів.

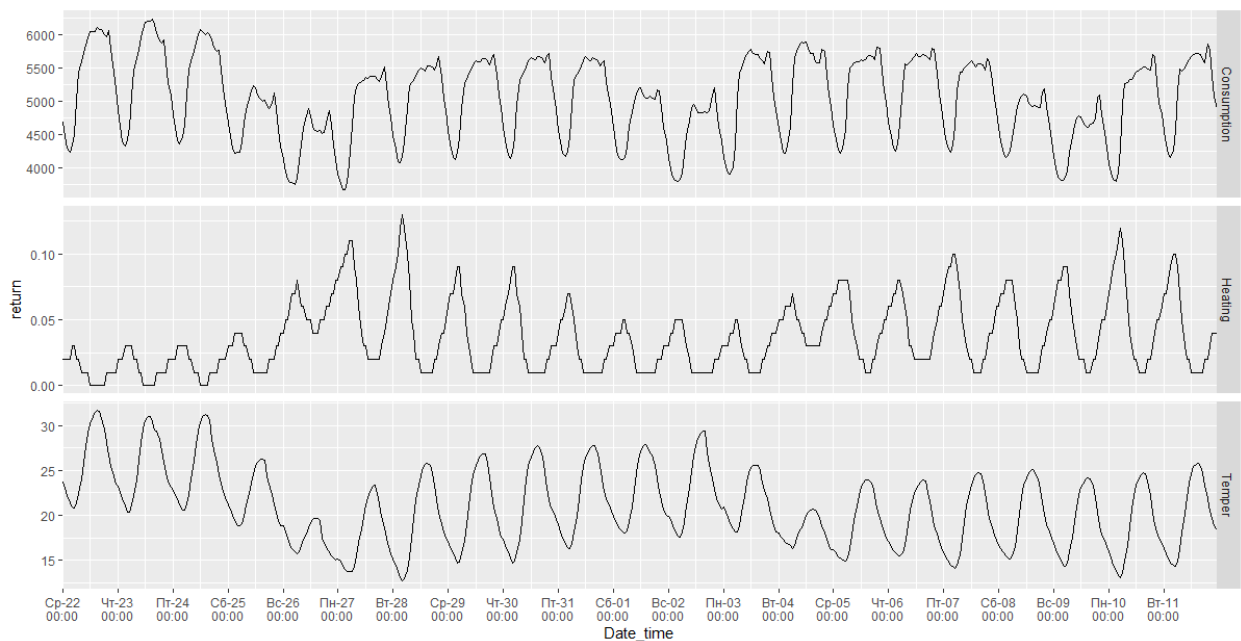


Рис. 2.10. Динаміка зміни факторів та цільового показника

Виходячи з коефіцієнтів кореляції, наведених у таблиці 2.2, в якості найбільш значущих чинників, що впливають на рівень споживання в моделі були залишені типи дня  $W(t)$ , година доби  $H(t)$ , середня температура в регіоні за одну годину  $T(t)$  та сонячність  $S(t)$ .

## Коефіцієнти кореляції споживання та зовнішніх факторів

| Фактор      | Споживання |
|-------------|------------|
| Година доби | 0,48       |
| День тижня  | 0,35       |
| Температура | 0,41       |
| Сонячність  | 0,26       |

Тому припустимо, що модель має такий вид (2.4):

$$\sum_{i=0}^m F(X_i(t)) = F(W(t)) + F(H(t)) + F_{temp}(T(t)) + F_{sun}(S(t)) + F_{prev}(t), \quad (2.4)$$

де  $F_{prev}(t) = F(Y_1(t - 24))$  – функція впливу змін попередньої доби на поточну, де 24 – лаг, знайдений вище автокореляційним аналізом

### 2.2.3. Формалізація моделі та уточнення параметрів

Одним з факторів, що більший вплив на електроспоживання є метеорологічний фактор. До метеофакторів відносять: температуру зовнішнього повітря, освітленість, вологість, швидкість вітру. Вони багато в чому визначають сезонні коливання і добову нерівномірність, а також нерегулярні коливання відхилення графіків споживання. Найбільш сильний вплив на споживання надає температура і освітленість. Вплив температури визначається витратою електроенергії на опалення будівель, вентиляцію, охолодження в холодильниках, кондиціонерах. Найбільш чутливий до температури витрата енергії в зимовий, опалювальний сезон, а також прилеглі до нього періоди. За існуючими оцінками, близько чверті частини енергетичного балансу йде на опалювальні потреби. Тому доцільно розглянути *залежність споживання від температури* (рис 2.11).

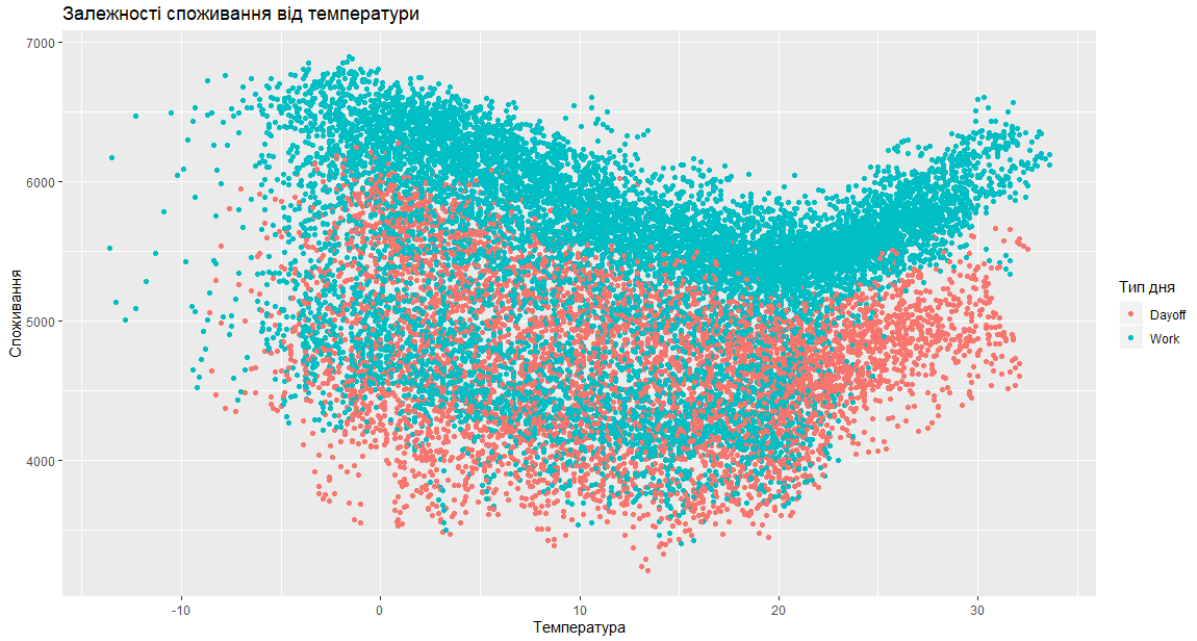


Рис. 2.11. Залежність споживання від температури, де червоні – вихідні дні, зелені – робочі

З аналізу рис 2.11 можна сказати, що залежність споживання від температури є нелінійною та данні є неоднорідними. Оскільки, вище було виявлено сезонні складові, доцільно розглянути дану залежність для кожного розбиття (рис 2.12).



Рис. 2.12. Залежність споживання від температури в рині години доби

Як видно з рис 2.12 залежність споживання від температури є нелінійною. Споживання зростає при температурі вище 20, коли починається підключення кондиціонерів та при температурі нижче 0, коли використовуються обігрівачі.

Графічне представлення регресійного аналізу наведено нижче (рис 2.13).

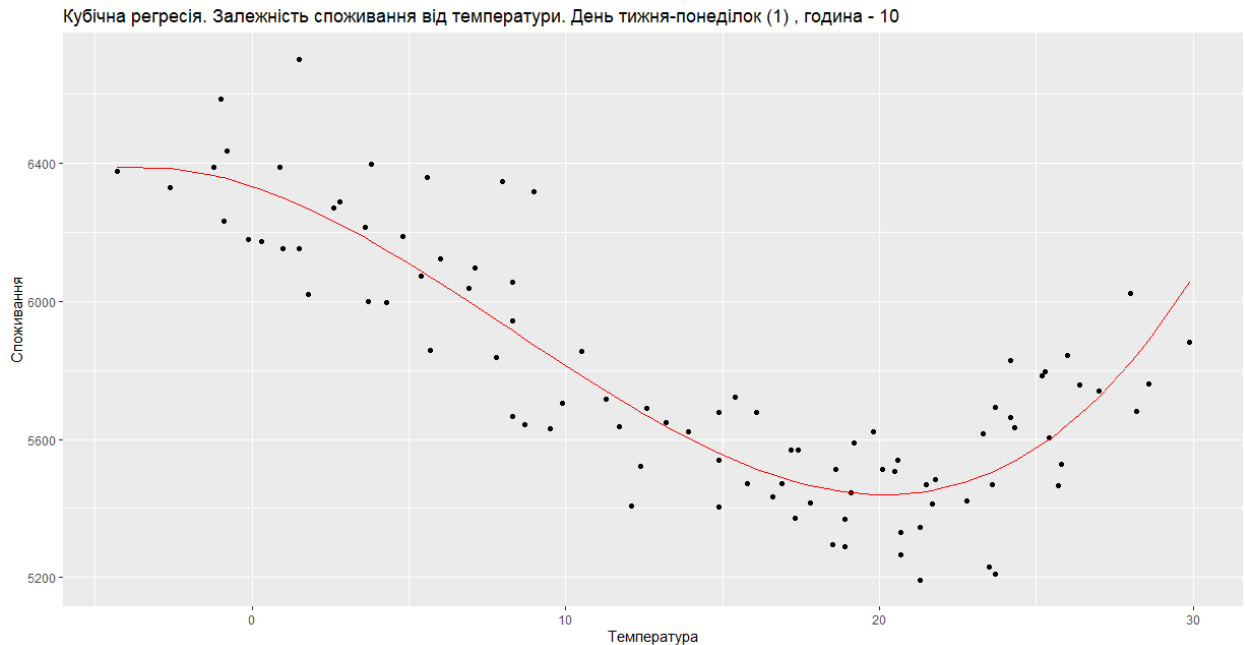


Рис. 2.13. Кубічна залежність споживання від температури (середа, 13 година)

Після проведення первинної обробки та аналізу даних було виявлено, що споживання має великі відмінності в робочі дні, вихідні та святкові. Доведено загальну тенденцію енергоспоживання: в понеділок – низьке споживання, до середи відбувається зростання споживання, до п'ятниці – спад, у вихідні та святкові дні споживання дуже низьке. Це пов'язано з великою кількістю споживачів – великих промислових підприємств. Також рівень споживання різний в залежності від години доби.

У зв'язку з цим дані були розбиті по днях тижня та годинам. Для побудови було створено 240 моделей (10 різних типів дня та 24 години в кожному).

З огляду на графік зображений на рис 2.13 обрана кубічна модель виду (2.5):

$$F_{temp}(T(t)) = a_{temp}(t)T^3(t) + b_{temp}(t)T^2(t) + c_{temp}(t)T(t) + d_{temp}(t), \quad (2.5)$$

де  $T(t)$  – температура у момент  $t$ , а коефіцієнти  $a_{temp}, b_{temp}, c_{temp}, d_{temp}$  знайдені за методом найменших квадратів та залежать від години доби  $H$  та від типу дня  $W$ . Дані коефіцієнти наведені у таблиці 2.3.

Таблиця 2.3

Коефіцієнти кубічної моделі (2.5)

| Тип дня | Година доби | a    | b     | c      | d       |
|---------|-------------|------|-------|--------|---------|
| 5       | 0           | 0,10 | -0,92 | -42,28 | 5141,80 |
| 5       | 1           | 0,09 | -0,70 | -44,48 | 4929,86 |
| 5       | 2           | 0,10 | -0,58 | -46,45 | 4724,44 |
| 5       | 3           | 0,10 | -0,52 | -45,75 | 4640,07 |
| 5       | 4           | 0,12 | -1,01 | -45,71 | 4754,14 |
| 5       | 5           | 0,13 | -1,75 | -46,46 | 5116,87 |
| 5       | 6           | 0,14 | -2,47 | -47,33 | 5858,63 |
| 5       | 7           | 0,13 | -2,27 | -47,54 | 6197,59 |
| 5       | 8           | 0,15 | -2,74 | -48,47 | 6296,95 |
| 5       | 9           | 0,16 | -3,36 | -42,73 | 6380,67 |
| 5       | 10          | 0,16 | -3,80 | -37,28 | 6441,29 |
| 5       | 11          | 0,16 | -4,13 | -32,25 | 6499,05 |
| 5       | 12          | 0,16 | -4,16 | -29,53 | 6520,67 |
| 5       | 13          | 0,15 | -4,08 | -26,22 | 6437,84 |
| 5       | 14          | 0,14 | -4,00 | -23,66 | 6394,04 |
| 5       | 15          | 0,14 | -3,94 | -25,50 | 6429,48 |
| 5       | 16          | 0,14 | -3,56 | -36,43 | 6517,72 |
| 5       | 17          | 0,16 | -4,09 | -35,53 | 6525,42 |
| 5       | 18          | 0,16 | -4,44 | -26,09 | 6447,16 |
| 5       | 19          | 0,13 | -3,41 | -25,13 | 6314,45 |
| 5       | 20          | 0,10 | -2,00 | -30,93 | 6088,43 |
| 5       | 21          | 0,12 | -1,76 | -40,30 | 5887,67 |
| 5       | 22          | 0,13 | -1,73 | -44,54 | 5656,81 |
| 5       | 23          | 0,09 | -0,72 | -45,11 | 5402,15 |
| 6       | 0           | 0,05 | 0,87  | -54,29 | 5021,27 |

Облік впливу метеофакторів вимагає вдосконалення, оскільки навіть невелика помилка в прогнозі може спричинити значні економічні витрати. Структура споживання електроенергії змінюється з часом, це відбивається в співвідношенні промислових і комунальних споживачів, що також призводить до підвищення впливу аспектів метеорологічного характеру. Кліматичні і географічні особливості, в сукупності з метеофакторами, виникненням природних і техногенних аномалій також призводять до необхідності їх аналізу.

Показники природного освітлення безпосередньо впливають на використання освітлювальних приладів. Зміна показників денної освітленості може призводити до різкого підвищення споживання електроенергії (в разі, наприклад, похмурої погоди, снігопадів та ін.). Добові коливання освітленості обумовлюють зміну значень навантаження енергетичних підприємств, особливо тих, значна частка потужності яких розподіляється по сектору промислових і побутових освітлювальних приладів. Таким чином, встановивши вплив освітленості на споживання енергії, можна корелювати ці значення.

Залежність *споживання від сонячності* відображена на рис 2.14

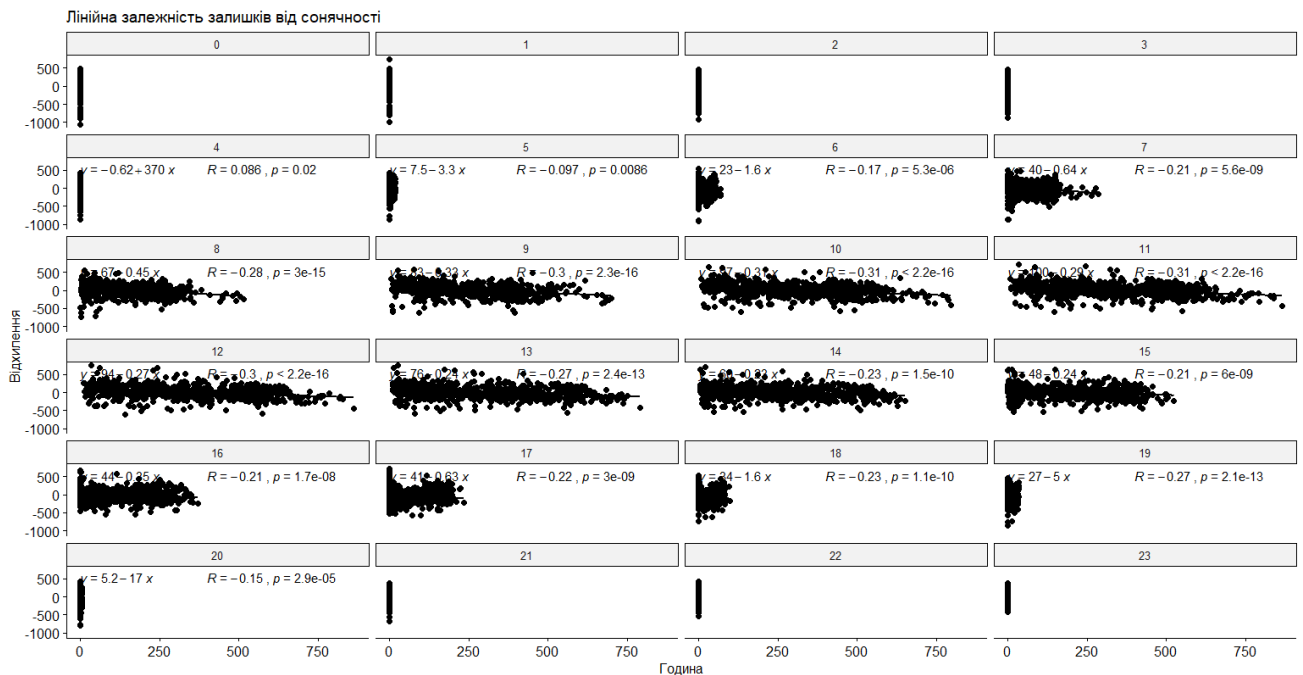


Рис. 2.14. Залежність споживання від сонячності



Більше детально лінійна залежність зображена на рис 2.15.

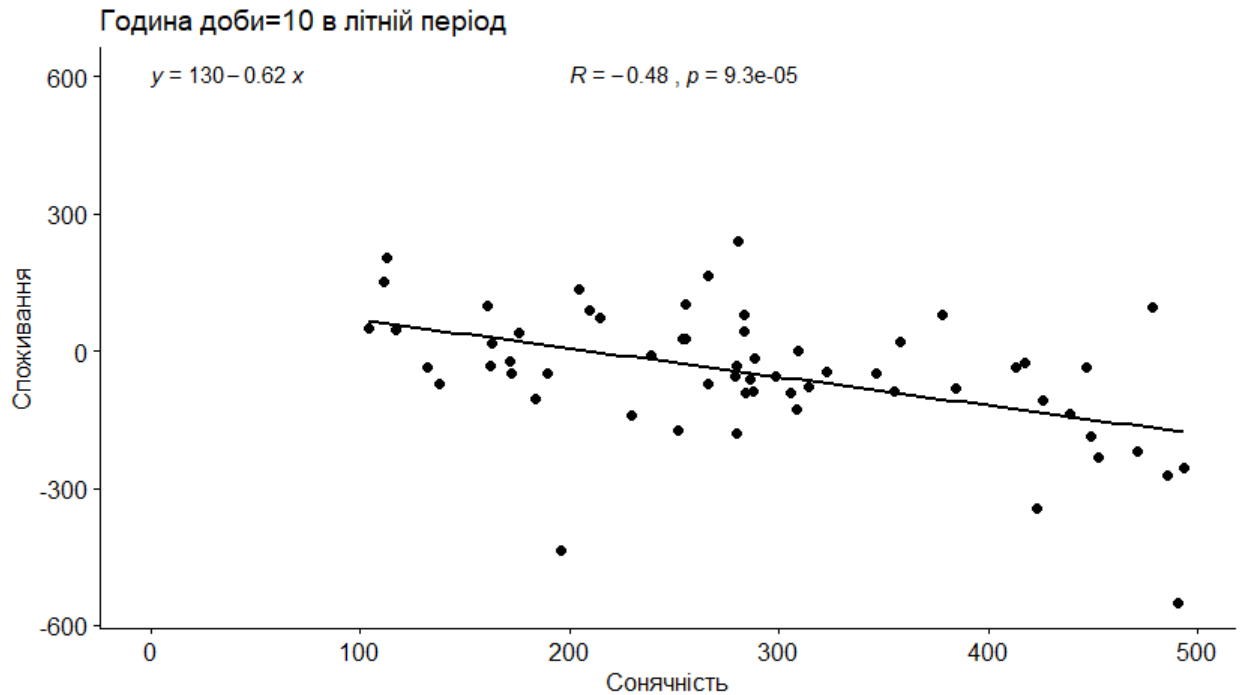


Рис. 2.15. Залежність споживання від сонячності о десятій годині

Залежність споживання від сонячності має лінійний вид:

$$F_{sun}(S(t)) = a_{sun}(t) + b_{sun}(t)S(t), \quad (2.6)$$

де  $S(t)$  – сонячність у момент  $t$ , а коефіцієнти  $a_{sun}, b_{sun}$  залежать від години доби  $H$ .

Сезонна зміна кількості освітленості найбільш помітно відбилася на енергоспоживанні в липні і серпні. Очевидно, що нерівномірність показників рівня сонячності впливає а рівень навантаження енергетичних мереж (в залежності від регіону).

Коефіцієнт кореляції має від'ємне значення, що означає наявність протилежної зв'язку: чим вище освітленість, тим нижче значення споживання.

В нічні години залежність відсутня. Коефіцієнти лінійної регресії (2.6) наведені в таблиці 2.4.

Коефіцієнти лінійної моделі (2.6) залежності споживання від сонячності

| Година доби | а     | б      |
|-------------|-------|--------|
| 0           | 0,08  | 0,00   |
| 1           | 0,34  | 0,00   |
| 2           | 0,54  | 0,00   |
| 3           | 0,63  | 0,00   |
| 4           | -0,62 | 369,99 |
| 5           | 7,49  | -3,34  |
| 6           | 22,91 | -1,61  |
| 7           | 39,81 | -0,64  |
| 8           | 67,09 | -0,45  |
| 9           | 83,19 | -0,33  |
| 10          | 96,61 | -0,31  |
| 11          | 99,97 | -0,29  |
| 12          | 93,50 | -0,27  |
| 13          | 76,11 | -0,24  |
| 14          | 59,88 | -0,23  |
| 15          | 47,91 | -0,24  |
| 16          | 43,88 | -0,35  |
| 17          | 40,95 | -0,63  |
| 18          | 33,76 | -1,56  |
| 19          | 27,23 | -4,97  |
| 20          | 5,23  | -16,96 |
| 21          | -0,44 | 0,00   |
| 22          | 0,58  | 0,00   |
| 23          | 0,00  | 0,00   |

Залежність від температури (2.5) та сонячності (2.6) об'єднаємо та позначимо:

$$\hat{Y}_{1_{global}}(t) = a_{temp}(t)T^3(t) + b_{temp}(t)T^2(t) + c_{temp}(t)T(t) + d_{temp}(t) + a_{sun}(t) + b_{sun}(t)S(t) \quad (2.7)$$

Для розробки технології було розглянуто дані за період останніх трьох років. За цей проміжок часу характеристики часового ряду змінюються. Тому доречно розглянути вплив споживання, наприклад, за останні 200 відліків на споживання у поточний момент.

Для цього розглянемо графік *часткової автокореляційної функції споживання* (рис 2.16), що залишилось після вилучення сезонних складових та вилучення залежності від метеофакторів.

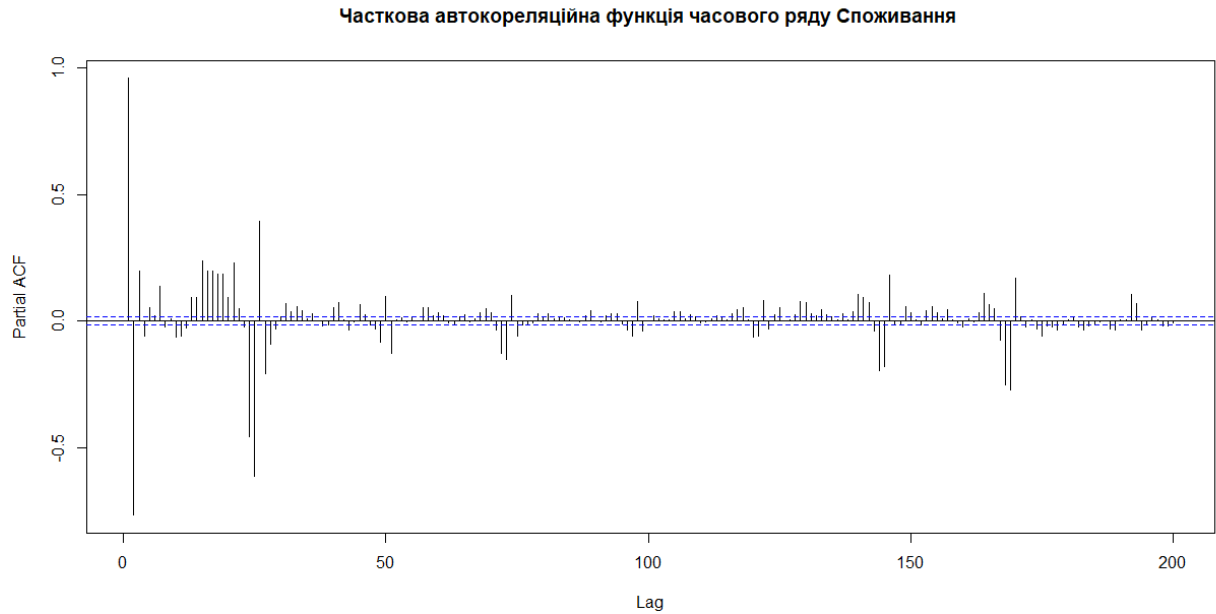


Рис. 2.16. Графік часткової автокореляційної функції споживання

Розгляд наведеного вище графіку часткової автокореляції залишків дозволяє зробити висновок, що найбільший вплив на залишки  $\Delta Y_1(t)$  мають значення залишків в момент  $(t - 24)$ . Дана залежність є лінійною (рис 2.17).

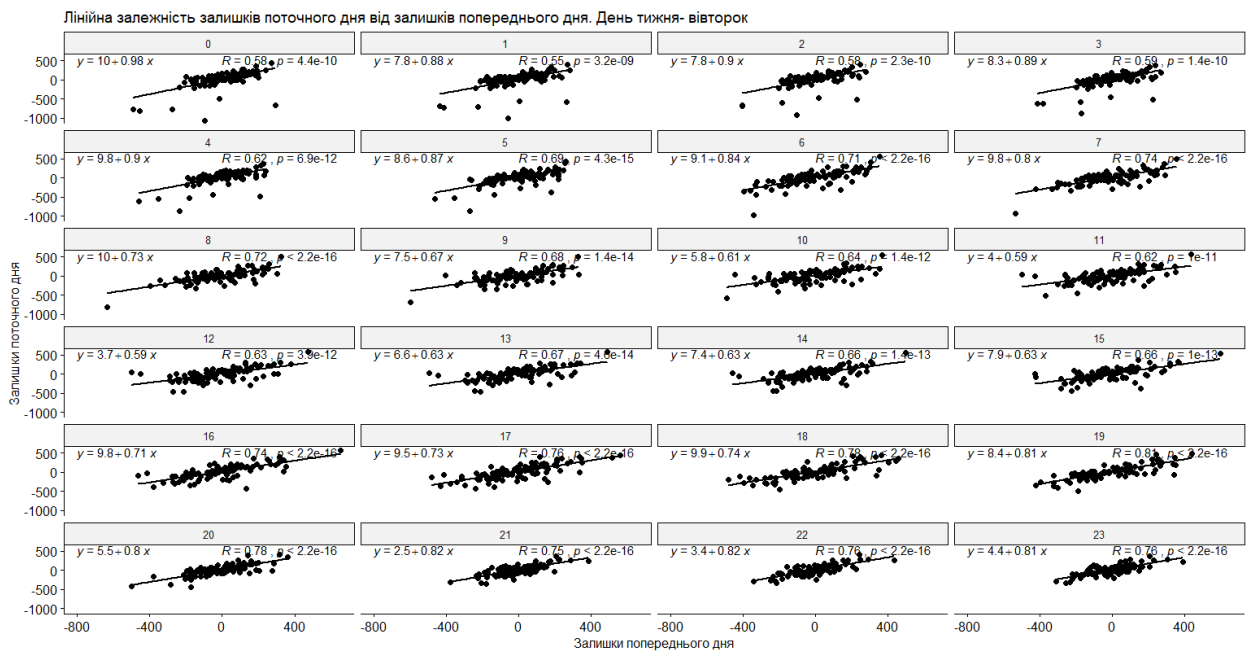


Рис. 2.17. Залежність споживання поточної доби від попередньої

Як видно з рис.2.17 залежність є лінійною, що відповідає наступному рівнянню (2.8):

$$F_{prev}(t) = a_{prev}(t) + b_{prev}(t)\Delta Y_1(t - 24), \quad (2.8)$$

$$\text{де } \Delta Y_1(t - 24) = Y_1(t - 24) - \hat{Y}_{1global}(t - 24)$$

де коефіцієнти  $a_{prev}, b_{prev}$  залежать від  $H$  та  $W$  та наведені у таблиці 2.5.

Таблиця 2.5

Коефіцієнти лінійної моделі (2.8) залежності споживання від попереднього дня

| Тип дня | Година доби | a      | b     |
|---------|-------------|--------|-------|
| 5       | 0           | -0,220 | 0,400 |
| 5       | 1           | -0,639 | 0,433 |
| 5       | 2           | -0,951 | 0,465 |
| 5       | 3           | -0,827 | 0,460 |
| 5       | 4           | -1,949 | 0,438 |
| 5       | 5           | -0,184 | 0,589 |
| 5       | 6           | -3,970 | 0,729 |
| 5       | 7           | -4,112 | 0,731 |
| 5       | 8           | -2,319 | 0,638 |
| 5       | 9           | -2,161 | 0,546 |
| 5       | 10          | -2,838 | 0,517 |
| 5       | 11          | -4,137 | 0,541 |
| 5       | 12          | -4,279 | 0,566 |
| 5       | 13          | -4,306 | 0,587 |
| 5       | 14          | -3,415 | 0,601 |
| 5       | 15          | -2,939 | 0,676 |
| 5       | 16          | -3,363 | 0,716 |
| 5       | 17          | -5,626 | 0,736 |
| 5       | 18          | -9,040 | 0,780 |
| 5       | 19          | -7,549 | 0,862 |
| 5       | 20          | -4,480 | 0,831 |

Тоді рівняння (2.4) можна представити в наступній формі:

$$\hat{Y}_1(t) = \hat{Y}_{1global}(t) + a_{prev}(t) + b_{prev}(t) \left( Y_1(t - 24) - \hat{Y}_{1global}(t - 24) \right) \quad (2.9)$$

Таким чином, рівняння (2.9) можна записати в розгорнутому вигляді:

$$\begin{aligned} \hat{Y}_1(t) = & a_{temp}(t)T^3(t) + b_{temp}(t)T^2(t) + c_{temp}(t)T(t) + d_{temp}(t) + \\ & + a_{sun}(t) + b_{sun}(t)S(t) + a_{prev}(t) + b_{prev}(t)[Y_1(t - 24) - \end{aligned} \quad (2.10)$$

$$-(a_{temp}(t - 24) T^3(t - 24) + b_{temp}(t - 24)T^2(t - 24) + c_{temp}(t - 24)T(t - 24) + d_{temp}(t - 24) + a_{sun}(t - 24) + b_{sun}(t - 24)S(t - 24))]$$

Отже, розроблена комбінована моделі виду (2.10) прогнозування споживання на добу наперед.

Загальні характеристики вхідних часових рядів відображено у таблиці 2.6.

Таблиця 2.6

Статистичні дані по вхідним даним для прогнозування споживання

| Показник        | Максимум | Середнє значення | Медіана | Мінімум | Стандартне відхилення |
|-----------------|----------|------------------|---------|---------|-----------------------|
| Час             | 23,00    | 11,50            | 11,00   | 0,00    | 6,90                  |
| Місяць          | 12,00    | 6,57             | 7,00    | 1,00    | 3,45                  |
| Сонячність      | 865,24   | 99,12            | 2,79    | 0,00    | 159,07                |
| Температура     | 33,60    | 12,27            | 12,30   | -13,60  | 9,28                  |
| Рівень спеки    | 0,96     | 0,24             | 0,13    | 0,00    | 0,24                  |
| Тип дня         | 11,00    | 4,34             | 4,00    | 1,00    | 2,28                  |
| Споживання      | 6894,00  | 5200,00          | 5289,00 | 3210,00 | 718,12                |
| Прогноз з МАВІР | 7079,80  | 5181,80          | 5292,90 | 802,20  | 790,72                |

Прилад роботи розробленої моделі наведено в наступному розділі

#### 2.2.4. Приклад роботи. Аналіз результатів

Результатом дипломної роботи є програмне забезпечення, у якому реалізовані моделі прогнозування споживання електроенергії та ціни на даний продукт.

На рис 2.18 та рис 2.19 відображені результати прогнозування споживання на добу наперед за розробленою моделлю (2.10) та фактичні значення.

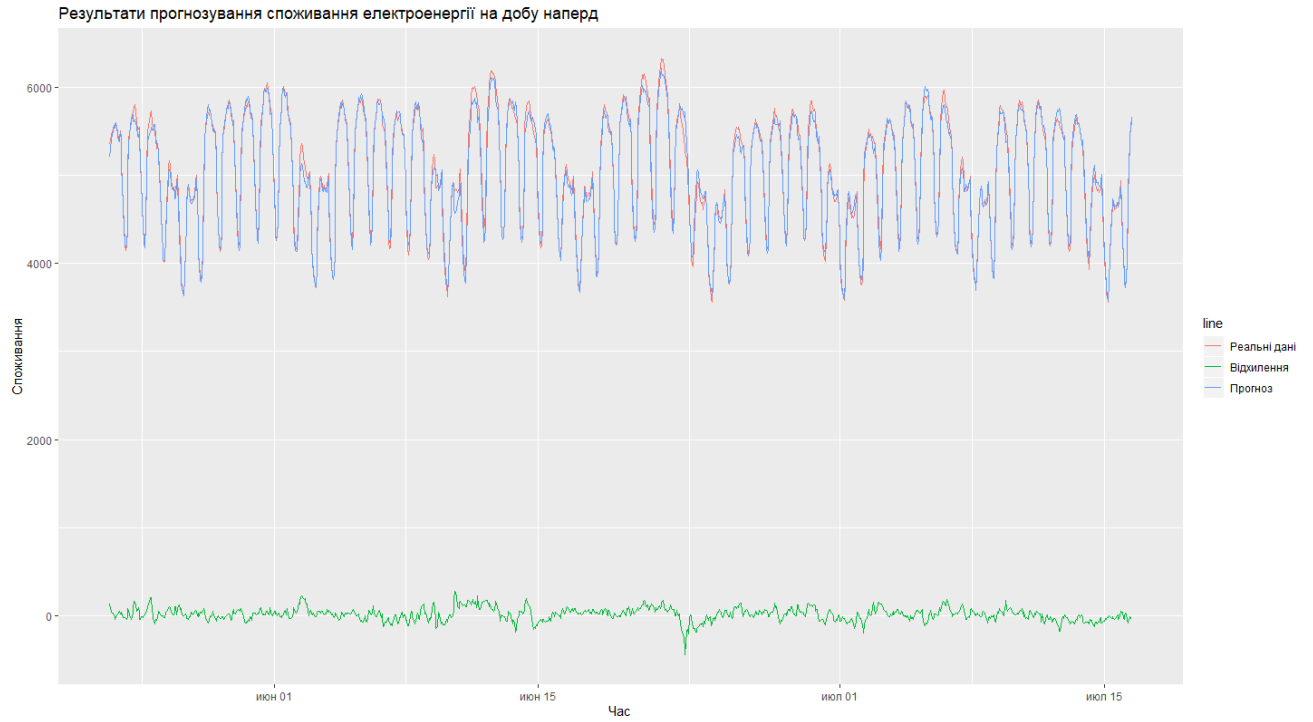


Рис. 2.18. Результат прогнозу споживання на добу наперед

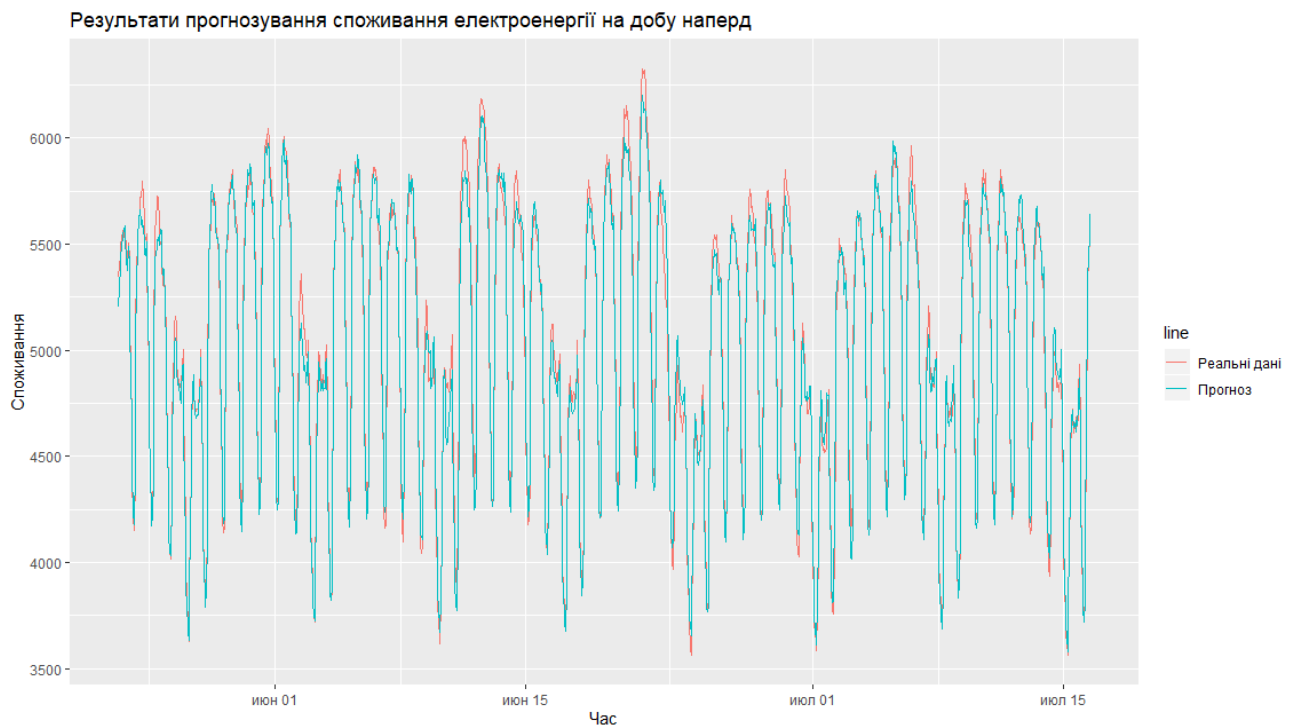


Рис. 2.19. Результат прогнозу споживання на добу наперед

При оцінці точності використовується прогноз на весь доступний період 26 місяців, починаючи з 30 листопада 2017 року.

Для оцінки ефективності прогнозування можуть бути визначені і використані різні індикації помилок, що описані у підрозділі 1.4.

Помилка прогнозування  $\varepsilon = Y_1 - \hat{Y}_1$  в даному випадку має наступні властивості (рис 2.20):

Математичне сподівання = 0

Середньоквадратичне відхилення = 125

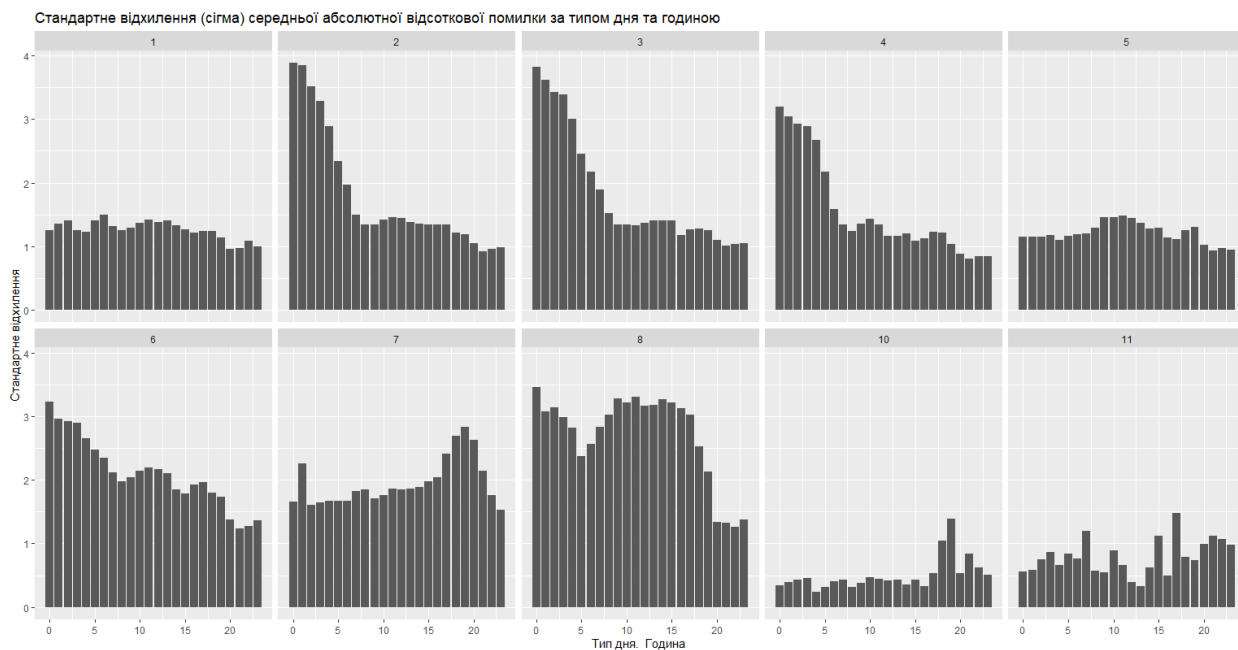


Рис. 2.20. Середньоквадратичне відхилення похибки моделі по кожному класу

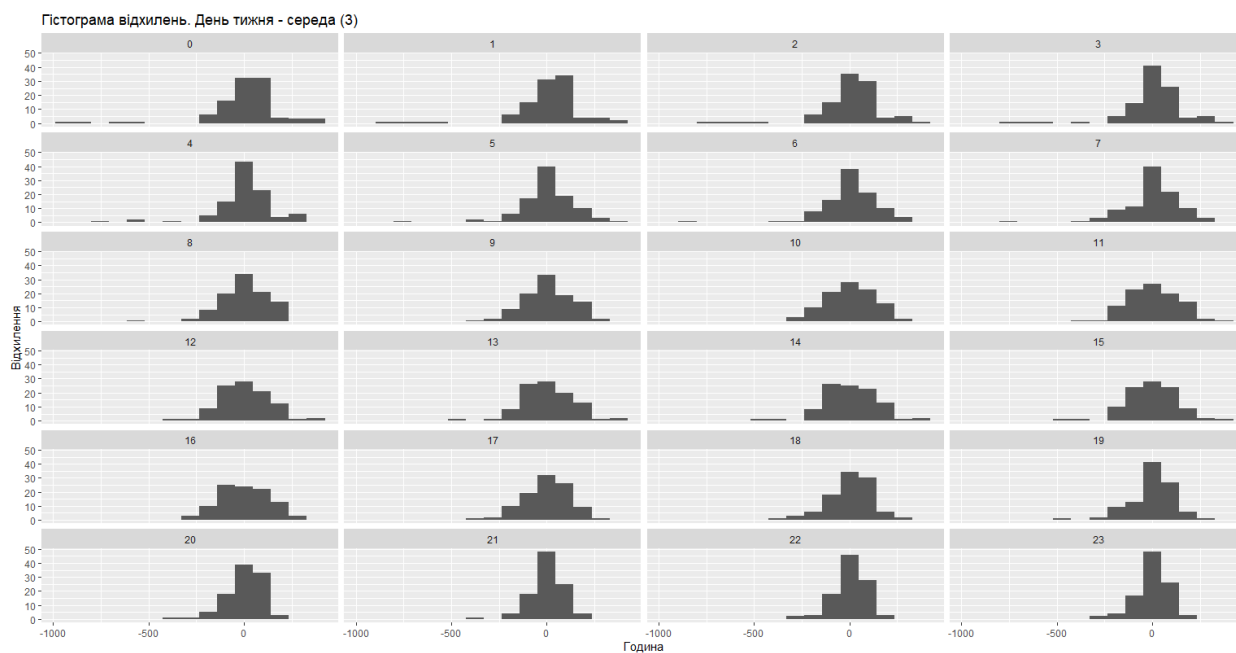


Рис. 2.21. Розподіл загальних відхилень по годинам (день тижня- середа)

Як видно з рис 2.21 похибка прогнозування є нормально розподіленою та з рис 2.18 можна сказати, що відхилення від фактичних значень являють собою білий шум.

Для порівняння зобразимо такі ж характеристики для похибок прогнозування існуючої системи МАВІР (рис 2.22, 2.23).

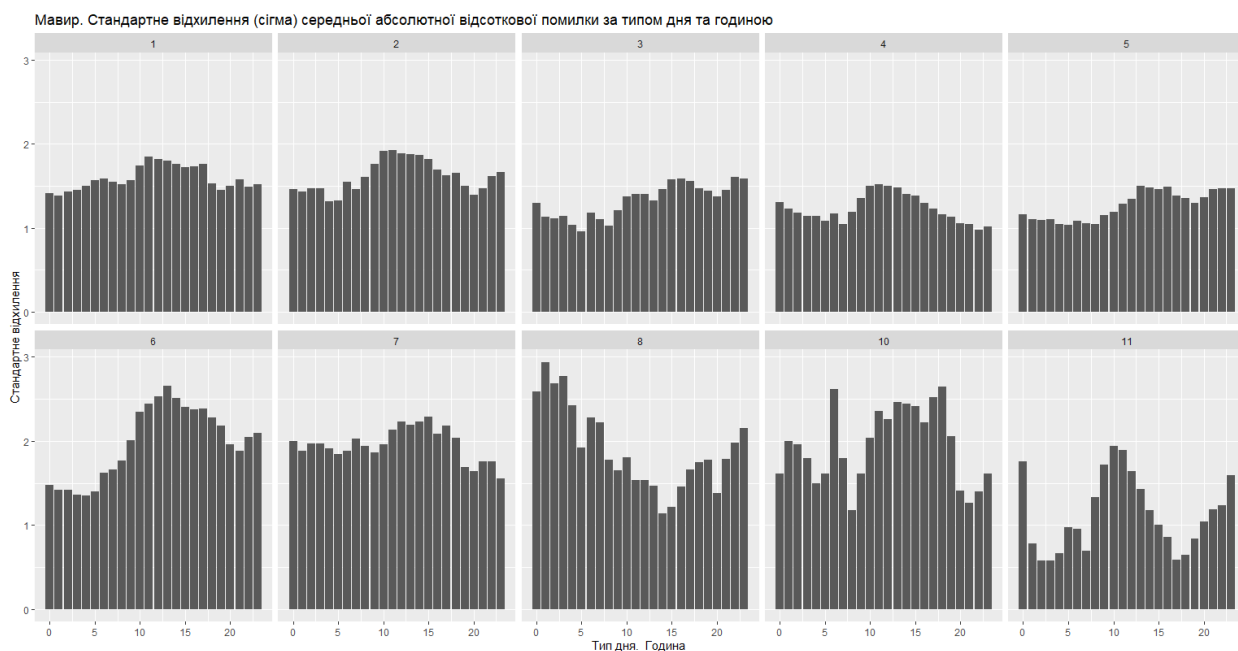


Рис. 2.22. Середньоквадратичне відхилення похибки МАВІР

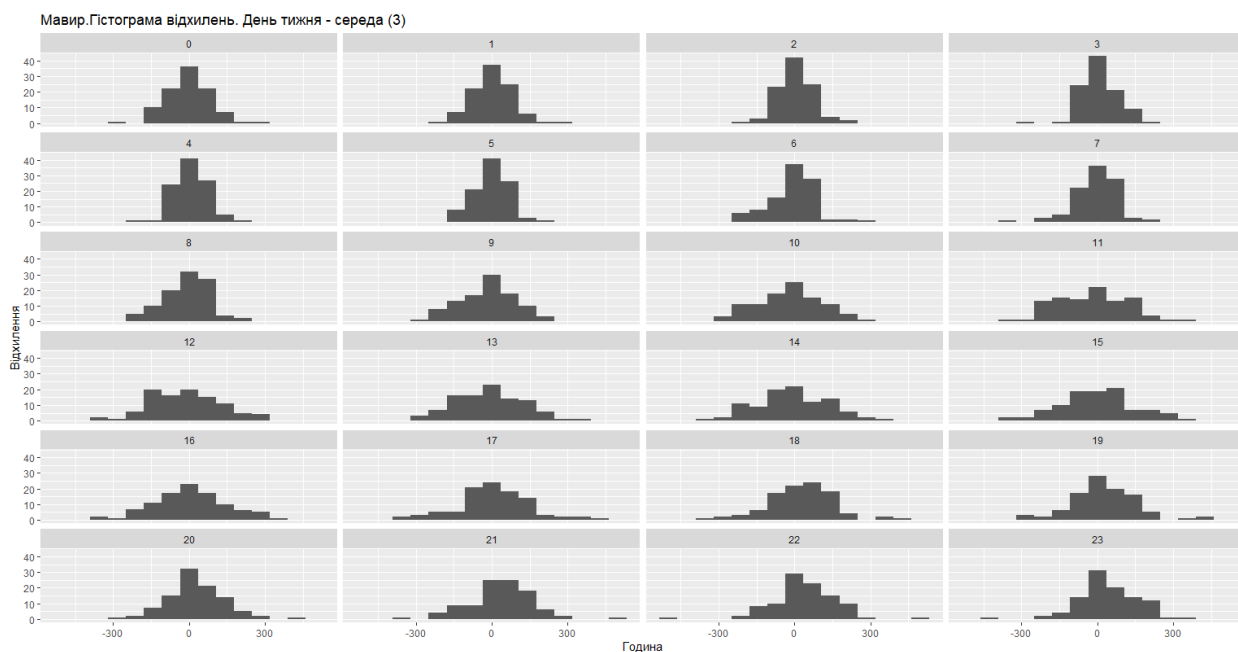


Рис. 2.23. Розподіл відхилень прогнозу МАВІР по годинам (день тижня- середа)



Порівняння загальної точності прогнозу з існуючою системою МАВІР :

У порівняльній таблиці 2.7 надані оцінки помилки прогнозування споживання електроенергії на добу наперед. Перші три моделі відповідають трьом крокам алгоритму. Четверта – результати прогнозування МАВІР.

Таблиця 2.7

Порівняльна таблиця точності прогнозування споживання електроенергії

| № моделі | Назва моделі                    | ME     | RMSE    | MAE     | MPE    | MAPE         |
|----------|---------------------------------|--------|---------|---------|--------|--------------|
| 1        | F_temper                        | 0,702  | 162,088 | 121,005 | -0,097 | 2,368        |
| 2        | F_temper+F_sun                  | 0,000  | 158,419 | 116,910 | -0,095 | 2,294        |
| <b>3</b> | <b>F_temper+F_sun+F_previos</b> | 0,000  | 125,381 | 88,181  | -0,058 | <b>1,744</b> |
| 4        | MAVIR                           | -0,268 | 127,270 | 93,802  | -0,047 | 1,816        |

На основі наведених оцінок можна зробити висновок, що описана вище модель (2.10) в загальному працює краще ніж існуюча МАВІР.

Для візуального розуміння нижче наведені емпіричні функції розподілу похибок (рис 2.24).

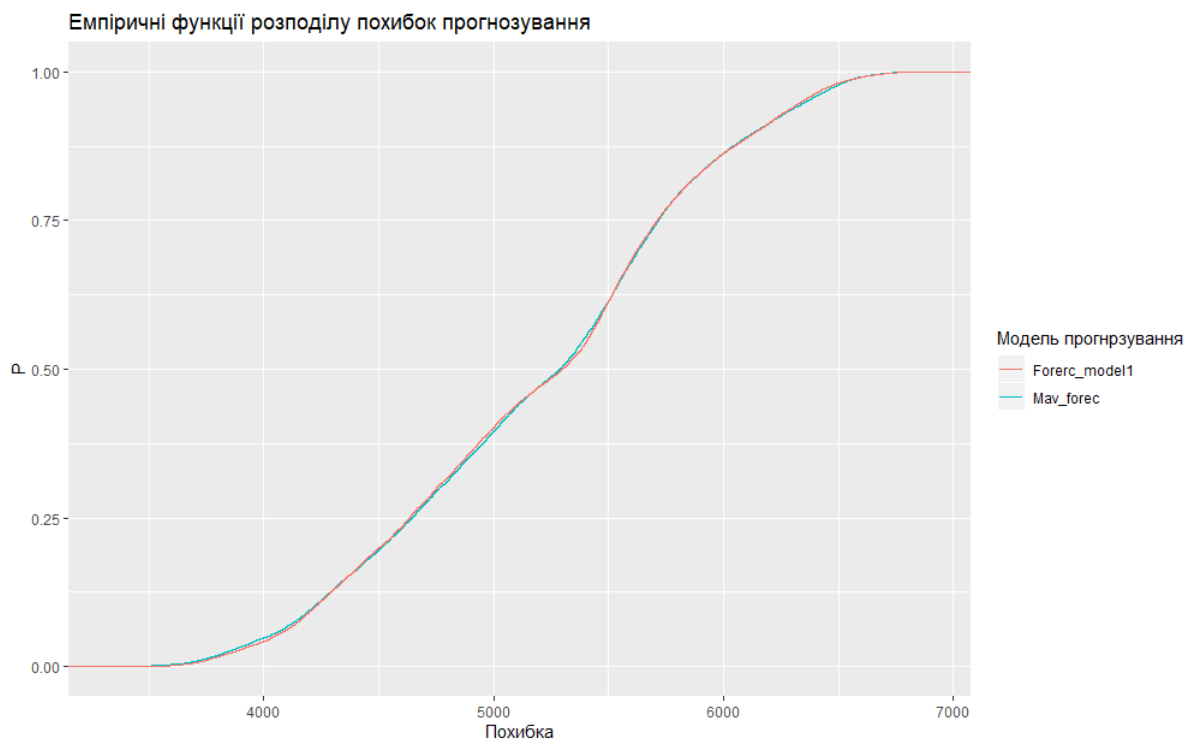


Рис. 2.24. Емпіричні функції розподілу похибок прогнозування

## 2.3. Модель прогнозування ціни на електроенергію на добу наперед

### 2.3.1. Первинний аналіз даних

Для більш адекватного розуміння моделі прогнозування ціни на добу наперед проводимо попередній аналіз даних.

Нижче наведено графік споживання електричної енергії в Угорщині в різних масштабах (рис 2.25-2.26).

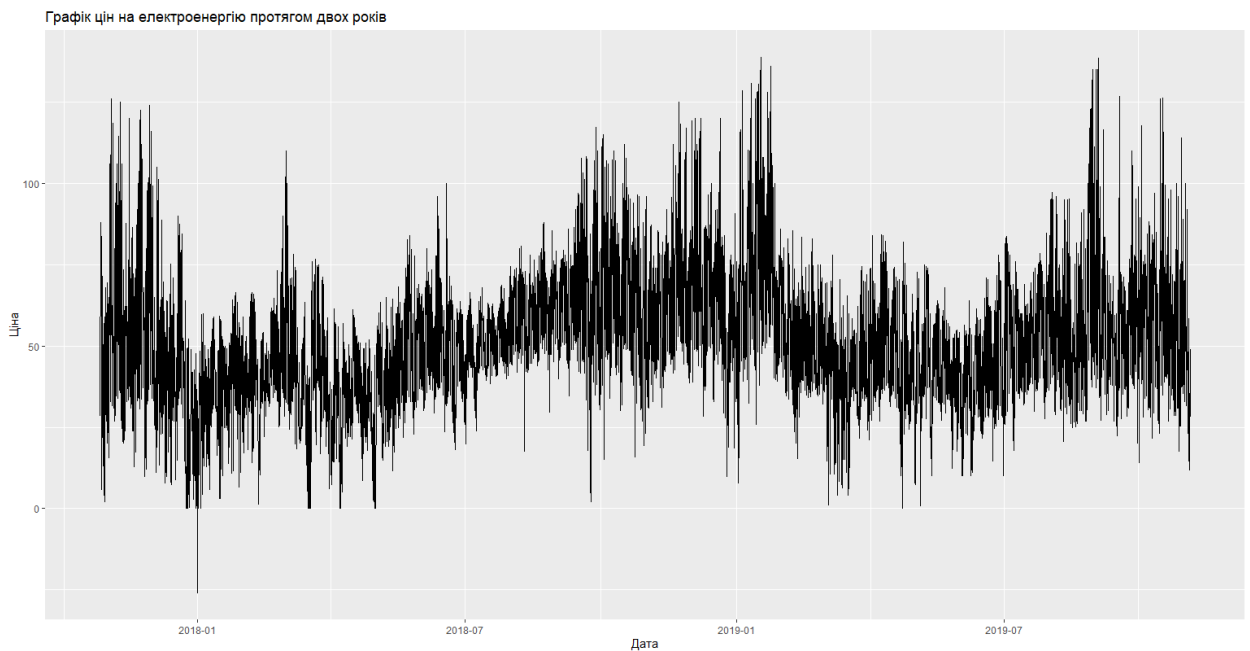


Рис. 2.25. Часовий ряд: ціна на електроенергію за два останні роки

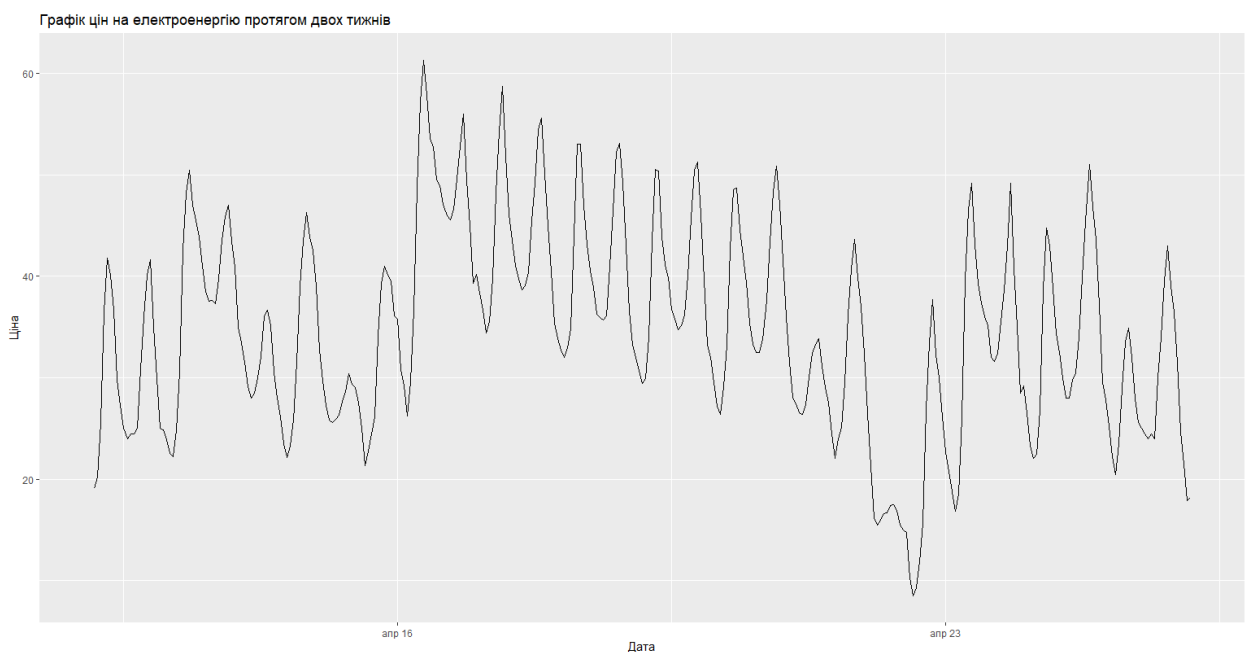


Рис. 2.26. Часовий ряд: ціна на електроенергію за два тижні

На рисунку 2.26 спостерігається, що волатильність ціни (або зміна ціни від одного дня до наступного) досить висока.

Виходячи з аналізу рисунку 2.25, можна припустити, що дані є неоднорідними. Присутні сезонні коливання, що свідчить про різну природу даних. Тому доцільно розглянути розбиття даних.

В січні спостерігається помітне зменшення ціни. Дивлячись на дати, можна припустити, що цей період відповідає новорічним святкам, що є причиною зменшення попиту. Також можна спостерігати тенденцію до зростання з жовтня по лютий, і цей ефект може бути пов'язаний з більш холодною погодою, що призводить до збільшення споживання через потребу в опаленні, що в свою чергу приводить до збільшення ціни. Щорічний сезонний ефект можна побачити більш чітко, глянувши на місячні значення наведені на рис 2.27.

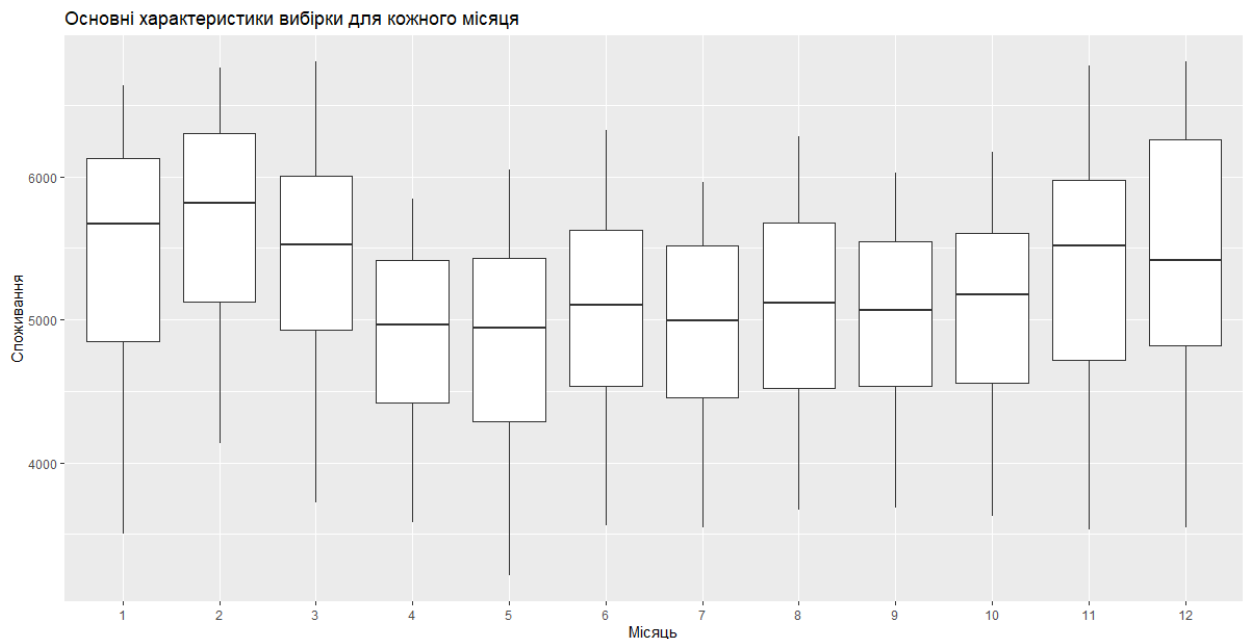


Рис. 2.27. Ціна на електроенергії в місячному розбитті

Якщо подивитися на рис 2.28, на якому показані середньодобові значення та стандартні відхилення для кожного типу тижня, то видно зниження ціни по суботах, неділях та у святкові дні, можливо, викликаного зменшенням промислового споживання.

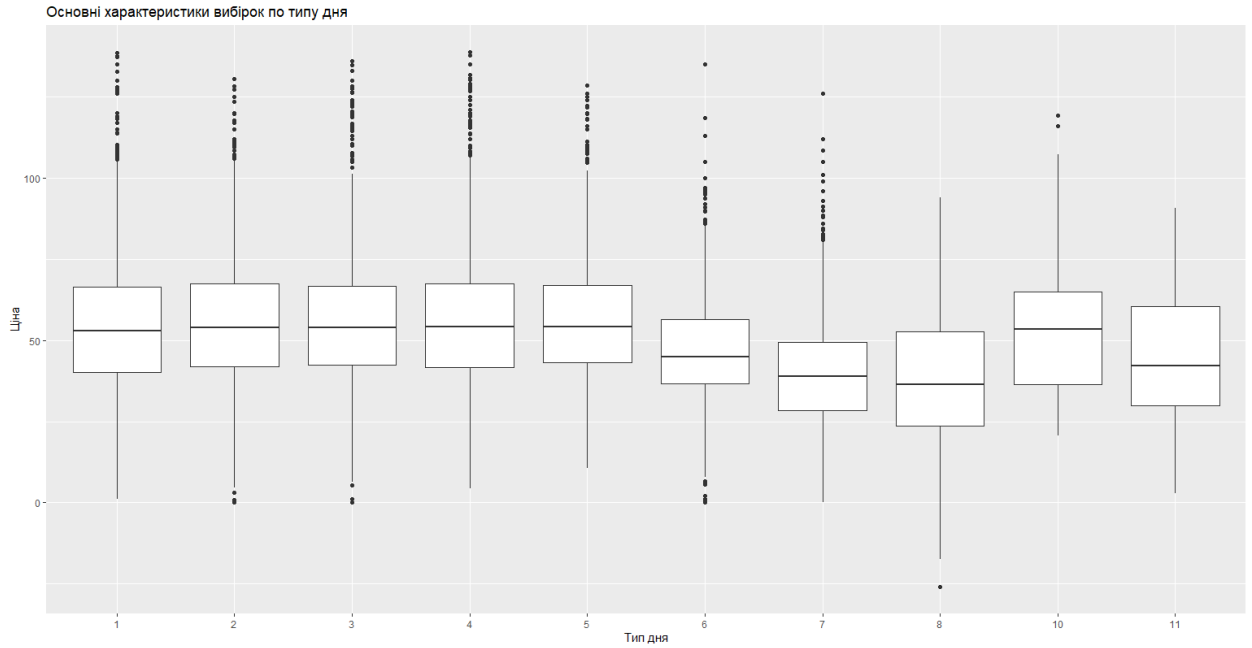


Рис. 2.27. Ціна на електроенергії в залежності від типу дня

Щоденна сезонність відображена на рис 2.28.

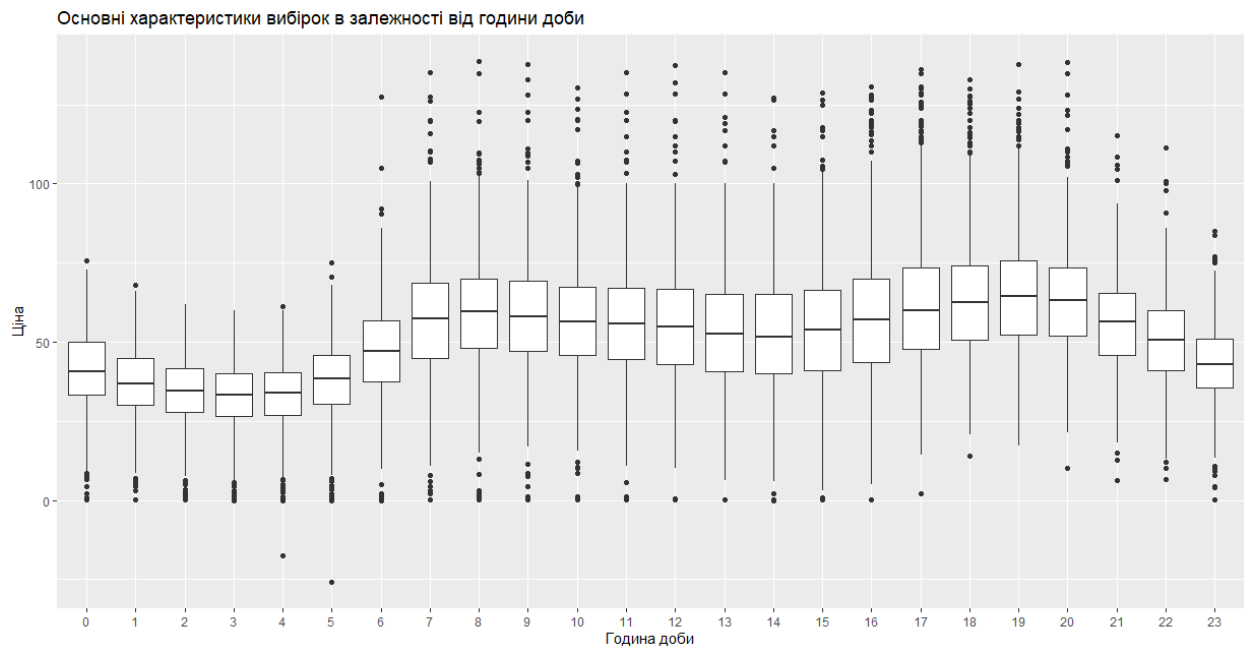


Рис. 2.28. Ціна на електроенергії в залежності від години доби

Отже, ЧР ціни має декілька сезонних складових, які мають бути враховані в моделі.

Для подальшого аналізу даних і підтвердження щоденної і щотижневої сезонності, що наведена вище, можуть бути перевірені функції автокореляції (рис 2.29)

Автокореляційна функція часового ряду Ціна

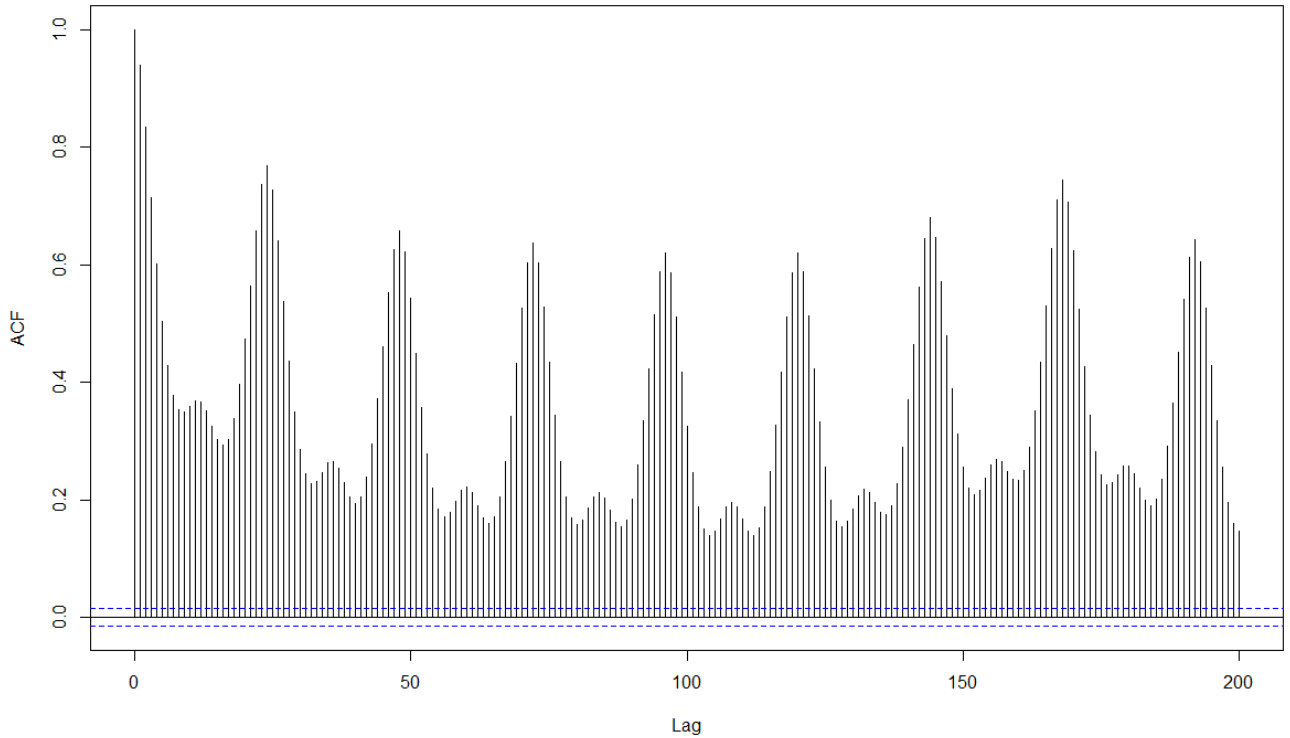


Рис. 2.29. Автокореляційна функція ціни на електроенергії

Дивлячись на рис 2.29 можна побачити лаг 24 і 168, що вказує на щоденну та тижневу сезонності, що підтверджує попередні припущення.

### 2.3.2. Побудова загальної моделі прогнозування ціни

Проаналізувавши графіки попереднього розділу пропустимо, що модель прогнозування ціни електричну енергії на день наперед є адитивною, виду:

$$Y_2(t) = \sum_{i=0}^m F(X_i(t)) + \varepsilon(t), \quad (2.11)$$

де  $F(X_i(t))$  – функція залежності споживання від зовнішнього фактору  $X_i(t)$   
 $\varepsilon(t)$  – похибка.

Припустимо, що  $\mathcal{E} = \{\varepsilon(t); t \in [1:T]\}$  має нормальний розподіл.

Подальша робота з ідентифікації моделі буде проходити наступним чином: знаходження фактору, що найбільше впливає на цільовий показник; побудова

регресії; вилучення з ЧР отриманої складової; оцінка похибки. Дані етапи застосовуються до кожного значимого зовнішнього фактору.

На рис 2.30 видно, що динаміка зміни ціни схожа з динамікою споживання.

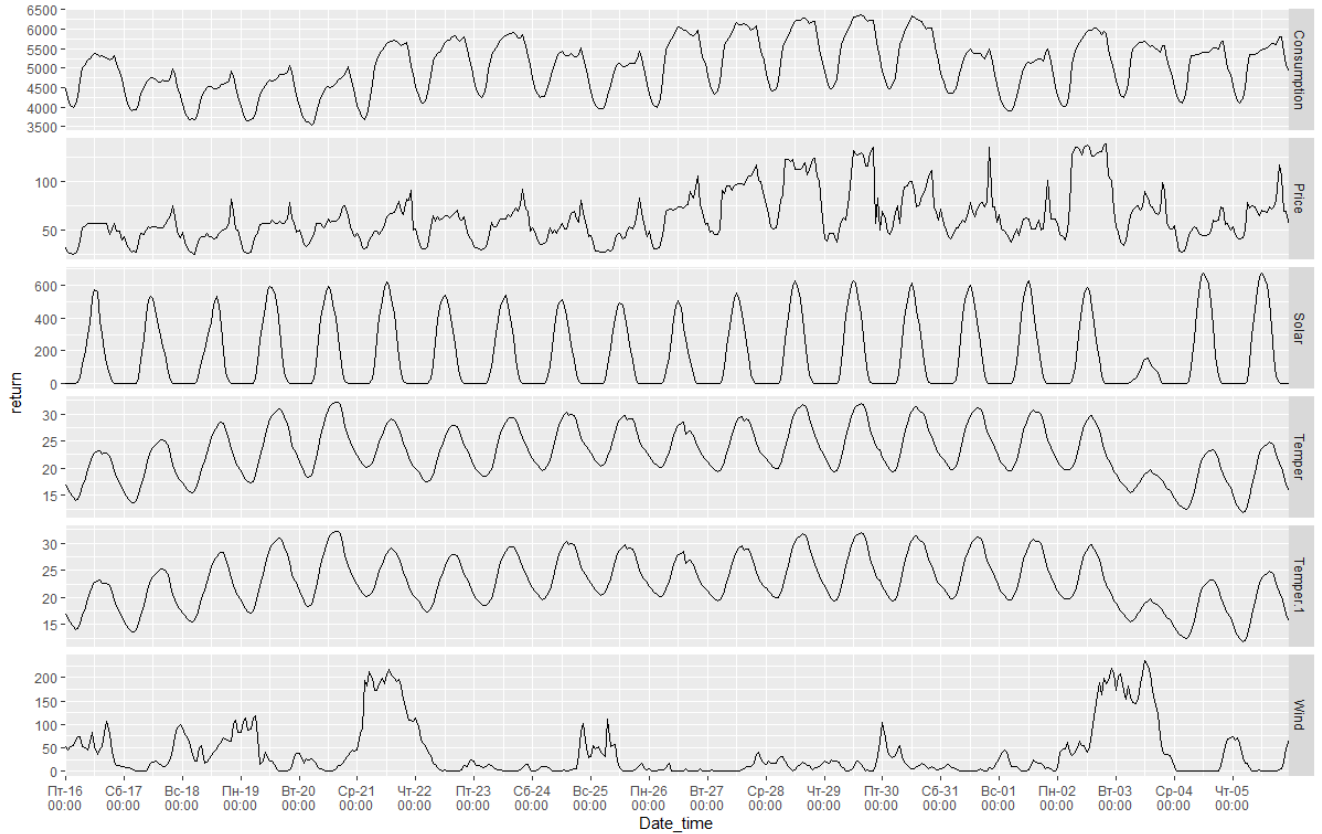


Рис. 2.30. Часові ряди факторів та цільового показника – ціни

Виходячи з коефіцієнтів кореляції, наведених у таблиці 2.8, в якості найбільш значущих чинників, що впливають на цін в моделі були залишені типи дня  $W(t)$ , година доби  $H(t)$ , споживання  $Y_1(t)$  та сонячність  $S(t)$ .

Таблиця 2.8

Коефіцієнти кореляції ціни та факторів

| Фактор      | Ціна  |
|-------------|-------|
| Споживання  | 0,65  |
| Година доби | 0,34  |
| День тижня  | 0,25  |
| Температура | 0,21  |
| Сонячність  | 0,2   |
| Вітер       | 0,098 |

По аналогії з ідентифікацією моделі прогнозування споживання будується модель прогнозування ціни.

З аналізу коефіцієнтів кореляції та графіків визначаємо, що модель складається з таких компонентів:

$$\sum_{i=0}^m F(X_i(t)) = F(W(t)) + F(H(t)) + F_{cons}(T(t)) + F_{sun}(S(t)) + F_{prev}(t), \quad (2.12)$$

де  $F_{prev}(t) = F(Y_2(t - 24))$  – функція впливу змін попередньої доби на поточну, де 24 – лаг, знайдений вище автокореляційним аналізом

### 2.3.3. Формалізація моделі та уточнення параметрів

Ринок електричної енергії базується на попиті – споживання та пропозиції – генерації. Тому одним з найголовніших чинників, що впливають на утворення ціни – є споживання. Дана залежність зображена на рис 2.31.

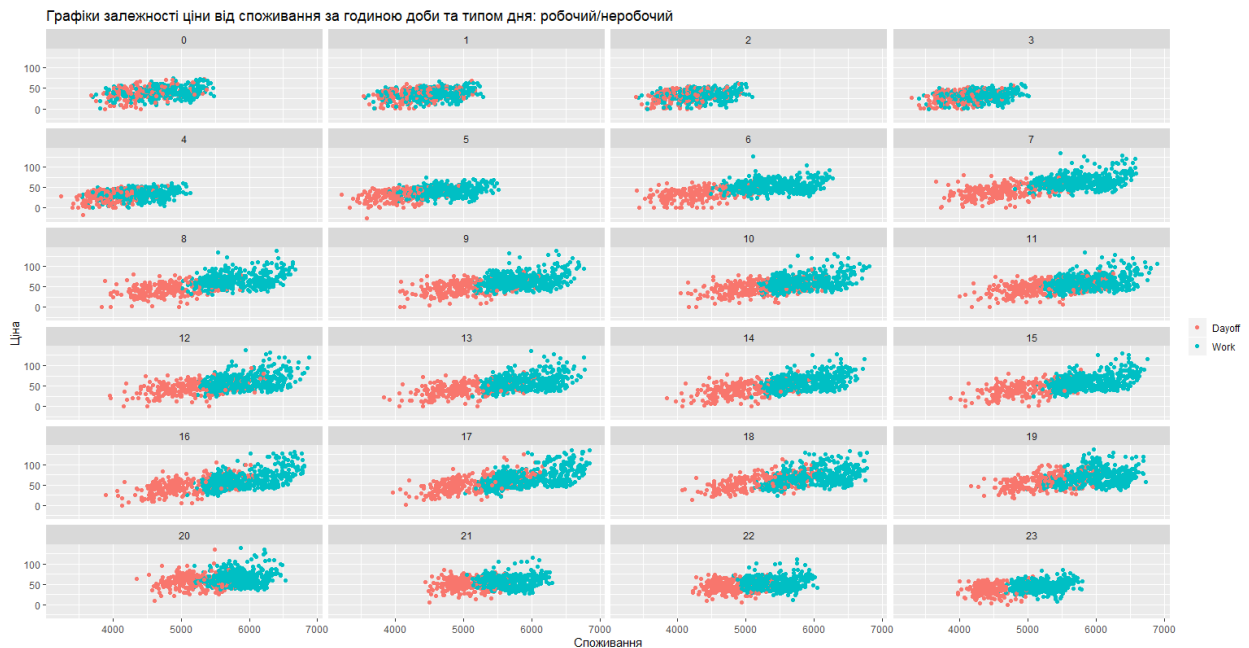


Рис. 2.31. Залежність ціни від споживання

Залежність ціни від споживання має лінійний вид:

$$F_{cons}(S(t)) = a_{cons}(t) + b_{cons}(t)Y_1(t), \quad (2.13)$$

де  $Y_1(t)$  – споживання у момент  $t$ , а коефіцієнти  $a_{cons}, b_{cons}$  залежать від години доби  $H$  та типу дня  $W$  та наведені у таблиці 2.9.

Коефіцієнти лінійної моделі (2.13) ціни від споживання

| Тип дня | Година доби | a       | b      |
|---------|-------------|---------|--------|
| 1       | 0           | 42,602  | -0,002 |
| 1       | 1           | 29,533  | 0,001  |
| 1       | 2           | 23,722  | 0,001  |
| 1       | 3           | 22,342  | 0,002  |
| 1       | 4           | 25,713  | 0,001  |
| 1       | 5           | 19,929  | 0,004  |
| 1       | 6           | 12,554  | 0,008  |
| 1       | 7           | 18,896  | 0,008  |
| 1       | 8           | 23,908  | 0,008  |
| 1       | 9           | 6,575   | 0,010  |
| 1       | 10          | -5,595  | 0,012  |
| 1       | 11          | -12,315 | 0,013  |
| 1       | 12          | -19,988 | 0,014  |
| 1       | 13          | -27,507 | 0,015  |
| 1       | 14          | -27,427 | 0,015  |
| 1       | 15          | -34,230 | 0,016  |
| 1       | 16          | -70,547 | 0,023  |

Значний вплив на ціну дає рівень сонячності (рис 2.32).

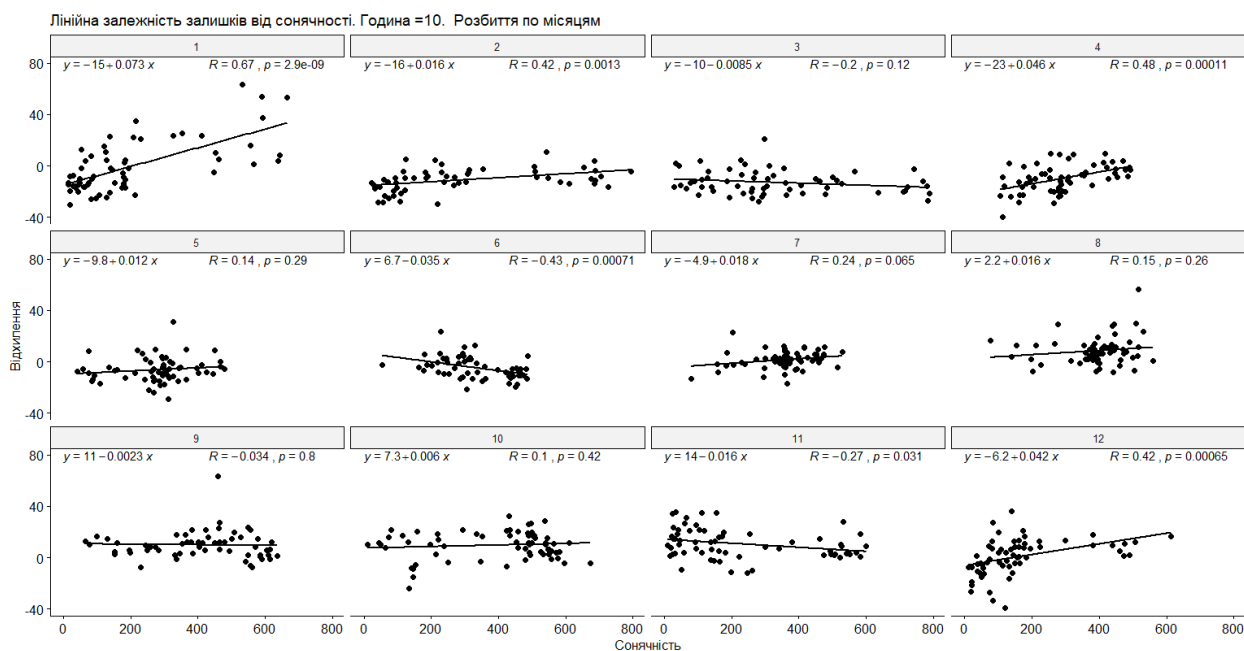


Рис. 2.32. Залежність ціни від сонячності



Залежність ціни від сонячності має лінійний вид:

$$F_{sun}(S(t)) = a_{sun}(t) + b_{sun}(t)S(t), \quad (2.14)$$

де  $S(t)$  – сонячність у момент  $t$ , а коефіцієнти  $a_{sun}, b_{sun}$  залежать від години доби  $H$  та типу дня  $W$ .

Залежність ціни від споживання та сонячності позначимо:

$$\hat{Y}_{2global}(t) = a_{cons}(t) + b_{cons}(t)Y_1(t) + a_{sun}(t) + b_{sun}(t)S(t) \quad (2.15)$$

Розгляд наведеного вище на рисунку 2.33 графіку часткової автокореляції залишків дозволяє зробити висновок, що найбільший вплив на залишки  $\Delta Y_2(t)$  мають значення залишків в момент  $(t - 24)$ .

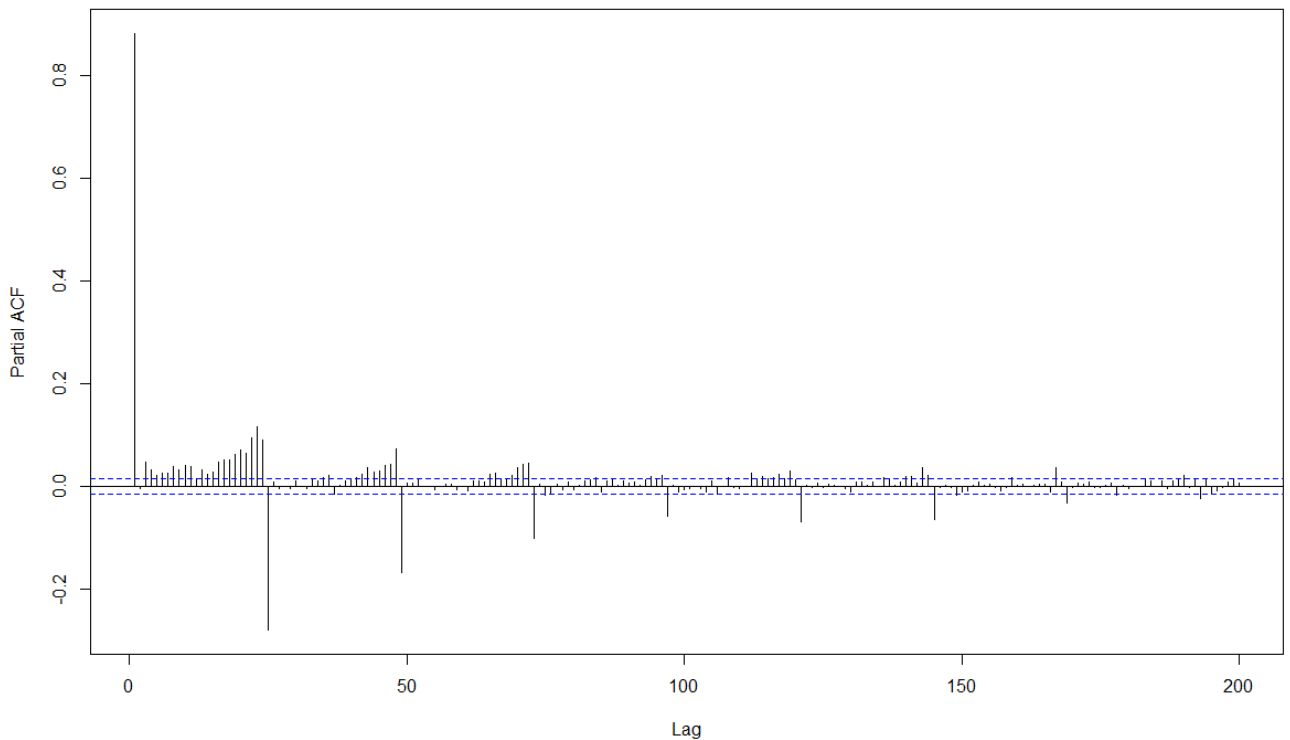


Рис. 2.33. Графік часткової автокореляції ціни

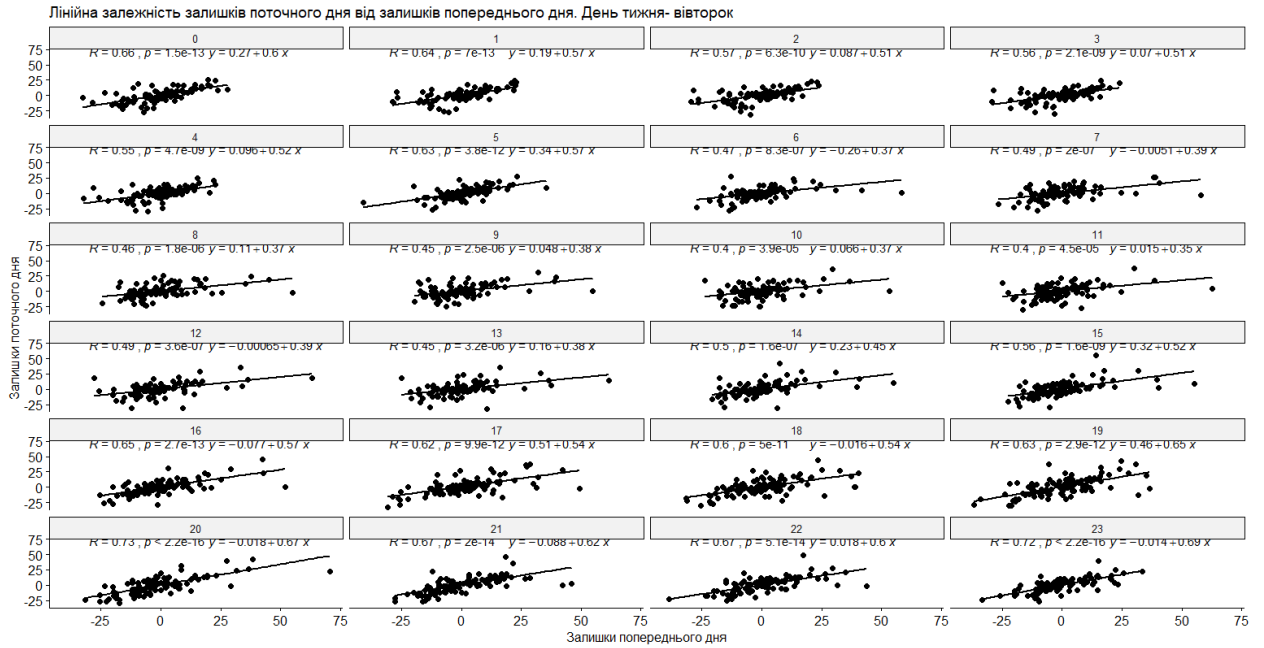


Рис. 2.34. Залежність ціни поточного дня від зміни ціни попереднього дня

Як видно з рис. 2.34 залежність є лінійною, що відповідає наступному рівнянню:

$$F_{prev}(t) = a_{prev}(t) + b_{prev}(t)\Delta Y_2(t - 24), \quad (2.16)$$

де  $\Delta Y_2(t - 24) = Y_2(t - 24) - \widehat{Y}_{2,global}(t - 24)$

де коефіцієнти  $a_{prev}, b_{prev}$  залежать  $H$  та  $W$  і наведені у таблиці 2.10.

Таблиця 2.10

Коефіцієнти лінійної авторегресійної моделі (2.16)

| Тип дня | Година доби | a      | b     |
|---------|-------------|--------|-------|
| 5       | 0           | -0,205 | 0,681 |
| 5       | 1           | -0,068 | 0,675 |
| 5       | 2           | -0,009 | 0,652 |
| 5       | 3           | -0,018 | 0,652 |
| 5       | 4           | 0,131  | 0,654 |
| 5       | 5           | 0,064  | 0,599 |
| 5       | 6           | -0,566 | 0,635 |
| 5       | 7           | -0,353 | 0,463 |
| 5       | 8           | -0,088 | 0,460 |
| 5       | 9           | 0,487  | 0,430 |
| 5       | 10          | 0,067  | 0,366 |
| 5       | 11          | -0,014 | 0,303 |
| 5       | 12          | -0,151 | 0,287 |

Тоді рівняння (2.12) можна представити в наступній формі:

$$\hat{Y}_2(t) = \hat{Y}_{2global}(t) + a_{prev}(t) + b_{prev}(t) \left( Y_2(t - 24) - \hat{Y}_{2global}(t - 24) \right) \quad (2.17)$$

Або, якщо розписати всі доданки, рівняння (2.17) можна записати в розгорнутому вигляді:

$$\begin{aligned} \hat{Y}_2(t) = & a_{cons}(t) + b_{cons}(t)Y_1(t) + a_{sun}(t) + b_{sun}(t)S(t) + \\ & + a_{prev}(t) + b_{prev}(t)[Y_1(t - 24) - \\ & - (a_{cons}(t - 24) + b_{cons}(t - 24)Y_1(t - 24) + \\ & + a_{sun}(t - 24) + b_{sun}(t - 24)S(t - 24))] \end{aligned} \quad (2.18)$$

Коефіцієнти комбінованої моделі (2.18) залежать від параметру  $t$ .

### 2.3.4. Приклад роботи. Аналіз результатів

На рис. 2.35 та рис 2.36 відображено результат прогнозування ціни на добу наперед за розробленою моделлю(2.18) та фактичні значення.

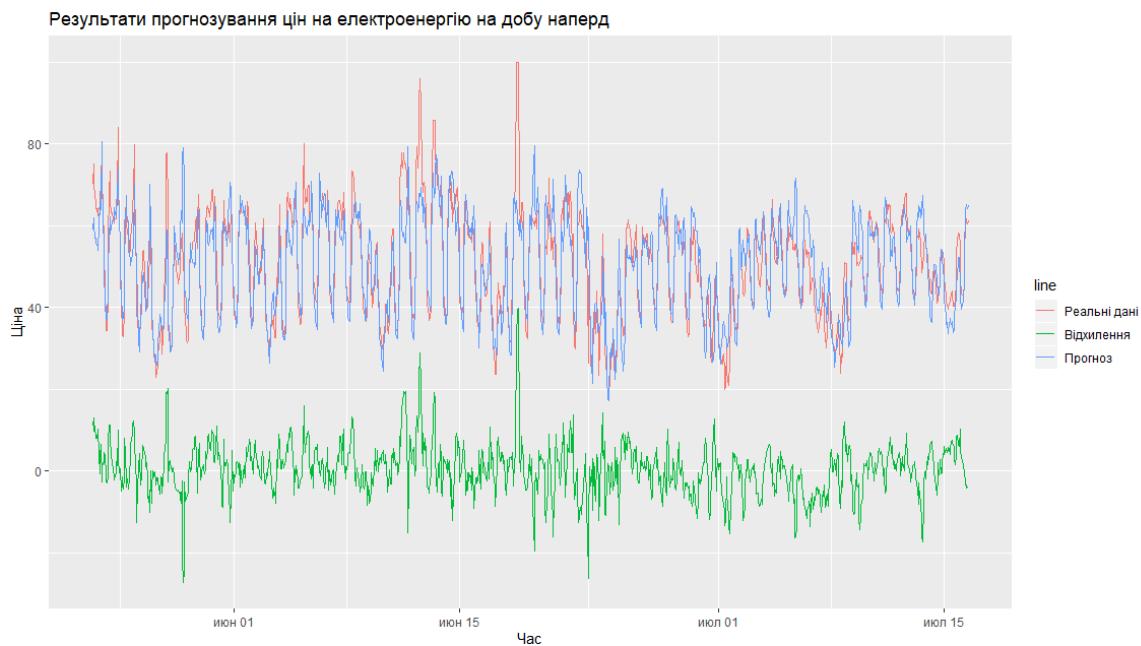


Рис. 2.35. Результат прогнозу ціни на добу наперед

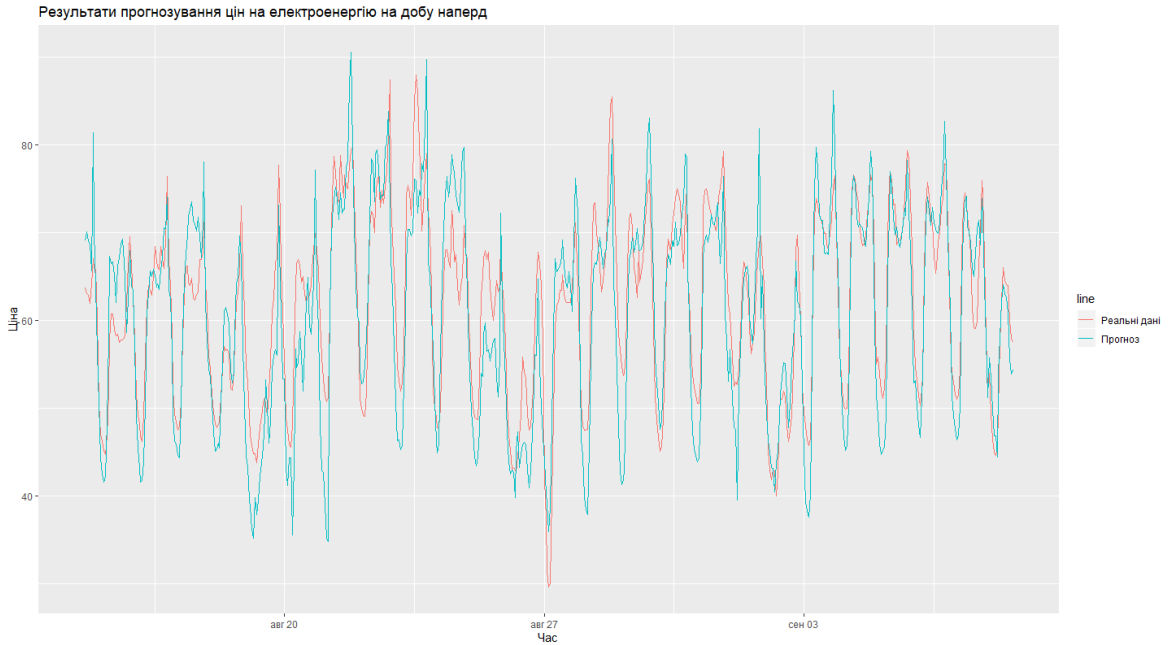


Рис. 2.36. Результат прогнозу ціни на добу наперед

Для оцінки ефективності прогнозування можуть бути визначені і використані різні індикації помилок. Оскільки ціни на електроенергію включають додатні, від’ємні і нульові значення використання відсоткових помилок не дає значущих результатів.

У даному розділі буде вказуватися в двох формах: середньоквадратична помилка (RMSE) і середня абсолютна помилка (MAE). Прогноз, який фокусується на мінімізації MAE, дасть прогнози медіани, тоді як мінімізація RMSE дасть прогнози середнього значення.

Помилка прогнозування  $\varepsilon = Y_2 - \hat{Y}_2$  в даному випадку має наступні властивості:

Математичне сподівання = 0

Середньоквадратичне відхилення = 9

Розподіл помилки прогнозування ціни на добу наперед в залежності від типу дня та години доби наведено на рис 2.37 – 2.38.

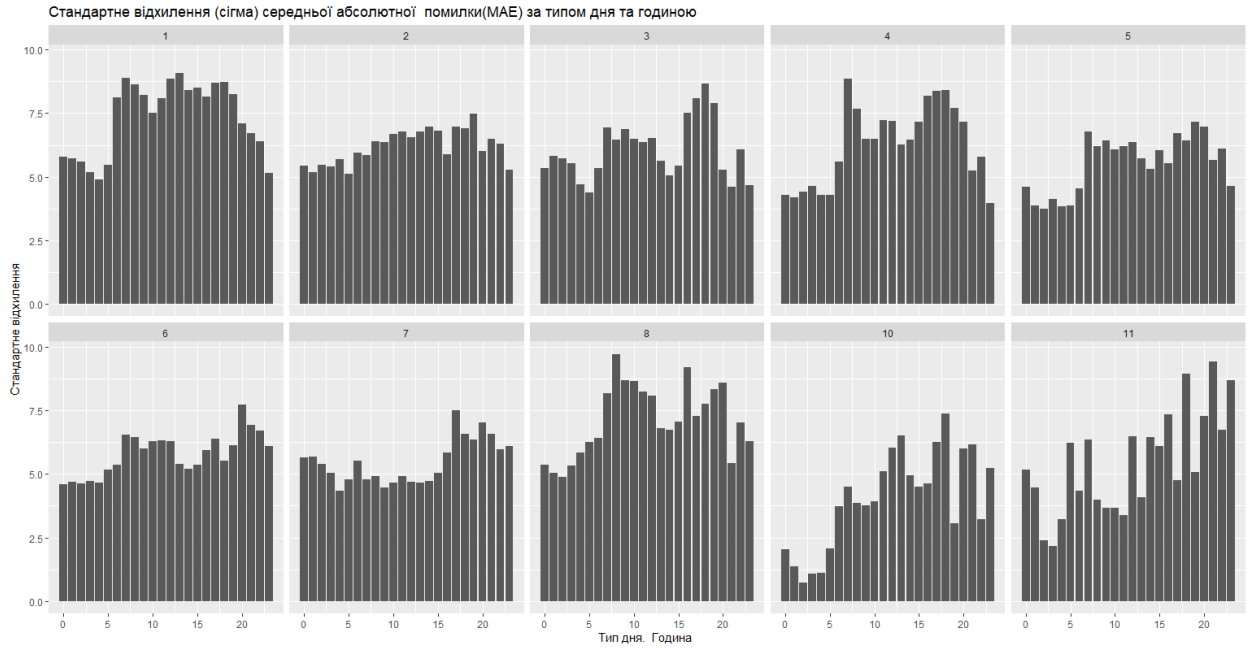


Рис. 2.37. Середньоквадратичне відхилення похибки моделі по кожному класу

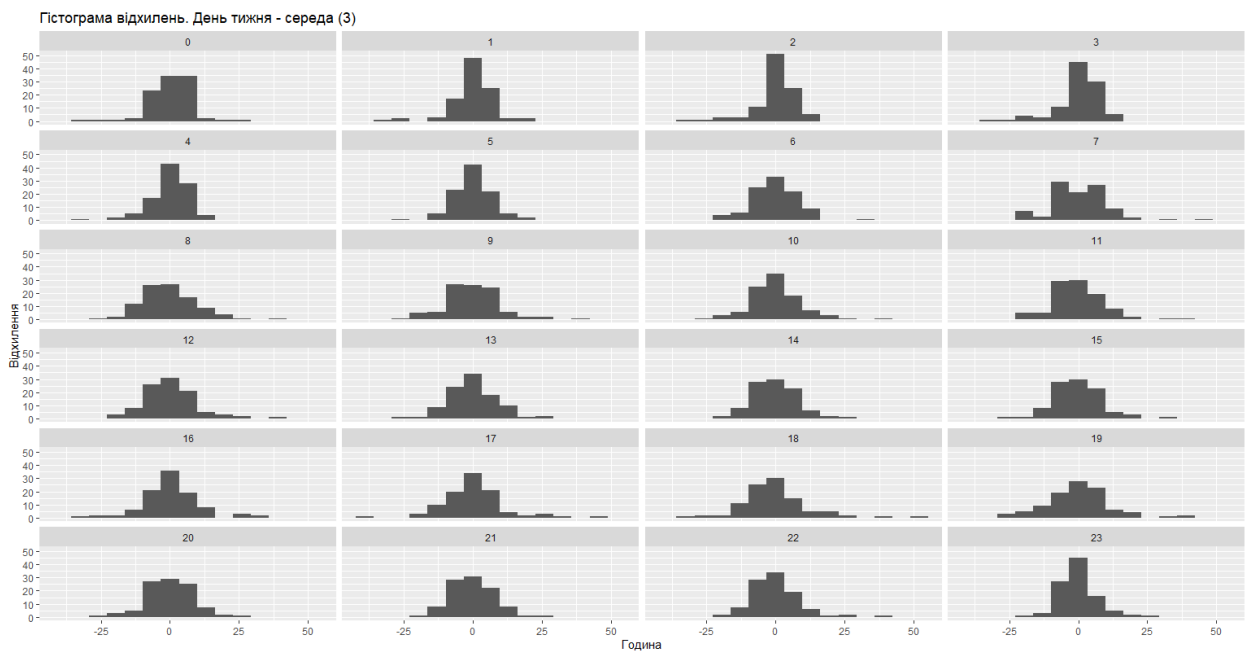


Рис. 2.38. Розподіл загальних відхилень по годинам (день тижня- середа)

Для порівняння точності прогнозування ціни на добу наперед була обрана модель Wattsight, яка враховує нелінійний вплив багатьох факторів. Розподіл похибки прогнозування моделі Wattsight зображена на рис 2.39.

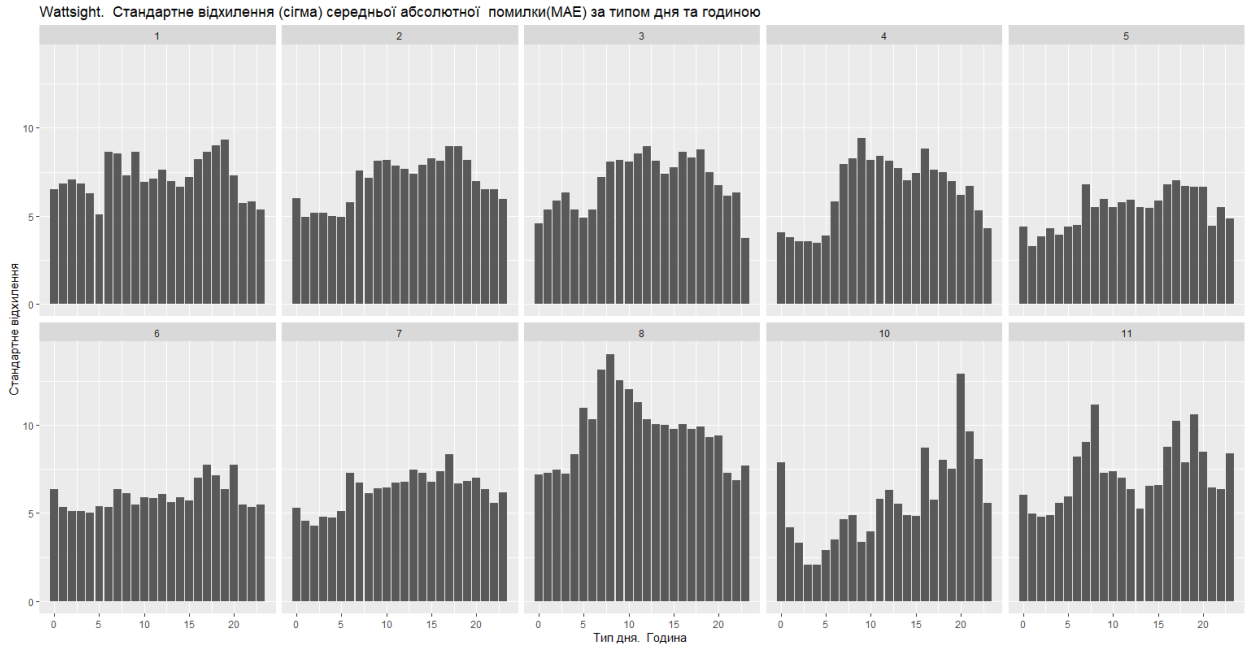


Рис. 2.39. Середньоквадратичне відхилення похибки Wttsight

У порівняльній таблиці 2.11 надані оцінки помилки прогнозування ціни на електроенергію на добу наперед. Перші три моделі відповідають трьом крокам алгоритму. Четверта – результати прогнозування Wattsight.

На основі наведених оцінок можна зробити висновок, що описана вище модель в загальному працює краще ніж існуюча з Wattsight.

Таблиця 2.11

Порівняльна таблиця помилок прогнозування ціни на електроенергію

| Номер | Модель                        | ME     | RMSE   | MAE          |
|-------|-------------------------------|--------|--------|--------------|
| 1     | G_cons                        | -0,223 | 13,316 | 10,287       |
| 2     | G_cons+G_sun                  | -0,010 | 11,255 | 8,421        |
| 3     | <b>G_cons+G_sun+G_previos</b> | 0,000  | 9,154  | <b>6,623</b> |
| 4     | Wattsight                     | -1,077 | 9,782  | 6,879        |

Але необхідно зазначити, що для прогнозування за розробленою моделлю було використано історично актуальні данні множини факторів, оскільки відсутня база прогнозних значень на минулий час. Що в свою чергу дещо покращує результати прогнозу.

В ході роботи було встановлено, що значний вплив на формування та встановлення ціни мають міждержавні комерційні перетоки. Але в моделі їх не було використано, оскільки ці дані не є достовірними на прогнозний проміжок часу.

Також в роботі було здійснено аналіз стійкості, що показав, що при додаванні в модель додаткових незначущих регресорів не відбувається істотної зміни точності прогнозування.

#### 2.4. Алгоритм ідентифікації моделі

Для прогнозування добових погодинних цін та споживання було використано підхід, що дозволяє проводити аналіз часових рядів. З цією метою була розроблена методика, що включає в себе наступні кроки (рис.2.40).

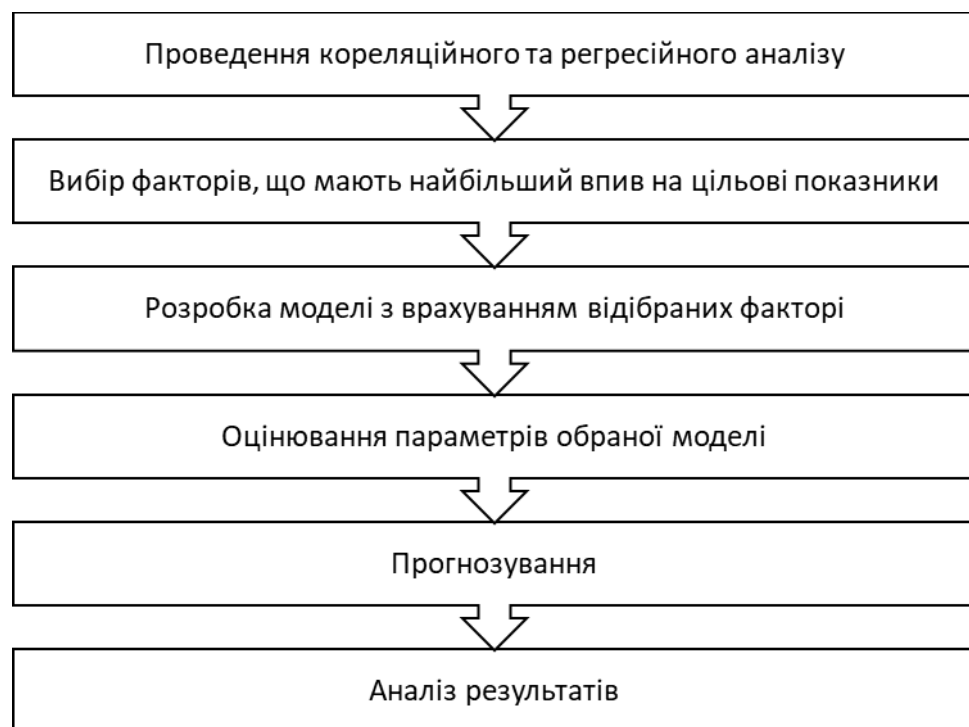


Рис. 2.40.Методика прогнозування цільових показників

## 2.5. Опис мови та середовища програмування

На даний момент налічуються десятки якісних статистичних пакетів, серед яких явними лідерами є SPSS, SAS і MatLab. Однак, в 2019 році, незважаючи на високу конкуренцію, Python та R стали найбільш використовуваними програмними продуктами для статистичного аналізу. Крім того, в останнє десятиліття R активно використовуються в бізнес-секторі: такі компанії-гіганти, як Google, Facebook, Ford і New York Times застосовують для збору, аналізу та візуалізації даних. Для того щоб зрозуміти причини зростаючої популярності мови R, звернемо увагу на загальні риси та відмінності від інших статистичних продуктів.

R – мова програмування для статистичної обробки даних і роботи з графікою, а також програмне середовище з відкритим вихідним кодом, що розвивається в рамках проекту GNU. Для Mac і Windows є штатний GUI.

У середовищі R реалізовані багато статистичних методів: лінійні і нелінійні моделі, перевірка статистичних гіпотез, аналіз часових рядів, класифікація, кластеризація, графічна візуалізація. Мова R дозволяє визначати власні функції. Багато функцій R написані на самому R. Для обчислювально складних задач є можливість реалізації функцій на C, C++ і Fortran. Графічні функції дозволяють створювати якісні графіки.

Хоча R найчастіше застосовується для статистичних обчислень, він також може використовуватися як засіб матричних обчислень. Як і MATLAB, R трактує результат будь-якої операції з числами як вектор одиничної довжини.

В R є можливість створення скриптів (програм), які потім завантажуються в R і інтерпретуються. Ще одна важлива перевага R – наявність численних розширень або пакетів. Є можливість установити будь-який доступний на CRAN пакетів.

R являє собою інтегрований набір програмних засобів для обробки даних, розрахунку та графічного відображення. Це включає:



- ефективний засіб обробки і зберігання даних,
- набір операторів для розрахунків на масивах, зокрема матриць,
- великий, узгоджений, інтегрований набір проміжних інструментів для аналізу даних,
- графічні засоби для аналізу даних і відображення на екрані або в друкованому вигляді,
- добре розроблена, проста і ефективна мова програмування, що включає в себе умовні вирази, цикли, визначені користувачем рекурсивні функції і засоби введення і виведення.

R, як і S, розроблений на основі справжньої комп'ютерної мови і дозволяє користувачам додавати додаткові функції. Велика частина системи написана на діалекті R мови S, що дозволяє користувачам легко слідувати прийнятним алгоритмічним рішенням. Для завдань, що вимагають великого обсягу обчислень, код C, C++ і Fortran може бути пов'язаний з написаним кодом і викликатися під час виконання. Досвідчені користувачі можуть написати C-код для безпосереднього управління R-об'єктами [60].

Функції, які були використані при розробці інформаційної технології:

- lm – визначення параметрів лінійної моделі
- nls – визначення параметрів нелінійної моделі
- loess – непараметричний підхід до визначення параметрів регресії
- predict – прогнозування за вказаною моделлю
- cor – обчислення кореляції
- ggplot, boxplot, ggscatter – побудова графіків
- accuracy – обчислення похибки прогнозування
- decompose – декомпозиція ЧР

## 2.6. Висновки до розділу 2

1. Для розробки моделі прогнозування ціни та споживання електричної енергії було обрано дані угорського ринку за останні три роки. Для прогнозування цільових показників інших країн від розробленої моделі залишається незмінним, необхідне лише перерахування коефіцієнтів.

2. Метою етапу розробки моделі прогнозування було прийняття рішення про формування множини незалежних змінних, що будуть включені в модель. Для розглянутого класу об'єктів було виділено наступні набори найбільш значущих параметрів: 1) споживання електроенергії конкретну годину певного дня тижня; 2) середня температура кількох часових відліків назад від прогнозованої точки; 3) тип дня : робочі, вихідні та святкові; 4) сонячність, що впливає на зміну комфортної температури та освітленості в приміщенні.

3. Під час проведення первинного аналізу ЧР споживання та ціни було виявлено, що дані є неоднорідними. Гістограми частот багатомодальні, що свідчить про різну природу показників. Тому запропоновано розбиття даних по типам дня та години, в результаті чого отримано 240 моделей прогнозування, об'єднання яких і дає початковий часовий ряд.

4. В результаті регресійного та кореляційного аналізу було розроблено адитивну комбіновану модель прогнозування споживання електричної енергії, що враховує кубічну залежність від температури, лінійну – від сонячності. Також врахований вплив відхилень попереднього дня.

5. Прогноз споживання з мінімальним стандартним відхилення середньої абсолютної відсоткової похибки  $<1,5$  вийшов для дня тижня «понеділок» та «п'ятниця», з максимальною  $3,1$  – для святкових днів. Відхилення мають нормальний розподіл з математичним сподіванням  $0$  та схожі на білий шум.

6. Модель прогнозування ціни на добу наперед являє собою комбінацію лінійних залежностей від споживання, сонячності та відхилень попереднього дня, що залежать від типу дня та години доби.

7. Для порівняння точності прогнозування ціни на добу наперед була обрана модель WattSight, яка враховує нелінійний вплив багатьох факторів. В результаті порівняння точності з WattSight, отримано, що MAE розробленої моделі = 6,62, а MAE WattSight = 6,88. Отже, в середньому запропонований метод є більш точним. Але, зауважимо, що за відсутності бази прогнозних значень на минулий час, перевірка відбувалась на актуальних даних, що в свою чергу покращує результат, оскільки відсутня похибка прогнозування зовнішніх факторів (температури, сонячності).

8. Визначено переваги розробленої інформаційної технології:

- швидкість знаходження коефіцієнтів моделі;
- можливість врахування довільної кількості зовнішніх факторів;
- простота та прозорість роботи;
- точність прогнозування.

9. Перспективами подальшої розробки технології може бути:

- аналіз та прогнозування викидів (нетипових значень) ціни;
- врахування в моделі прогнозування ціни таких факторів, як генерація вітру, сонця, доступна робоча потужність теплової та атомної генерації;
- врахування ринкової інтеграції, що може бути реалізовано шляхом одночасного прогнозування ціни на сусідніх ринках електроенергії;

## ВИСНОВКИ

У ході виконання кваліфікаційної роботи розроблено інформаційну технологію аналізу динаміки цін ринку електричної енергії країн Євросоюзу на прикладі даних угорського ринку.

Під час виконання роботи отримані такі результати:

1. Досліджено ринки електричної енергії країн Євросоюзу, розглянуто структуру споживання та генерації різних країн. З'ясовано, що для учасників оптового ринку електроенергії і потужності найбільш актуальні коротко- і середньострокові прогнози ціни на електроенергію на ринку на добу вперед.

2. Проведено аналіз існуючих методів та засобів короткострокового прогнозування часових рядів. Література з прогнозування цін зазвичай поділяється на п'ять областей: теоретичні моделі ігор, фундаментальні методи, моделі з використанням статистичних властивостей та методи штучного інтелекту.

3. Розроблена вперше запропонована інформаційна технологія прогнозування споживання та ціни на електроенергію на добу наперед, що представляє собою комбінацію лінійних та нелінійних залежностей цільових показників та зовнішніх факторів. Ключовим в моделюванні цін на електроенергію є врахування сезонності. Ціна на електроенергію демонструє сезонність на трьох рівнях: щоденний і щотижневий, а в деякій мірі - річний.

4. На основі запропонованої технології розроблено програмне забезпечення на мові програмування R, що дозволяє ефективно обробляти та зберігати дані, проводити складні обчислення та зручно візуалізувати роботу алгоритму.

5. Проведено експериментальні дослідження запропонованих методів на реальних даних. Під час тестування була оцінена похибка прогнозування. В результаті чого було зроблено висновок, що якість прогнозування розробленої інформаційної технології є статистично кращою за існуючі моделі МАВІР (прогнозування споживання) та Wattsight (прогнозування ціни на електроенергію).

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. R. Weron Electricity price forecasting: a review of the state-of-the-art with a look into the future/ *Int J Forecast*, 30 (4) (2014), pp. 1030-1081,
2. T. Jamasb, M. Pollitt Electricity market reform in the European union: review of progress toward liberalization & integration / *Energy J*, 26 (2005), pp. 11-41.
3. Антикризисные стратегии развития региональной энергетики : моногр. / Е. В. Оборина, Д. В. Волошин, С. Г. Ажнакин, К.Э. Шурда ; НАН Украины. Ин-т проблем рынка и экономико-эколог. исслед. - О.: Фенікс, 2010. - 283 с. - ISBN 978-966-438-170-0.
4. Дев'яткіна, С. С. Альтернативні джерела енергії : навч. посіб. / С. С. Дев'яткіна, Т. Ю. Шкварницька ; Нац. авіац. ун-т. - К., 2006. - 92 с. - Бібліогр.: с. 83.
5. Про ринок електричної енергії: Закон України від 13 квітня 2017 р. № 2019-VIII //Відомості Верховної Ради (ВВР), 2017, № 27-28, ст.312.
6. Укренерго [Електронний ресурс] – Режим доступу до ресурсу: <https://ua.energy/novuj-rynok-e-e/implementatsiya/>
7. R. Weron Electricity price forecasting: a review of the state-of-the-art with a look into the future / *Int J Forecast*, 30 (4) (2014), pp. 1030-1081.
8. R. Weron, A. Misiorek Forecasting spot electricity prices: a comparison of parametric and semiparametric time series models / *Int J Forecast*, 24 (4) (2008), pp. 744-763.
9. J. Crespo Cuaresma, J. Hlouskova, S. Kossmeier, M. Obersteiner Forecasting electricity spot-prices using linear univariate time-series models / *Appl Energy*, 77 (1) (2004), pp. 87-106.
10. Z. Yang, L. Ce, L. Lian Electricity price forecasting by a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods / *Appl Energy*, 190 (2017), pp. 291-305.

11. F.J. Nogales, J. Contreras, A.J. Conejo, R. Espínola Forecasting next-day electricity prices by time series models / IEEE Trans Power Syst, 17 (2) (2002), pp. 342-348.
12. A. Cruz, A. Muñoz, J. Zamora, R. Espínola The effect of wind generation and weekday on Spanish electricity spot price forecasting / Electr Power Syst Res, 81 (10) (2011), pp. 1924-1935.
13. A. Misiolek, S. Trueck, R. Weron Point and interval forecasting of spot electricity prices: linear vs. non-linear time series models / Stud Nonlinear Dyn Econometr, 10 (3) (2006), pp. 1-36
14. A.K. Diongue, D. Guégan, B. Vignal Forecasting electricity spot market prices with a k-factor GIGARCH process / Appl Energy, 86 (4) (2009), pp. 505-510.
15. A. Conejo, M. Plazas, R. Espinola, A. Molina Day-ahead electricity price forecasting using the wavelet transform and ARIMA models / IEEE Trans Power Syst, 20 (2) (2005), pp. 1035-1042.
16. Щелкалин В. Н. Трендовый подход прогнозирования временных рядов на основе метода «Гусеница»–SSA // Материалы 14-й Международной научно-технической конференции SAIT 2012. Киев, 24 апреля 2012 г. Киев: УНК «ИПСА» НТУУ «КПИ», 2012. С. 258-259.
17. Vahabie A.H., Yousefi M.M.R., Araabi B.N., Lucas C., Barghinia S. Combination of Singular Spectrum Analysis and Autoregressive Model for short term load forecasting // IEEE LAUSANNE POWERTECH. 2007. P. 1090-1093.
18. Щелкалин В.Н. Декомпозиционный подход прогнозирования временных рядов на основе метода «Гусеница»–SSA // Материалы 14-й Международной научно-технической конференции SAIT 2012. Киев, 24 апреля 2012 г. К.: УНК «ИПСА» НТУУ «КПИ», 2012. С. 260-261.
19. Голяндина Н. Э. Метод «Гусеница»–SSA: прогноз временных рядов. СПб. : С.-Петербургский государственный университет, 2004. 52 с.
20. Степанов Д.В., Голяндина Н.Э. Варианты метода «Гусеница»–SSA для прогноза многомерных временных рядов // Труды IV Международной

конференции «Идентификация систем и задачи управления» SICPRO'05. Москва, 25-28 января 2005. М.: Институт проблем управления им. В.А. Трапезникова РАН, 2005. С. 1831-1848.

21. Бокс Дж., Дженкинс Г. Анализ временных рядов. Прогноз и управление. Вып. I / Пер. с англ. М.: Мир, 1974. 406 с.

22. Евдокимов А.Г., Тевяшев А.Д. Об одном алгоритме выявления скрытой периодичности // Автоматизированные системы управления и приборы автоматики. 1977. Вып. 43. С. 60-64.

23. Qiang Zhan, Ben De Wang, Bin He, Yong Peng, Ming Lei Ren. Singular Spectrum Analysis and ARIMA Hybrid Model for Annual Runnoff Forecasting // Water Resource Manag. 2011.

24. Евдокимов А. Г., Тевяшев А. Д. Оперативное управление потокораспределением в инженерных сетях. Харьков: Вища школа, 1980. 144 с.

25. Lawrance A.J., Kottegoda N. T. Stochasting modelling of riverflow time series // J. R. Stat. Soc. A. 1977. Vol. 140, No. 1. P. 1-47.

26. Fernando D.A.K., Jayawardena W.A. Generation and forecasting of monsoon rainfall data // In: Proc. of the 20th WEDC conference. Colombo, Sri Lanka, 1994.

27. Yurekli K., Kurunca A., Ozturkb F. Application of linear stochastic models to monthly flow data of Kelkit Stream // Ecol. Model. 2005. Vol. 183, No. 1. P. 67-75.

28. Broomhead D. S., King G. P. Extracting qualitative dynamics from experimental data // Physica D. 1986. Vol. 20. P. 217-236.

29. Fraedrich K. Estimating the dimension of weather and climate attractor // J. Atmos. Sci. 1986. Vol. 43. P. 419-432.

30. Vautard R., Ghil M. Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series // Physica D. 1989. Vol. 35. P. 395 424.

31. Бэнн Д. В., Фармер Е. Д. Сравнительные модели прогнозирования электрической нагрузки: пер. с англ. М.: Энергоатомиздат, 1987. 200 с.

32. Седов А.В. Моделирование объектов с дискретно-распределенными параметрами: декомпозиционный подход. М.: Наука, 2010. 438 с.
33. Broomhead D. S., King G. P. Extracting qualitative dynamics from experimental data // *Physica D*. 1986. Vol. 20. P. 217-236.
34. Fraedrich K. Estimating the dimension of weather and climate attractor // *J. Atmos. Sci.* 1986. Vol. 43. P. 419-432.
35. Vautard R., Ghil M. Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series // *Physica D*. 1989. Vol. 35. P. 395-424.
36. Ghil M., Vautard R. Interdecadal oscillations and the warming trend in global temperature time series // *Nature*. 1991. Vol. 350. P. 324-327.
37. Yiou P., Baert E., Loutre M.F. Spectral analysis of climate data // *Surv. Geophys.* 1996. Vol. 17. P. 619-663.
38. Orhan Altu ğ Karabiber, George Xydis. Electricity Price Forecasting in the Danish Day-Ahead Market Using the TBATS, ANN and ARIMA Methods // *Energies* 2019, 12(5), 928;
39. JesusLago, FjoDe Ridder. Forecasting day-ahead electricity prices in Europe: The importance of considering market integration // *Applied Energy*, 1 February 2018, Pages 890-903.
40. Unger, E.A.; Ulfarsson, G.F.; Gardarsson, S.M.; Matthiasson, T. The effect of wind energy production on cross-border electricity pricing: The case of western Denmark in the Nord Pool market. *Econ. Anal. Policy* 2018, 58, 121–130.
41. Таран А.С. Учет особенностей работы ТЭЦ при управлении режимами на оптовом рынке электроэнергии // Труды всероссийской научно-технической конференции «Электроэнергетика глазами молодежи». – Екатеринбург: УрФУ. – 2010. – Т.1.
42. HUPX Ltd. [Электронный ресурс] – Режим доступа до ресурсу: <https://hupx.hu/en/about-us/company-info> Дата звернення :2 січня 2020.



43. Портал об электроразрядной инфраструктуре, электромобилях и объектах электроэнергетики. [Электронный ресурс] – Режим доступа до ресурсу: [http://proenergo.blogspot.com/p/blog-page\\_5.html](http://proenergo.blogspot.com/p/blog-page_5.html)
44. Протокол к энергетической хартии по вопросам энергетической эффективности и соответствующим экологическим аспектам // румыния. Регулярный обзор 2019
45. ДП «НЕК «УКРЕНЕРГО» Огляд стану та перспективи розвитку міждержавних електричних мереж. Київ – 12/2018
46. Ghosh D. "Game Theoretic Bidding Strategies for Auctions in Green Electricity Markets" (2010). Honors Scholar Theses. Paper 149.
47. Bastian J, Zhu J, Banunaryanan V, Mukherji R. Forecasting energy prices in a competitive market. IEEE Comput Appl Power 1999(July):40–5.
48. Deb R, Albert R, Hsue Lie-Long, Brown N. How to incorporate volatility and risk in electricity price forecasting / Electricity J 2000(May):1–16.
49. Anbazhagan S., Kumarappan N. Day-ahead deregulated electricity market price forecasting using neural network input featured by DCT / Energy Conversion and Management. Vol.78. 2014. pp. 711-719.
50. Воловиков С.А., Горячев Ю.В., Коновалов Д.М. Методы прогнозирования временных рядов: Учеб. пособие. - М.: Изд-во МГСУ «Союз», 1999. – 29 с.
51. Мишулина О. А. Статистический анализ и обработка временных рядов.— М.: МИФИ, 2004.— С. 180.
52. Афанасьев В.Н., Юзбашев М. М. Анализ временных рядов и прогнозирование , Инфра-М, 2010
53. MAVIR [Электронный ресурс] – Режим доступа до ресурсу: <https://www.mavir.hu/web/mavir-en/data-publication>
54. ENTSOE [Электронный ресурс] – Режим доступа до ресурсу: <https://transparency.entsoe.eu/dashboard/show>

55. Wattsight [Электронный ресурс] – Режим доступа до ресурсу:  
<https://app.wattsight.com/#tab/power/56/45>
56. HUPX [Электронный ресурс] – Режим доступа до ресурсу:  
<https://hupx.hu/en/>
57. Pogodaiklimat [Электронный ресурс] – Режим доступа до ресурсу:  
<http://www.pogodaiklimat.ru/forecast.php?id=hu>
58. Timeanddate [Электронный ресурс] – Режим доступа до ресурсу:  
<https://www.timeanddate.com/counters/newyearlocation.html>
59. Open Power System Data [Электронный ресурс] – Режим доступа до ресурсу: <https://open-power-system-data.org/data-sources>
60. CRAN [Электронный ресурс] – Режим доступа до ресурсу:  
<https://www.r-project.org/about.html>