

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»**

**Навчально-науковий інститут прикладного системного аналізу
Кафедра математичних методів системного аналізу**

До захисту допущено:

Завідувач кафедри

_____ Оксана ТИМОЩУК

«__» _____ 2025 р.

Дипломна робота

на здобуття ступеня бакалавра

за освітньо-професійною програмою «Системний аналіз і управління»

спеціальності 124 «Системний аналіз»

**на тему: «Моделі та прогнози демографічних змін у Німеччині
з урахуванням міграційних процесів»**

Виконала:

студентка IV курсу, групи КА-15

Лоїк Ірина Валеріївна _____

Керівник:

Ст. викладач кафедри ММСА, PhD, Левенчук Людмила Борисівна _____

Консультант з економічного розділу:

Доцент, к. е. н., Рощина Надія Василівна _____

Консультант з нормоконтролю:

к.ф.-м.н., Статкевич Віталій Михайлович _____

Рецензент:

Доцент кафедри ШІ, PhD, Гуськова Віра Геннадіївна _____

Засвідчую, що у цій дипломній роботі
немає запозичень з праць інших авторів
без відповідних посилань.

Студентка _____

Київ – 2025 року

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Навчально-науковий інститут прикладного системного аналізу
Кафедра математичних методів системного аналізу

Рівень вищої освіти – перший (бакалаврський)

Спеціальність – 124 «Системний аналіз»

Освітньо-професійна програма «Системний аналіз і управління»

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ Оксана ТИМОЩУК

« ____ » _____ 2025 р.

ЗАВДАННЯ

на дипломну роботу студентці

Лоїк Ірині Валеріївни

1. Тема роботи « **Моделі та прогнози демографічних змін у Німеччині з урахуванням міграційних процесів** », керівник роботи Левенчук Людмила Борисівна, ст. викладач кафедри ММСА, PhD, затверджені наказом по університету від «26» травня 2025 р. № 1759-с
2. Термін подання студентом роботи 11.06.2025 р.
3. Вихідні дані до роботи: Статистичні дані з міжнародних джерел: World Bank, Human Mortality Database, UN WPP 2024.
4. Зміст роботи: 1 – Дослідження предметної області; 2 – Теоретичні основи та вибір методів моделювання; 3 – Реалізація прогнозової моделі та аналіз результатів; 4 – Функціонально-вартісний аналіз програмного продукту.
5. Перелік ілюстративного матеріалу: презентація.

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Економічний	Рощина Н. В. , к. е. н., доц.		

7. Дата видачі завдання _____

Календарний план

№ з/п	Назва етапів виконання дипломної роботи	Термін виконання етапів роботи	Примітка
1	Аналіз предметної області та формулювання завдання	31.03.2025 - 06.04.2025	виконано
2	Вивчення літератури та обробка даних за темою роботи	07.03.2025 - 13.04.2025	виконано
3	Підготовка математичних моделей, розробка програмного коду їх реалізацій	14.04.2025 - 18.05.2025	виконано
4	Робота над розділами під час переддипломної практики	14.04.2025 - 18.05.2025	виконано
5	Оформлення та узгодження пояснювальної записки і розділів роботи	19.05.2025 - 01.06.2025	виконано
6	Попередній захист дипломної роботи	05.06.2025 - 06.06.2025	виконано
7	Захист дипломної роботи	16.06.2025- 20.06.2025	виконано

Студентка

Ірина ЛОЇК

Керівник

Людмила ЛЕВЕНЧУК

РЕФЕРАТ

Дипломна робота: 93 с., 10 табл., 19 рис., 3 дод., 18 джерел
ДЕМОГРАФІЯ, НІМЕЧЧИНА, ARIMA, СЦЕНАРНИЙ АНАЛІЗ, ЧАСОВІ
РЯДИ, МІГРАЦІЯ, PYTHON, ПРОГНОЗУВАННЯ

Тема: моделі та прогнози демографічних змін у Німеччині з
урахуванням міграційних процесів.

У роботі розглянуто питання моделювання демографічних змін у
Німеччині з урахуванням зовнішньої міграції, зниження фертильності та
старіння населення. Для цього застосовано сучасні методи аналізу часових
рядів, включаючи авторегресійні моделі (ARIMA), експоненційне
згладжування та логістичні моделі.

Реалізовано сценарне прогнозування з варіацією ключових факторів,
зокрема рівня народжуваності та обсягів нетто-міграції. Прогноз побудовано
на період до 2035 року.

Об'єкт дослідження: демографічна динаміка Федеративної Республіки
Німеччина.

Предмет дослідження – математичні моделі, що описують тенденції
чисельності населення з урахуванням міграційних і народжувальних
процесів.

Мета роботи – підготувати обґрунтований прогноз демографічної
ситуації в Німеччині на основі аналізу реальних статистичних даних,
визначити вплив змін у міграційній політиці та динаміці TFR на майбутню
чисельність населення.

Методи дослідження – ARIMA(p,d,q), перевірка стаціонарності (ADF),
побудова сценаріїв, аналіз залишків, побудова трендів, Python.

В роботі створено програмну систему, яка дозволяє здійснювати
обчислення та візуалізацію результатів, змінювати вхідні параметри
прогнозу та перевіряти чутливість моделей до змін у припущеннях.

ABSTRACT

Thesis: 93 pages, 10 tables, 19 figures, 3 appendices, 18 references
DEMOGRAPHY, GERMANY, ARIMA, SCENARIO ANALYSIS, TIME
SERIES, MIGRATION, PYTHON, FORECASTING.

Topic: models and forecasts of demographic changes in Germany considering migration processes.

This thesis addresses the modeling of demographic changes in Germany, taking into account external migration, declining fertility, and population aging. Modern time series analysis methods are applied, including autoregressive models (ARIMA), exponential smoothing, and logistic models.

Scenario forecasting is implemented with variation in key factors, particularly total fertility rate and net migration volume. The forecast horizon extends to the year 2035.

Object of research: demographic dynamics of the Federal Republic of Germany.

Subject of research: mathematical models describing population change trends with consideration of migration and fertility processes.

Purpose of the work: to develop a substantiated demographic forecast for Germany based on real statistical data, and to assess the impact of migration policy changes and TFR dynamics on future population size.

Methods of research: ARIMA(p,d,q), stationarity testing (ADF), scenario modeling, residual analysis, trend construction, Python programming.

A custom software tool has been developed that enables automated calculation and visualization of results, adjustment of input forecast parameters, and sensitivity analysis of the model under various assumptions.

ЗМІСТ

ВСТУП.....	8
РОЗДІЛ 1 ДОСЛІДЖЕННЯ ПРЕДМЕТНОЇ ОБЛАСТІ.....	9
1.1 Актуальність демографічного прогнозування в сучасній Європі	9
1.2 Основні демографічні показники	10
1.2.1 Сумарний коефіцієнт народжуваності (TFR).....	11
1.2.2 Нетто-міграція.....	11
1.2.3 Смертність та вікова структура населення.....	12
1.3 Демографічна ситуація в Німеччині.....	14
1.3.1 Історичні та сучасні тенденції.....	14
1.3.2 Міграційна динаміка та її вплив.....	16
1.4 Висновки до розділу 1	16
РОЗДІЛ 2 ТЕОРЕТИЧНІ ОСНОВИ ТА МЕТОДИ МОДЕЛЮВАННЯ.....	18
2.1 Структура та побудова моделі демографічного прогнозу	18
2.2 Огляд методів прогнозування	19
2.3 Сценарний підхід: роль в прогнозах населення.....	21
2.4. Модель ARIMA в демографічному прогнозуванні.....	22
2.5 Логістична модель у демографічному прогнозуванні.....	23
2.6 Трансформація Вох–Сох у демографічному прогнозуванні.....	25
2.8 Критерії оцінки якості прогнозів.....	28
2.9 Висновки до розділу 2	30
РОЗДІЛ 3 РЕАЛІЗАЦІЯ МОДЕЛІ ТА АНАЛІЗ РЕЗУЛЬТАТІВ.....	31
3.1 Обґрунтування вибору країни, показників і методів.....	31
3.2 Вибір моделей і змінних.....	32

	7
3.3 Джерела даних	33
3.4. Прогнозування сумарного коефіцієнта народжуваності (TFR).....	33
3.5. Прогнозування нетто-міграції населення Німеччини до 2035 року.....	37
3.6. Сценарне прогнозування міграції.....	41
3.7. Прогнозування очікуваної тривалості життя	42
3.8 Сценарне прогнозування чисельності населення Німеччини до 2035 року.....	46
3.9 Висновки до розділу 3	49
РОЗДІЛ 4 ФУНКЦІОНАЛЬНО-ВАРТІСНИЙ АНАЛІЗ ПРОГРАМНОГО ПРОДУКТУ.....	51
4.1 Постановка задачі проектування	52
4.2 Обґрунтування функцій програмного продукту	52
4.3 Обґрунтування системи параметрів програмного продукту	57
4.4 Аналіз експертного оцінювання параметрів	59
4.5 Аналіз рівня якості варіантів реалізації функцій.....	63
4.6 Економічний аналіз варіантів розробки ПП	65
4.7 Вибір кращого варіанту ПП техніко-економічного рівня.....	69
4.8. Висновки до розділу 4	70
ВИСНОВКИ	71
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ	72
ДОДАТОК А (ілюстративний матеріал доповіді).....	74
ДОДАТОК Б (статистичні дані).....	83
ДОДАТОК В (код програмного продукту).....	86

ВСТУП

Демографічні процеси мають визначальний вплив на соціально-економічний розвиток держав і формування довгострокових суспільних тенденцій. Особливої уваги ці питання набувають у європейських країнах, зокрема в Німеччині, де останніми десятиліттями спостерігається стійка тенденція до зниження народжуваності, старіння населення та зростання ролі міграції як компенсуючого чинника.

Дослідження зосереджується на аналізі демографічної динаміки Німеччини та побудові прогностичних моделей на основі статистичних даних. Особливу увагу приділено показникам народжуваності, смертності, міграційного приросту та загальної чисельності населення. Використання математичних моделей дозволяє не лише виявити наявні тенденції, а й оцінити можливі сценарії розвитку демографічної ситуації в майбутньому.

Для досягнення поставлених цілей у роботі застосовано сучасні методи аналізу часових рядів, зокрема ARIMA-моделі, експоненційне згладжування, логістичну модель, а також сценарний підхід. Для оцінки якості прогнозів та вибору найбільш адекватної моделі використовувалися статистичні критерії, такі як середня абсолютна відсоткова похибка (MAPE), коефіцієнт детермінації (R^2), критерій Акаїке (AIC) та статистика Дарбіна-Уотсона (DW). Обробка даних та побудова моделей здійснювалися у середовищі Python.

Робота структурована наступним чином. У першому розділі здійснено загальний огляд демографічної ситуації в Німеччині та її особливостей. Другий розділ присвячено теоретичному аналізу методів прогнозування та порівнянню їхніх характеристик. У третьому розділі реалізовано практичне моделювання, побудовано прогноз чисельності населення до 2035 року та виконано порівняння результатів за різними сценаріями. У четвертому розділі виконано функціонально-вартісний аналіз використаної моделі прогнозування.

РОЗДІЛ 1 ДОСЛІДЖЕННЯ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Актуальність демографічного прогнозування в сучасній Європі

У ХХІ столітті країни Європи опинилися перед масштабними демографічними викликами, які мають далекосяжні соціальні, економічні та політичні наслідки. Найбільш помітними з них є: зниження рівня народжуваності, стрімке старіння населення, міграційні кризи, скорочення чисельності населення в ряді країн, а також нерівномірний розподіл населення в межах регіонів. Ці процеси суттєво змінюють структуру європейських суспільств і вимагають адаптації державної політики.

Більшість країн Європейського Союзу мають сумарний коефіцієнт народжуваності нижчий за рівень простого відтворення населення. Це призводить до зменшення чисельності працездатного населення та зростання демографічного навантаження на пенсійну і медичну системи [1, 2].

Старіння населення в Європі також досягає критичних показників. Медіанний вік у таких країнах, як Німеччина та Італія, перевищує 46 років. Частка осіб віком 65+ постійно зростає, а демографічне навантаження на осіб працездатного віку посилюється. Наприклад, у Німеччині в 2021 році на кожні 100 осіб віком 20–65 років припадало 37 осіб віком 65+, і ця цифра, за прогнозами, зросте до 70 упродовж найближчих десятиліть [3].

Міграція стала компенсаторним механізмом демографічного скорочення. Проте вона створює нові виклики: зокрема, потребу в інтеграційній політиці, адаптації інфраструктури, зміні етнічної та релігійної структури населення [1].

Усе це вимагає науково обґрунтованого, статистично точного прогнозування демографічних змін. Сучасні інструменти, такі як ARIMA-моделювання, сценарний аналіз та статистичне моделювання, дозволяють

створювати надійні прогнози чисельності населення, вікової структури та демографічного навантаження.

Таким чином, актуальність демографічного прогнозування в Європі обумовлена потребою в прийнятті стратегічних управлінських рішень, плануванні соціальної, освітньої та медичної інфраструктури, а також забезпеченні стійкого розвитку в умовах демографічного переходу.

1.2 Основні демографічні показники

Аналіз демографічних процесів базується на вивченні основних компонентів змін чисельності населення: народжуваності, смертності, міграції та вікової структури. У демографії ці елементи утворюють основу балансової моделі зміни населення, де приріст визначається як сума природного приросту та міграційного сальдо [4].

Хоча в довгостроковій перспективі внесок міграції часто є менш стабільним, ніж природний приріст, обидва фактори залишаються ключовими для аналізу чисельності та динаміки населення. У сучасних підходах до прогнозування, таких як когортно-компонентна модель, ці компоненти використовуються для обчислення чисельності населення в майбутньому з урахуванням віку та статі [5].

Зміни в демографічній поведінці — зокрема рівень народжуваності та міграційна активність — значною мірою залежать від зовнішніх соціально-економічних умов. Економічні кризи, нестабільність, нерівність, або навпаки, зростання добробуту можуть стимулювати або стримувати народжуваність і міграцію, що, у свою чергу, впливає на структуру та чисельність населення.

1.2.1 Сумарний коефіцієнт народжуваності (TFR)

Сумарний коефіцієнт народжуваності (Total Fertility Rate, TFR) є ключовим індикатором у демографічному аналізі та прогнозуванні. Він відображає середню кількість дітей, яких народить жінка за все життя за умови збереження поточного вікового розподілу народжуваності [4].

TFR суттєво впливає не лише на абсолютну чисельність населення, але й на його вікову структуру. В країнах із низьким рівнем фертильності поступово зростає частка осіб похилого віку, що призводить до демографічного старіння [6].

Більшість міжнародних прогнозів, зокрема прогнози ООН, базуються на сценарному підході до оцінки TFR. Найбільш поширеними є три варіанти: середній (Medium Variant), низький (Low Variant) — на 0,5 дитини менше, і високий (High Variant) — на 0,5 дитини більше базового. Такий підхід дозволяє враховувати невизначеність динаміки фертильності [7].

Особливістю фертильності як показника є її залежність від комплексу соціальних, економічних і культурних чинників. До них належать рівень освіти, урбанізація, зайнятість жінок, доступ до медичних послуг, соціальна політика тощо [4, 5].

1.2.2 Нетто-міграція

Нетто-міграція є основним показником, що відображає міграційний баланс країни. Вона визначається як різниця між числом осіб, що прибули до країни на постійне проживання, і числом тих, хто виїхав за її межі за певний період [4]. Цей показник може бути як позитивним (коли імміграція переважає), так і негативним (при переважанні еміграції).

У демографічному прогнозуванні нетто-міграція відіграє важливу роль, особливо в країнах із низькою народжуваністю та зростаючим рівнем старіння населення. У таких умовах імміграція компенсує природне скорочення населення та сприяє збереженню чисельності економічно активного населення [1, 5].

Особливість міграційних процесів полягає у високій нестабільності: на відміну від народжуваності та смертності, вони надзвичайно чутливі до політичної ситуації, воєн, економічних криз, змін у законодавстві. Це ускладнює побудову достовірних довгострокових прогнозів. Як правило, у демографічних моделях використовується сценарний підхід або статистичне моделювання (наприклад, ARIMA) для опису тенденцій міграції [5, 7].

Крім абсолютного значення, важливим є також коефіцієнт міграційного приросту на 1000 осіб, який дозволяє порівнювати країни з різною чисельністю населення. У когортно-компонентному моделюванні також використовуються вікові та статеві міграційні коефіцієнти для точнішого прогнозування впливу міграції на структуру населення [4, 6].

1.2.3 Смертність та вікова структура населення

Смертність є другою основною складовою природного руху населення після народжуваності. У демографічному аналізі використовуються такі основні показники: коефіцієнт загальної смертності (Crude Death Rate, CDR), очікувана тривалість життя при народженні (Life Expectancy at Birth), а також віковоспецифічні коефіцієнти смертності [4, 6].

Коефіцієнт загальної смертності визначає кількість смертей на 1000 осіб середньої чисельності населення на рік. Цей показник дає загальне уявлення про рівень смертності, але не враховує вікову структуру, тому його аналітична

цінність обмежена. Значно точнішими є віковоспецифічні коефіцієнти, які дозволяють порівнювати смертність у різних вікових групах та за статтю [4].

Очікувана тривалість життя при народженні є найважливішим агрегованим показником стану здоров'я населення та якості життя. Зростання цього показника у більшості європейських країн зумовлено розвитком медицини, покращенням умов життя та гігієни. Проте воно також спричиняє старіння населення — збільшення частки осіб похилого віку [5, 6].

У прогнозуванні чисельності населення смертність зазвичай враховується через побудову таблиць смертності за віком і статтю. Замість безпосереднього прогнозування коефіцієнтів смертності, більшість моделей ґрунтуються на припущеннях щодо подальшого зростання очікуваної тривалості життя [7]. На основі цих припущень формуються відповідні ймовірності смерті (q_x) для кожної вікової групи.

Щодо вікової структури, у більшості країн Європи, включно з Німеччиною, спостерігається процес старіння населення. Це зумовлено зниженням народжуваності та збільшенням тривалості життя. Поступово частка осіб віком 65+ зростає, тоді як питома вага молоді та осіб працездатного віку зменшується [1, 3].

Старіння населення створює серйозні виклики для системи соціального забезпечення, охорони здоров'я, ринку праці та пенсійного забезпечення. У моделюванні демографічних процесів цей аспект враховується через зміщення вікових груп у майбутньому прогнозі, що дозволяє оцінити навантаження на систему соціальної підтримки та потенціал економічного розвитку.

1.3 Демографічна ситуація в Німеччині

Станом на 2024 рік Німеччина є однією з найбільших за чисельністю населення країн Європейського Союзу, але водночас однією з найбільш демографічно вразливих [8].

Населення країни складає понад 83 мільйони осіб, однак у довгостроковій перспективі очікується його скорочення. Основними демографічними характеристиками є низький рівень народжуваності (TFR \approx 1,5), висока тривалість життя (понад 81 рік) та суттєве старіння населення. За прогнозами UN DESA, до 2050 року частка осіб віком 65+ перевищить 30% від загальної чисельності населення [9].

У поєднанні з цими трендами особливу роль відіграє міграція, яка стала головним компенсатором негативного природного приросту. За офіційними даними Destatis, у 2022 році міграційне сальдо склало понад 1,5 млн осіб, що стало ключовим фактором зростання населення, попри перевищення смертності над народжуваністю [10].

1.3.1 Історичні та сучасні тенденції

Демографічна ситуація в Німеччині упродовж XX і початку XXI століття формувалася під впливом глибоких політичних, економічних і соціальних змін, включно з наслідками двох світових воєн, процесом возз'єднання країни, а також масштабними хвилями міграції.

У післявоєнний період у Західній Німеччині (ФРН) спостерігався так званий baby boom — демографічний вибух, що припав на 1950–1960-ті роки. Абсолютний пік народжуваності зафіксовано у 1964 році, коли в Німеччині народилося 1,36 млн дітей. Відтоді кількість новонароджених почала

зменшуватись, і в 2011 році було зафіксовано історичний мінімум — лише 663 000 новонароджених [8].

Сумарний коефіцієнт народжуваності (TFR) у ці роки перевищував 2,5 дитини на одну жінку, але починаючи з 1970-х років почалося тривале зниження фертильності. До початку 1990-х TFR зменшився до приблизно 1,4 і залишається стабільно низьким до сьогодні [9].

Паралельно зі скороченням народжуваності відбувалося поступове зростання тривалості життя. Згідно з даними World Population Prospects, очікувана тривалість життя при народженні в Німеччині зросла з приблизно 67 років у 1950-х до понад 81 року у 2023 році [9]. Ці зміни, у поєднанні з низькою народжуваністю, призвели до значного демографічного старіння населення.

Після об'єднання Німеччини у 1990 році спостерігалися різкі регіональні демографічні зсуви: у Східній Німеччині (колишній НДР) народжуваність різко впала, а значна частина молоді мігрувала на захід, що призвело до депопуляції східних регіонів і порушення внутрішнього демографічного балансу [6].

У 2000-х роках чисельність населення Німеччини почала зменшуватись природним шляхом. Згідно з даними Destatis, у період 2003–2013 років смертність стабільно перевищувала народжуваність, що зумовило негативний природний приріст. Єдиним компенсуючим фактором стали міграційні надходження, які значно зросли після кризи 2015 року — зокрема, через притік біженців і робочої сили [10, 11].

Останні демографічні прогнози свідчать, що Німеччина перебуває на завершній стадії демографічного переходу, з характерними ознаками: низька народжуваність, висока тривалість життя та швидке старіння населення. Це створює передумови для тривалого демографічного тиску і необхідність структурних змін у соціальній політиці, імміграційній системі та системі охорони здоров'я [6, 9].

1.3.2 Міграційна динаміка та її вплив

Починаючи з початку 2000-х років, міграція стала основним фактором, який компенсує природне скорочення населення Німеччини. У період з 2003 по 2013 рік рівень смертності стабільно перевищував народжуваність, що призводило до негативного природного приросту. У цей час саме міграційне сальдо дозволяло уникнути загального зменшення чисельності населення [10].

Особливо значним став притік мігрантів у 2015 році, коли в країну прибуло близько 890 тисяч осіб, що шукали притулку [11]. Після цього міграційні надходження залишаються стабільно високими. Наприклад, у 2022 році чистий приріст населення за рахунок міграції склав понад 1,5 мільйона осіб, що повністю компенсувало природне скорочення [10].

За даними Prospects [1], внесок міжнародної міграції в загальні демографічні зміни в Німеччині оцінюється як критично важливий у коротко- та середньостроковій перспективі. Міграція сприяє оновленню працездатного населення та уповільнює процес старіння, однак не може повністю змінити тренд зростання частки літніх осіб [9].

Таким чином, без урахування міграції довгостроковий прогноз населення Німеччини був би значно більш песимістичним. Але навіть за її збереженням на нинішньому рівні демографічне старіння залишається основною тенденцією, що зумовлює необхідність реформ у сфері соціальної політики, охорони здоров'я та ринку праці.

1.4 Висновки до розділу 1

У першому розділі було розглянуто основні аспекти демографічного аналізу в контексті сучасної Європи та Німеччини. Демографічне

прогнозування набуває особливого значення в умовах низької народжуваності, зростаючої тривалості життя та посилення міграційних процесів. Було охарактеризовано ключові показники — TFR, смертність, міграційне сальдо — та описано їхнє значення у прогнозуванні майбутньої чисельності населення.

У фокусі дослідження опинилась Німеччина як приклад країни з завершеним демографічним переходом. Історичні та сучасні тенденції свідчать про глибокі структурні зміни, серед яких: скорочення природного приросту, старіння населення, регіональні дисбаланси та зростаюча залежність від імміграції.

Отримані дані та огляд літератури сформували основу для подальшого моделювання демографічної динаміки, що буде представлено в наступних розділах.

РОЗДІЛ 2 ТЕОРЕТИЧНІ ОСНОВИ ТА МЕТОДИ МОДЕЛЮВАННЯ

2.1 Структура та побудова моделі демографічного прогнозу

Для реалізації демографічного прогнозування важливо обґрунтувати загальну логіку побудови моделі та обрати відповідні інструменти. Коректно визначена структура моделі дає змогу врахувати особливості вхідних даних, типові властивості часових рядів та характер змінних, що підлягають прогнозуванню.

На початковому етапі здійснюється аналіз статистичних показників, що характеризують демографічну ситуацію, зокрема рівень народжуваності, смертності та міграції. Далі перевіряється стаціонарність часових рядів, оскільки нестійкість у даних може істотно вплинути на точність прогнозу. Для цього використовуються, зокрема, тест Дікі-Фуллера, тест KPSS тощо.

Наступним кроком є визначення типу моделі (AR, ARMA, ARIMA), встановлення її порядку та виявлення потенційної сезонності. Залежно від структури даних та статистичних характеристик, можуть бути також застосовані логістичні моделі, експоненційне згладжування або сценарний підхід.

Після побудови моделей-кандидатів виконується оцінювання їхньої якості за допомогою таких критеріїв, як середня абсолютна відсоткова похибка (MAPE), коефіцієнт детермінації (R^2), інформаційний критерій Акаїке (AIC) та статистика Дарбіна-Уотсона (DW). Найбільш релевантна модель обирається для подальшого прогнозування.

2.2 Огляд методів прогнозування

У демографічному аналізі застосовуються різноманітні підходи до прогнозування залежно від природи досліджуваного явища, типу даних та горизонту прогнозу. Найбільш поширеними є методи часових рядів, регресійні моделі, логістичні функції та імовірнісні підходи.

Методи часових рядів, зокрема ARIMA (AutoRegressive Integrated Moving Average), дозволяють враховувати автокореляцію, тренди та сезонні компоненти в динаміці показників, і тому широко застосовуються для прогнозування змін народжуваності, тривалості життя чи міграції. ARIMA є одним із найбільш універсальних інструментів для роботи з нестационарними демографічними рядами, оскільки включає компоненти авторегресії, інтегрування та ковзного середнього, що забезпечує гнучкість моделі.

Для змін, що мають обмежений або асимптотичний характер (наприклад, тривалість життя), використовуються логістичні моделі. Вони описують зростання, яке сповільнюється у часі, і добре підходять для випадків, коли показник наближається до природної межі. Логістична крива дозволяє врахувати насичення та наслідки структурних змін у суспільстві.

Крім того, в демографічних дослідженнях іноді використовуються методи ETS (Exponential smoothing), які добре працюють для короткострокового прогнозу при наявності стабільного тренду або сезонності. Однак ці методи менш адаптивні до складних структур часових рядів.

Демографічне прогнозування — це процес побудови обґрунтованих оцінок майбутньої чисельності населення, його вікової структури, а також змін у народжуваності, смертності та міграції. Методологічні підходи до побудови таких прогнозів істотно відрізняються за своєю складністю, припущеннями, гнучкістю та чутливістю до вхідних даних.

Основні підходи, що найчастіше використовуються в науковій і прикладній демографії, наведено нижче.

1. Когортно-компонентна модель є класичним детерміністичний підходом у якому населення розподіляється на віково-статеві когорти, і для кожної такої групи окремо прогнозуються зміни, пов'язані з народжуваністю, смертністю та міграцією.
2. Стохастичні (ймовірнісні) моделі є підходом, де основні демографічні змінні (наприклад, TFR або міграція) розглядаються як випадкові процеси. Одним із найвідоміших прикладів є байєсівська модель для прогнозування фертильності, яка була розроблена Alkema [5] і застосована у прогнозах World Population Prospects. Такий підхід дозволяє формувати інтервали довіри та враховувати структурну невизначеність майбутнього, що є особливо важливим для довгострокового прогнозування.
3. Сценарне моделювання базується на створенні кількох альтернативних припущень щодо майбутньої динаміки ключових показників — таких як народжуваність, тривалість життя, нетто-міграція. Зазвичай застосовуються три сценарії: Low, Medium та High. Сценарний підхід не враховує імовірності реалізації кожного варіанта, проте дозволяє аналізувати чутливість системи до змін ключових параметрів.
4. Машинне навчання та нейромережі є сучасними методи на основі штучного інтелекту застосовуються переважно для індивідуальних прогнозів, наприклад, тривалості життя конкретної особи або ймовірності міграції. Хоча потенціал цих методів значний, вони ще не досягли стійких результатів у макродемографічному моделюванні через обмежену точність і високу стохастичність вхідних змінних [6].

2.3 Сценарний підхід: роль в прогнозах населення

Сценарне прогнозування є ключовим елементом сучасних демографічних оцінок, особливо у випадках, коли довгострокова точність передбачень залежить від змін зовнішніх умов — таких як соціальна політика, міграційні потоки, медична інфраструктура або економічні кризи. Методологічно сценарний підхід базується не на фіксованій оцінці майбутнього, а на формуванні множини можливих траєкторій розвитку, що дає змогу врахувати невизначеність і побудувати гнучкі стратегії управління [17].

В основі сценарного підходу лежить ідея побудови альтернативних припущень щодо ключових демографічних параметрів: сумарної народжуваності (TFR, англ. Total Fertility Rate), очікуваної тривалості життя (LE, англ. Life Expectancy) та нетто-міграції.

У прогнозах UN DESA[16], починаючи з 2010-х років, кожен з цих компонентів моделюється у трьох основних варіантах.

1. Low — фіксує гіпотетичне погіршення показника (наприклад, зниження TFR до 1.3);
2. Medium — найімовірніший сценарій (наприклад, стабілізація TFR на рівні 1.6);
3. High — сприятливий розвиток подій (помірне зростання до 1.9).

Однак, головним недоліком є відсутність ймовірнісної ваги сценаріїв: усі варіанти вважаються умовно рівноймовірними. Крім того, сценарії не мають автоматичного механізму адаптації до нових даних — на відміну від стохастичних моделей, які оновлюються при зміні вхідних рядів [17].

Тому в сучасній практиці демографічного прогнозування часто поєднують сценарні моделі з когортно-компонентним або ARIMA-аналізом: сценарії задають «рамки» змінних, а динаміку в межах цих рамок уточнюють статистичні моделі [1].

2.4. Модель ARIMA в демографічному прогнозуванні

Серед широкого спектру методів прогнозування часових рядів особливе місце посідає модель ARIMA (Autoregressive Integrated Moving Average), яка широко застосовується для моделювання демографічних змін, зокрема динаміки фертильності, смертності або міграції. Основна ідея моделі полягає в тому, щоб описати часовий ряд як комбінацію його власних попередніх значень, похибок прогнозу та тренду, усуненого шляхом диференціювання.

Модель ARIMA є узагальненням трьох підходів: авторегресії (AR), інтегрованої складової (I), тобто диференціювання для досягнення стаціонарності, та моделі ковзного середнього (MA). Математичне представлення моделі ARIMA наведено у формулі (2.1):

$$\phi(B)(1 - B)^d y_t = \theta(B)\varepsilon_t \quad (2.1)$$

де y_t — значення часового ряду в момент часу;

B — оператор зсуву назад;

D — порядок диференціювання для досягнення стаціонарності;

$\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ — поліном авторегресії порядку p ;

$\theta(B) = 1 + \theta_1 B + \dots + \theta_p B^p$ — поліном ковзного середнього порядку;

ε_t — білий шум (незалежні нормально розподілені випадкові змінні з нульовим середнім).

У контексті демографії застосування ARIMA дозволяє врахувати внутрішні залежності показників у часі, зокрема затримки, інерційність змін та циклічність.

У прогнозах, наприклад, рівня народжуваності (TFR), модель ARIMA дозволяє побудувати як точковий, так і інтервальний прогноз на кілька років уперед. Вона також ефективна для моделювання нетто-міграції, яка має складну структуру автокореляції, але при цьому достатньо довгі ряди спостережень для надійної оцінки параметрів.

Застосування ARIMA в демографічному аналізі підтримується численними джерелами. Наприклад, у дослідженні Keilman і Mazzucco [5] автори застосовують ARIMA-моделі для короткострокового прогнозування фертильності, акцентуючи увагу на їхній гнучкості та здатності адаптуватися до особливостей емпіричних даних. У монографії Yusuf, Martins і Swanson [7] метод ARIMA розглядається як базовий інструмент екстраполяції трендів демографічних показників, особливо у випадках, коли відсутня складна структурна інформація або прогноз здійснюється в умовах високої невизначеності.

Попри свою ефективність, ARIMA має і обмеження. Вона не враховує вплив зовнішніх змінних (наприклад, політики чи економічних криз) і працює тільки з внутрішньою динамікою часового ряду. Також модель передбачає лінійність зв'язків, що не завжди є характерним для демографічних процесів. Саме тому в практиці прогнозування часто поєднують ARIMA з іншими методами — стохастичними, сценарними або компонентними.

Таким чином, ARIMA є одним з основних інструментів демографічного прогнозування, який при належній перевірці та адаптації до специфіки даних забезпечує надійні результати і служить основою для формування сценаріїв майбутнього розвитку населення.

2.5 Логістична модель у демографічному прогнозуванні

Логістична модель — це один із класичних підходів до моделювання зростання, що часто застосовується в демографії для прогнозування показників, які мають природну межу, зокрема очікуваної тривалості життя при народженні. Її теоретичне підґрунтя бере початок із моделі Вергуля, яка спочатку використовувалася для опису біологічних популяцій, але з часом знайшла широке застосування у соціальних науках і демографії [1].

Сутність логістичної моделі полягає в описі зростання змінної, яке з часом уповільнюється та асимптотично наближається до граничного рівня L , тобто межі, яку не може перевищити величина через біологічні, соціальні або інституційні обмеження. У випадку тривалості життя такою межею є умовно максимально можлива тривалість життя людини в майбутньому.

Аналітичний вигляд логістичної функції (2.2) має форму:

$$f(t) = \frac{L}{1 + \exp(-k(t-t_0))} \quad (2.2)$$

де $f(t)$ — прогнозоване значення показника в рік;

L — граничне (асимптотичне) значення показника;

K — швидкість росту (градієнт кривої);

t_0 — рік, у якому зростання має точку перегину (інфлексії).

Значення параметрів, та зазвичай визначаються емпіричним шляхом — за допомогою методу найменших квадратів або інших оптимізаційних процедур. Відомо, що темпи зростання тривалості життя поступово знижуються з часом. Саме тому логістична модель краще підходить, ніж лінійна чи експоненціальна, які переоцінюють майбутні значення.

Водночас вона має і певні обмеження: не враховує випадкові флуктуації та може недооцінювати вплив раптових змін, таких як медичні інновації або пандемії. Тому часто логістичну модель поєднують з ARIMA або іншими стохастичними підходами для кращого врахування варіативності.

Таким чином, логістична модель виступає зручною та надійною основою для прогнозу показників, які мають верхню межу розвитку, зокрема тривалості життя. У межах демографічного прогнозування її застосування є виправданим з огляду на біологічні межі та історичну тенденцію сповільнення зростання показника LE.

2.6 Трансформація Вох–Сох у демографічному прогнозуванні

У практиці побудови демографічних прогнозів одним із важливих кроків є попередня обробка даних, зокрема стабілізація дисперсії та наближення розподілу до нормального. Для досягнення цих цілей у статистичному аналізі широко використовується трансформація Вох–Сох — параметричне перетворення, що дозволяє зробити дані більш відповідними вимогам лінійних і авторегресивних моделей.

Трансформація задається наступною формулою (2.3):

$$y^{(\lambda)} = \begin{cases} \frac{y^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \ln(y), & \lambda = 0 \end{cases} \quad (2.3)$$

де $y^{(\lambda)}$ — трансформоване значення змінної;

y — початкове позитивне значення;

λ — параметр трансформації.

Основна мета цієї трансформації — стабілізувати дисперсію часового ряду та зробити розподіл залишків ближчим до нормального. Це особливо важливо для моделей типу ARIMA, які передбачають гомоскедастичність та нормальність похибок.

Оптимальне значення параметра λ зазвичай підбирається шляхом максимізації логарифмічної функції правдоподібності або за допомогою евристичних критеріїв, таких як мінімізація коефіцієнта варіації [18].

У демографічному прогнозуванні трансформація Вох–Сох ефективно застосовується до таких показників, як очікувана тривалість життя (LE), TFR або коефіцієнти міграції, особливо якщо вони демонструють зростаючу дисперсію в часі або асиметричний розподіл. Наприклад, для ряду тривалості життя в Німеччині було проведено логарифмічну трансформацію перед побудовою ARIMA-моделі з метою покращення залишків і точності прогнозу.

Після побудови моделі на трансформованих даних обов'язковим є виконання оберненої трансформації (2.4), яка для $\lambda \neq 0$ має вигляд:

$$y = (\lambda y^{(\lambda)} + 1)^{\frac{1}{\lambda}} \quad (2.4)$$

а для $\lambda=0$ (2.5):

$$y = \exp(y^{(0)}) \quad (2.5)$$

Таким чином, Вох–Сох трансформація виконує подвійну функцію: вона забезпечує математичну коректність моделі (через приведення даних до умов застосування ARIMA) та підвищує точність прогнозу за рахунок зменшення гетероскедастичності.

2.7. Інші підходи до прогнозування демографічних змін

Окрім ARIMA-моделі, логістичного підходу та трансформації Вох–Сох, у демографічному прогнозуванні застосовуються й інші математико-статистичні методи. Деякі з них можуть використовуватися як допоміжні інструменти для первинного аналізу, порівняння результатів або перевірки гіпотез.

Цей метод передбачає апроксимацію змінної у часі за допомогою прямої за формулою (2.6):

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t \quad (2.6)$$

де y_t — значення показника в момент часу t ;

β_0, β_1 — коефіцієнти моделі;

ε_t — випадкова похибка.

Лінійна регресія часто застосовується для короткострокового прогнозування за умов сталого тренду, але вона не враховує автокореляцію залишків і може бути непридатною в умовах структурних зламів. Її використання в демографії обмежене переважно до первинного аналізу тенденцій, зокрема для візуалізації довгострокових змін TFR або LE [1].

Цей метод ґрунтується на зваженому середньому попередніх значень часового ряду, де ваги експоненційно зменшуються в минуле. Найбільш

поширеною є модель простого експоненційного згладжування (SES), яка задається формулою (2.7):

$$\hat{y}_{t+1} = \alpha y_t + (1 - \alpha) \hat{y}_t \quad (2.7)$$

де:

\hat{y}_{t+1} — прогнозоване значення;

$\alpha \in [0, 1]$ — параметр згладжування.

У демографічному прогнозуванні експоненційне згладжування застосовується рідше, але може бути корисним для згладжування флуктуацій у річних показниках народжуваності або міграції, особливо при нестабільному зовнішньому середовищі.

Подібно до логістичної моделі, криві типу Гомпертца використовуються для прогнозування процесів, що мають межу зростання, зокрема тривалості життя або рівня урбанізації. Наприклад, функція Гомпертца (2.8) задається як:

$$y(t) = L * \exp(-\exp(-k(t - t_0))) \quad (2.8)$$

де параметри аналогічні логістичній моделі, але функція має іншу динаміку кривизни. Такі підходи використовуються переважно для візуального моделювання обмеженого зростання. Стохастичні підходи дозволяють враховувати випадкові флуктуації, що притаманні демографічним показникам, особливо у довгостроковому горизонті. Як приклад — моделі народження–смерті або рівняння Іто, які поєднують тренд і шум. Їх реалізація складніша, але дозволяє отримувати довірчі інтервали та моделювати невизначеність. Особливо це актуально для міграції, де вплив зовнішніх факторів важко передбачити [5].

Попри наявність широкого спектру підходів, у даній роботі було вирішено обмежитися застосуванням моделей ARIMA, логістичної функції та трансформації Вох–Сох, оскільки саме вони найкраще відповідають структурі даних, цілям дослідження та вимогам до точності прогнозу. Інші методи були протестовані на етапі попереднього аналізу, однак не були обрані як основні інструменти для побудови сценаріїв майбутньої демографічної динаміки.

2.8 Критерії оцінки якості прогнозів

Оцінювання якості прогнозних моделей у демографії є критичним етапом, який дозволяє обґрунтувати вибір математичного підходу, виявити слабкі сторони моделей та підвищити достовірність результатів. У сучасній статистичній практиці використовують комбінацію трьох груп критеріїв: метричних похибок (MAE, RMSE тощо), інформаційних критеріїв (AIC, BIC), аналізу залишків (автокореляція, нормальність, гетероскедастичність).

ARIMA-моделі вимагають, щоб часова серія була стаціонарною — тобто мала стабільне середнє, дисперсію та автоковаріаційну структуру. Для перевірки цього зазвичай використовується тест Дікі-Фуллера (ADF):

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \dots + \delta_{p-1} \Delta y_{t-p} + \varepsilon_t \quad (2.9)$$

Нульова гіпотеза $H_0: \gamma=0$ (ряд має одиничний корінь, тобто нестаціонарний).

Якщо $p\text{-value} < 0.05$ — стаціонарність досягнута.

Також перевіряється гетероскедастичність залишків. Це нестабільність дисперсії залишків, що порушує припущення класичних моделей. Для її виявлення використовують тест Бройша–Пагана або візуальний аналіз графіків залишків. У ARIMA моделі гетероскедастичність може свідчити про потребу в Вох-Сох трансформації або використанні GARCH-типів моделей.

Наступний критерій AIC (Akaike Information Criterion) Акаїке дозволяє порівняти моделі з різною кількістю параметрів:

$$AIC = 2k - 2\ln(L) \quad (2.10)$$

де

k — кількість параметрів моделі;

L — максимальне значення функції правдоподібності.

Менше значення AIC означає кращу модель.

Також використовується BIC (Bayesian Information Criterion):

$$BIC = \ln(n) * k - 2 \ln(L) \quad (2.11)$$

де n — кількість спостережень.

Також важливими є метрики точності: MAE (2.11), MSE (2.12), RMSE (2.13):

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t| \quad (2.12)$$

$$MSE = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2 \quad (2.13)$$

$$RMSE = \sqrt{MSE} \quad (2.14)$$

Ці показники вимірюють середнє відхилення прогнозу від реального значення.

Тест Льюнга–Бокса використовується для перевірки автокореляції залишків:

$$Q = n(n+2) \sum_{k=1}^h \frac{\hat{p}_k^2}{n-k} \quad (2.15)$$

де p^k — автокореляція порядку;

k, n — кількість спостережень.

Нульова гіпотеза: залишки незалежні. Якщо $p\text{-value} < 0.05$, модель залишає інформацію необробленою.

Тест Джарка–Бера для перевірки нормальності:

$$JB = \frac{n}{6} \left(S^2 + \frac{(K-3)^2}{4} \right) \quad (2.16)$$

де S — асиметрія (skewness),

K — ексцес (kurtosis).

Нормальність залишків — ключова умова для надійності інтервальних прогнозів.

2.9 Висновки до розділу 2

У другому розділі було здійснено огляд сучасних методів прогнозування демографічних процесів. Розглянуто ключові типи моделей, що застосовуються для роботи з часовими рядами: авторегресивні (AR), комбіновані ARIMA-моделі, методи експоненційного згладжування, логістичні моделі, а також holdout-підхід і сценарне прогнозування.

Особливу увагу приділено ARIMA-моделі як базовому інструменту для побудови демографічних прогнозів на основі наявних історичних даних. Також обґрунтовано доцільність використання альтернативних підходів для порівняння точності та поведінки прогнозу за різними умовами.

Окрім моделей, теоретично описано основні етапи підготовки даних до прогнозування: перевірку стаціонарності, диференціювання, логарифмування та інші трансформації, які можуть знадобитися для коректної роботи моделей. Також розглянуто критерії оцінки точності прогнозу — середню абсолютну відсоткову похибку (MAPE), середньоквадратичну похибку (MSE), корінь середньоквадратичної похибки (RMSE) та інформаційний критерій Акаїке (AIC). Для візуального аналізу розподілу похибок у літературі часто використовується boxplot — цей підхід теж було коротко охарактеризовано.

Отримані теоретичні висновки стали основою для формування практичного підходу до прогнозування показників народжуваності, смертності та міграції в Німеччині, що реалізується у наступному розділі.

РОЗДІЛ 3 РЕАЛІЗАЦІЯ МОДЕЛІ ТА АНАЛІЗ РЕЗУЛЬТАТІВ

3.1 Обґрунтування вибору країни, показників і методів

У практичній частині дослідження об'єктом аналізу обрано Федеративну Республіку Німеччина — країну, що протягом останніх десятиліть демонструє складну демографічну динаміку. Стійке зниження рівня народжуваності, зростання тривалості життя та значна залежність від міграційного приросту створюють численні виклики для соціально-економічного розвитку. Тому Німеччина є показовим прикладом для аналізу складних взаємозв'язків між демографічними показниками та моделювання майбутніх змін.

Для реалізації прогностичної моделі було виокремлено три ключові компоненти демографічної динаміки: сумарний коефіцієнт народжуваності (TFR), як показник інтенсивності відтворення населення; нетто-міграцію, яка виступає компенсаційним механізмом природного скорочення; очікувану тривалість життя при народженні, як індикатор стану смертності.

Ці показники є основою когортно-компонентного підходу до прогнозування чисельності населення, оскільки вони безпосередньо визначають зміну демографічної піраміди у часі. Для кожного з них було підібрано відповідний інструментарій прогнозування: ARIMA-моделі, експоненційне згладжування, логістичне наближення, а також сценарний аналіз для оцінки альтернативних варіантів розвитку.

3.2 Вибір моделей і змінних

У практичній частині дослідження передбачається побудова окремих моделей прогнозування для кожного з трьох ключових демографічних показників: сумарного коефіцієнта народжуваності (TFR), очікуваної тривалості життя при народженні (e_x) та нетто-міграції (net migration). Вибір цих змінних обґрунтований тим, що вони є основними компонентами демографічного балансу і безпосередньо визначають чисельність населення в середньостроковій перспективі.

Для моделювання TFR буде використано ARIMA-модель, яка дозволяє враховувати трендові властивості та автокореляційну структуру ряду. Такий підхід є доцільним для змінної, що має обмежену сезонність і демонструє інерційність змін. Параметри ARIMA будуть підібрані автоматично за допомогою інформаційних критеріїв (AIC, BIC), а також з урахуванням тестів на автокореляцію залишків.

Очікувана тривалість життя буде прогнозуватись за допомогою логістичної моделі, яка дозволяє моделювати процеси з асимптотичним наближенням до межі. Така форма функції добре відображає реальну динаміку зростання e_x у розвинених країнах. Крім того, для цього показника буде застосована Вох–Сох трансформація, після чого здійсниться побудова альтернативної ARIMA-моделі, що дасть змогу порівняти якість обох підходів.

Для ряду нетто-міграції також буде використана ARIMA-модель, оскільки цей показник має нестабільну, але трендову природу і складно піддається моделюванню за допомогою структурних рівнянь. Перед побудовою моделі буде здійснено тест на стаціонарність, після чого обрано конфігурацію ARIMA з найкращими показниками точності.

3.3 Джерела даних

Для побудови моделей прогнозування було використано офіційні відкриті статистичні джерела, що містять репрезентативні часові ряди демографічних показників для Німеччини.

Сумарний коефіцієнт народжуваності (TFR) отримано з бази даних Світового банку, яка охоплює щорічні значення цього показника з 1960 року для більшості країн світу [14]. У межах дослідження використано дані за період з 1990 по 2023 рік.

Інформацію про очікувану тривалість життя при народженні (e_x) взято з бази Human Mortality Database [15]. Використано таблицю смертності для Німеччини у форматі 1×1 , з якої було обрано дані лише для віку 0. Досліджуваний часовий діапазон становив 1990–2023 роки; значення змінної було очищено від пропусків для забезпечення цілісності ряду.

Нетто-міграційне сальдо (net migration) також отримано з офіційного ресурсу Світового банку [13], де наведено щорічні оцінки різниці між кількістю іммігрантів та емігрантів у форматі «осіб на рік». Для моделі обрано дані по Німеччині, починаючи з 1990 року.

Перед початком моделювання всі набори даних було перевірено на повноту, очищено від пропущених значень і приведено до єдиного числового формату для подальшого використання у програмному середовищі Python.

3.4. Прогнозування сумарного коефіцієнта народжуваності (TFR)

Сумарний коефіцієнт народжуваності (TFR) визначає середню кількість дітей, яку народжує одна жінка протягом репродуктивного періоду життя. З

метою попереднього аналізу було побудовано графік часової динаміки показника за період 1990–2023 рр., представлений на рисунку 3.1.



Рисунок 3.1 – Динаміка TFR у Німеччині

Як можна побачити на рисунку 3.1, у 1990-х роках спостерігається зниження рівня народжуваності, після чого з 2007 року відбулося часткове відновлення показника. Утім, загальний рівень TFR протягом усього періоду залишався нижчим за 1.6, що свідчить про стабільно низьку фертильність та потенційну нестационарність ряду.

Для перевірки стаціонарності проведено тест Дікі–Фуллера. Результати показали значення $p\text{-value} \approx 0.21$, що суттєво перевищує критичний рівень значущості 0.05. Таким чином, нульова гіпотеза про наявність одиничного кореня не була відхилена, що підтверджує нестационарний характер ряду.

З метою забезпечення коректного прогнозування застосовано перше диференціювання, тобто порядок інтеграції встановлено як $d = 1$.

Для моделювання було обрано клас моделей ARIMA (AutoRegressive Integrated Moving Average), які є стандартним інструментом для обробки трендових нестационарних рядів.

Далі було реалізовано повний перебір ARIMA($p,1,q$) з параметрами $p=0\dots$ і $q=0\dots5$, а також оцінено моделі за такими критеріями:

AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion), MAE, RMSE, MAPE, Durbin–Watson (DW), Ljung–Box.

У таблиці 3.1 представлено порівняння кількох моделей, які показали найкращі результати. Хоча відмінності між ними не є суттєвими, модель ARIMA(2,1,1) забезпечила збалансовані показники точності, стабільності та простоти. Вона продемонструвала низьке значення похибок ($MAE \approx 0.086$; $RMSE \approx 0.24$), відсутність автокореляції залишків (Ljung–Box $p > 0.05$) та прийнятні значення інформаційних критеріїв ($AIC = -104.69$, $BIC = -98.71$).

Таблиця 3.1 – Метрики адекватності ARIMA-моделей прогнозу TFR у Німеччині

Модель	AIC	BIC	DW	MAE	Ljung-P
ARIMA(2,1,1)	-104.69	-98.71	1.2	0.0866	1.0
ARIMA(1,1,5)	-105.81	-95.33	1.192	0.089	1.0
ARIMA(2,1,3)	-104.71	-95.74	1.194	0.0885	1.0

Після оцінки параметрів модель була використана для побудови прогнозу TFR на період 2024–2035 років на рисунку 3.2.

На графіку нижче наведено як фактичні, так і прогнозні значення. Видно, що модель передбачає стабілізацію показника на рівні ~ 1.35 , без суттєвого зростання.



Рисунок 3.2 – Прогноз TFR у Німеччині за моделлю ARIMA(2,1,1)

Для перевірки припущення нормальності залишків моделі ARIMA(2,1,1) було побудовано гістограму та Q–Q графік, представлені на рисунку 3.3.

На гістограмі помітна асиметрія розподілу, що частково пояснюється структурним зломом у динаміці TFR після 2020 року.

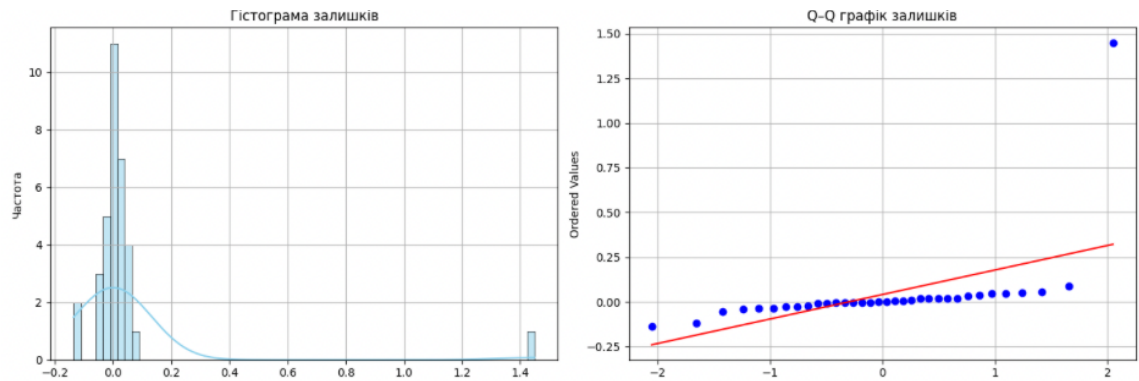


Рисунок 3.3 – Перевірка нормальності залишків моделі ARIMA(2,1,1)

Хоча тест Джарке-Бера вказує на статистично значуще відхилення від нормального розподілу ($p\text{-value} < 0.05$), тест Льюнга–Бокса ($p = 1.000$) не виявив автокореляції в залишках, що дозволяє вважати модель адекватною для прогнозування.

Детальні прогностні значення наведено на рисунку 3.4

Рік	
2024	1.364
2025	1.353
2026	1.349
2027	1.347
2028	1.347
2029	1.347
2030	1.346
2031	1.346
2032	1.346
2033	1.346
2034	1.346
2035	1.346

Рисунок 3.4 – Прогнозовані значення TFR

3.5. Прогнозування нетто-міграції населення Німеччини до 2035 року

З метою побудови реалістичного та статистично обґрунтованого прогнозу нетто-міграції населення Німеччини до 2035 року було застосовано комбінований підхід, який об'єднує оцінку тренду та моделювання залишків за допомогою ARIMA-моделі.

На початку було перевірено стаціонарність часового ряду NetMigration за допомогою тесту Дікі-Фуллера (ADF). Отримане р-значення становило 0.0219, що нижче за критичні значення на рівнях 1%, 5% і 10%. Це дало підстави відхилити нульову гіпотезу про наявність одиничного кореня, а отже вважати ряд стаціонарним. Тому було прийнято рішення не виконувати диференціювання ($d = 0$) при побудові ARIMA-моделі.

Для побудови трендової компоненти було обрано інтервал 2010–2023 років, як найбільш репрезентативний для поточної фази демографічної динаміки. Було реалізовано лінійну регресійну модель (3.1) у вигляді:

$$T_t = \beta_0 + \beta_1 t \quad (3.1)$$

Крім того, для кращого згладжування короткострокових коливань було застосовано фільтр Савицького–Голяя (Savitzky–Golay) (3.2) з параметрами `window_length=5`, `polyorder=2`. Це дозволило сформулювати згладжений тренд \hat{T}_t , на основі якого було обчислено залишки:

$$e_t = Y_t - \hat{T}_t \quad (3.2)$$

Для прогнозування нетто-міграції у Німеччині було використано поєднання лінійної регресії (тренд на основі даних 2010–2023 років) та моделі ARIMA(1,0,1), застосованої до залишків після згладження трендової компоненти. Такий гібридний підхід дозволяє врахувати як структурні зміни у часі, так і випадкові відхилення.

Оцінка моделі залишків здійснювалася за допомогою кількох критеріїв якості, зокрема Durbin–Watson, Ljung–Box та тесту нормальності D'Agostino,

а також стандартних показників MAE та RMSE. Результати наведено в таблиці 3.2.

Таблиця 3.2 – Метрики адекватності моделі ARIMA(1,0,1), побудованої на залишках трендової компоненти нетто-міграції в Німеччині.

DW	Ljung-Box	D'Agostino	MAE	RMSE
2.27	0.6125	0.8089	64.402	83.823

Отримані результати свідчать про статистичну адекватність побудованої моделі ARIMA(1,0,1), застосованої до залишків трендової компоненти. Значення коефіцієнта Дарбіна–Уотсона ($DW = 2.27$) вказує на відсутність автокореляції у залишках. Ljung–Box тест також не виявив статистично значущої автокореляції ($p = 0.6125 > 0.05$), а тест D'Agostino ($p = 0.8089$) підтвердив наближеність розподілу залишків до нормального. Значення MAE ≈ 64 тис. та RMSE ≈ 84 тис. є прийнятними з огляду на масштаби зміни нетто-міграції в Німеччині (з середнім значенням понад 300 тис.), що підтверджує задовільну точність моделі для подальшого використання у прогнозі.

На рисунку 3.5 представлено фінальний прогноз нетто-міграції в Німеччині до 2035 року, сформований шляхом поєднання трендового компоненту та залишків моделі ARIMA(1,0,1), симульованих із заданим сидом.

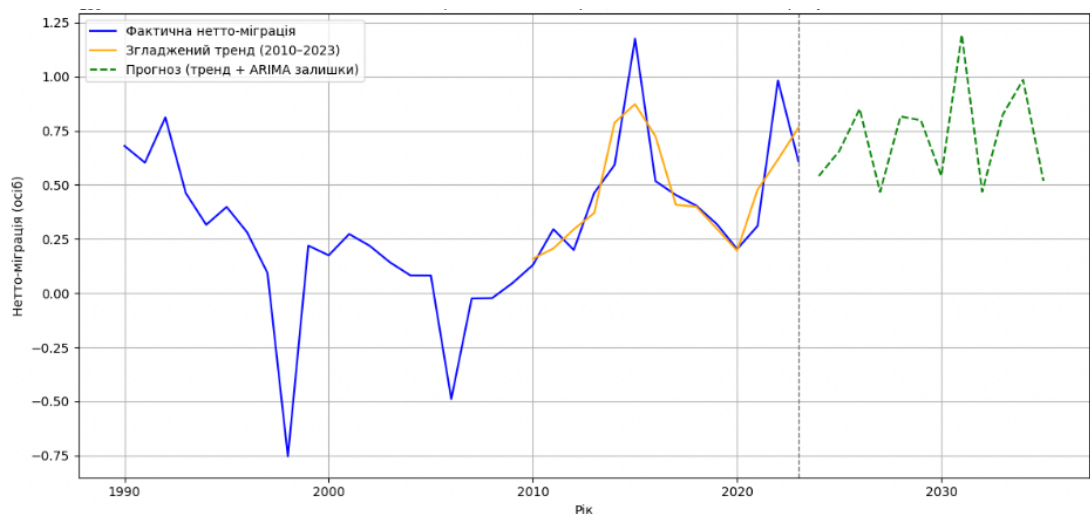


Рисунок 3.5 – Прогноз нетто-міграції у Німеччині

На графіках автокореляції (ACF), зображених на рисунку 3.6, та часткової автокореляції (PACF) зображених на рисунку 3.7, значення автокореляцій перебувають у межах довірчих інтервалів. Це означає відсутність статистично значущої автокорельованості на лагах 1–6, що підтверджує, що модель не пропустила важливі лаги і не потребує донастроювання.

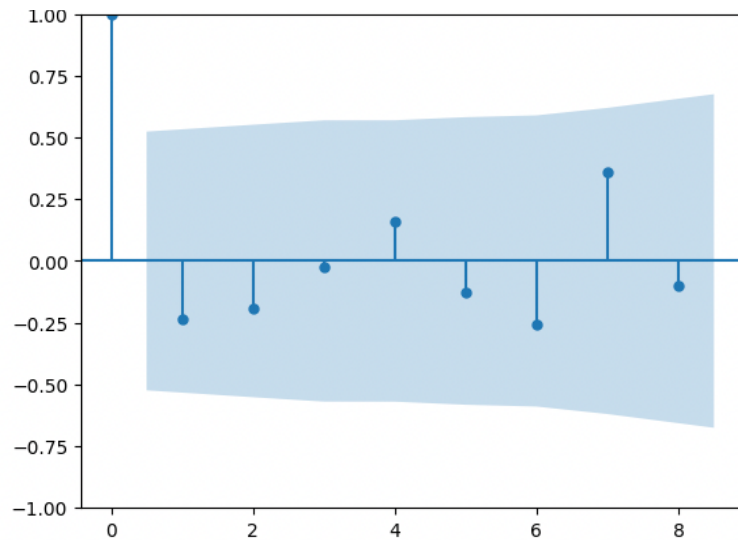


Рисунок 3.6 – ACF залишків моделі ARIMA(1,0,1)

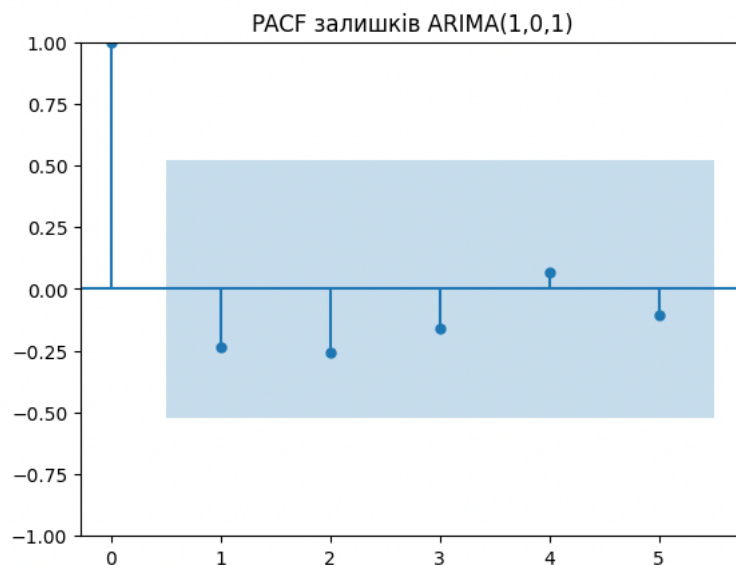


Рисунок 3.7 – PACF залишків моделі ARIMA(1,0,1)

На гістограмі залишків на рисунку 3.8 видно симетричний розподіл, центрований біля нуля, без різких піків або зміщення. Розподіл не має викидів, що також узгоджується з нормальністю.

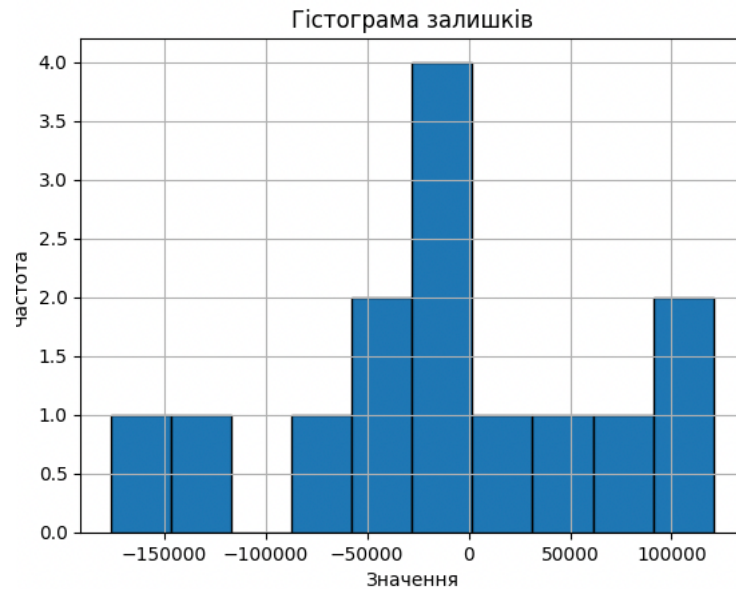


Рисунок 3.8 – Гістограма залишків

На Q–Q графіку на рисунку 3.9 спостерігається, що більшість точок лежать близько до діагональної прямої. Це вказує на узгодження емпіричного розподілу залишків з нормальним розподілом, що підтверджує припущення про нормальність залишків моделі. Невеликі відхилення на краях графіка можуть свідчити про незначні відхилення від нормальності, однак загалом модель відповідає критеріям адекватності для подальшого використання в прогнозуванні.

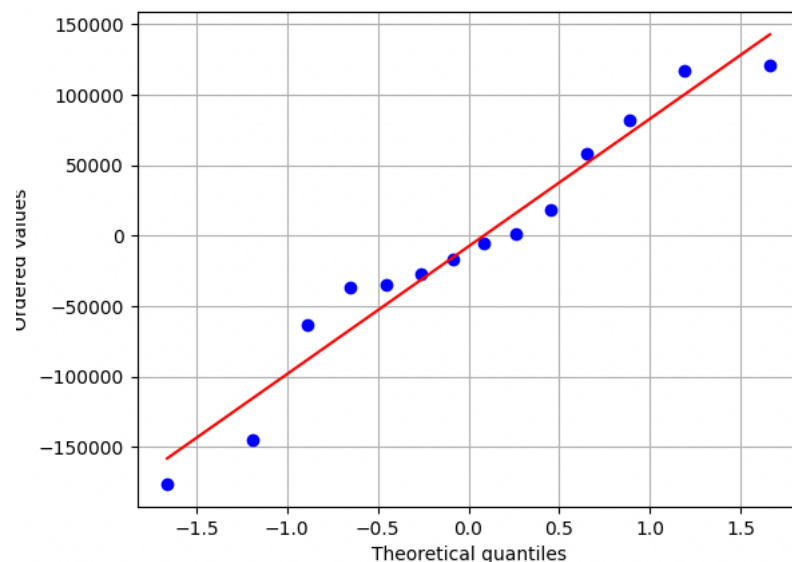


Рисунок 3.9 – Q–Q графік

На основі проведених перевірок можна зробити висновок, що побудована модель $ARIMA(1,0,1)$ є статистично адекватною для моделювання залишків тренду нетто-міграції. Вона успішно враховує часову структуру даних та забезпечує залишки, які не мають автокореляції, відповідають нормальному розподілу та можуть використовуватись для подальшої генерації прогнозу.

3.6. Сценарне прогнозування міграції

Для побудови прогнозу нетто-міграції на період 2024–2035 років також було застосовано сценарний підхід. Прогноз сформовано у вигляді трьох варіантів: базового, оптимістичного та песимістичного.

У кожному сценарії майбутні значення нетто-міграції обчислювались за експоненційною моделлю:

$$M_t = M_{2023}(1 + g)^t + \varepsilon_t \quad (3.3)$$

де g — щорічна зміна залежно від сценарію (0.003 для базового, 0.01 для оптимістичного, -0.01 для песимістичного);

$\varepsilon_t \sim N(0, \sigma^2)$ — випадкова флуктуація (білий шум) із фіксованою дисперсією.

Параметри шуму задано наступним чином: базовий: $\varepsilon_t \sim N(0, 12000^2)$; оптимістичний: $\varepsilon_t \sim N(0, 8000^2)$; песимістичний: $\varepsilon_t \sim N(0, 8000^2)$.

На рисунку 3.10 наведено результати сценарного прогнозування нетто-міграції на період 2024–2035 років. Фактичні значення за 1990–2023 рр. використано як база для розрахунку трендів та флуктуацій.



Рисунок 3.10– Сценарне прогнозування нетто-міграції

Базовий сценарій демонструє стабільне помірне зростання міграційних потоків, оптимістичний — більш динамічне зростання, тоді як песимістичний сценарій моделює поступове зниження показника. Візуалізація підтверджує значну варіативність прогнозу залежно від припущень щодо темпів приросту, що свідчить про необхідність врахування сценарного аналізу у довгостроковому демографічному плануванні.

3.7. Прогнозування очікуваної тривалості життя

Очікувана тривалість життя при народженні (e_x) є критично важливим показником, що відображає смертність населення і безпосередньо впливає на демографічний баланс. У роботі розглянуто дві моделі прогнозування e_x : ARIMA(2,2,2) та логістичну регресію.

Для обґрунтування вибору моделі ARIMA(2,2,2) було проведено порівняння декількох конфігурацій ARIMA(p,d,q), результати якого наведено в таблиці 3.6а. Оцінювання здійснювалося за основними критеріями: AIC, BIC, середня абсолютна похибка (MAE), коефіцієнт Дарбіна–Уотсона (DW) для залишків, а також Ljung–Box тест на автокореляцію.

Модель ARIMA(2,2,2) має одне з найкращих значень AIC і стабільні залишки ($DW = 1.998$, Ljung–Box $p = 0.98$), що вказує на відсутність автокореляції. Хоча деякі моделі з нижчим порядком диференціювання (наприклад, ARIMA(1,1,0) та ARIMA(3,1,2)) мають менші значення MAE, вони демонструють значну автокореляцію залишків ($DW \approx 0.99$, Ljung–Box $p < 0.65$), що ставить під сумнів їх адекватність. Модель ARIMA(0,2,0) виявилася найменш придатною через дуже високі значення AIC та BIC.

Отже, за сукупністю статистичних показників у таблиці 3.3, модель ARIMA(2,2,2) була обрана як базова для побудови прогнозу очікуваної тривалості життя.

Таблиця 3.3 – Порівняльні метрики моделей ARIMA(p,d,q) для прогнозу очікуваної тривалості життя в Німеччині

Модель	AIC	BIC	DW	MAE	Ljung-P
ARIMA(2,2,2)	-12.43	-5.6	1.998	3.76	0.98
ARIMA(1,2,1)	-13.81	-9.07	1.192	3.77	0.98
ARIMA(0,2,0)	16.71	17.9	1.98	3.96	0.95

Та також було перевірено на історичних даних, модель показала велику точність, зображено на рисунку 3.11, а саме MAE = 0.14 років, MAPE = 0.18%.

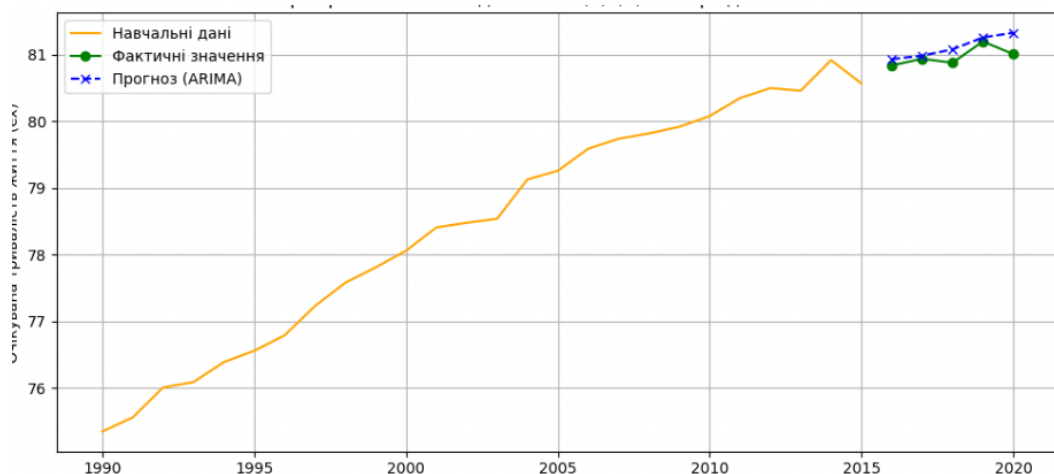


Рисунок 3.11 – Out-of-sample перевірка на періоді 2016–2020

Водночас було протестовано логістичну модель, що має вигляд:

$$ex(t) = \frac{L}{1 + \exp(-k(t - t_0))} \quad (3.4)$$

де $L=85.3$, $k=0.057$, $t_0=2003$ — параметри, підібрані методом найменших квадратів. Логістичний прогноз демонструє плавне наближення ex до верхньої межі виживаності.

Як видно з таблиці 3.6, логістична модель демонструє трохи кращу точність за основними метриками ($MAE = 0.12$; $RMSE = 0.15$; $MAPE = 0.16\%$), а також має дуже високе значення $R^2 = 0.9931$, що свідчить про високий рівень апроксимації емпіричних даних. Однак $ARIMA(2,2,2)$ забезпечує високу якість залишків ($DW = 1.998$, $Ljung\text{--}Box\ p = 0.98$) та має низьке значення AIC , що робить її придатною для коротко- та середньострокового прогнозування.

Таблиця 3.4 – Порівняння метрик якості моделей прогнозу очікуваної тривалості життя в Німеччині

Модель	MAE (років)	MAPE	RMSE	R^2	DW	Jarque- bera p
ARIMA	0.14	0.18	0.24	-	1.99	0.78
Logistic	0.12	0.16	0.15	0.99	-	0.70

Отже на рисунку 3.12 було побудовано логістичну модель та $ARIMA(2,2,2)$.

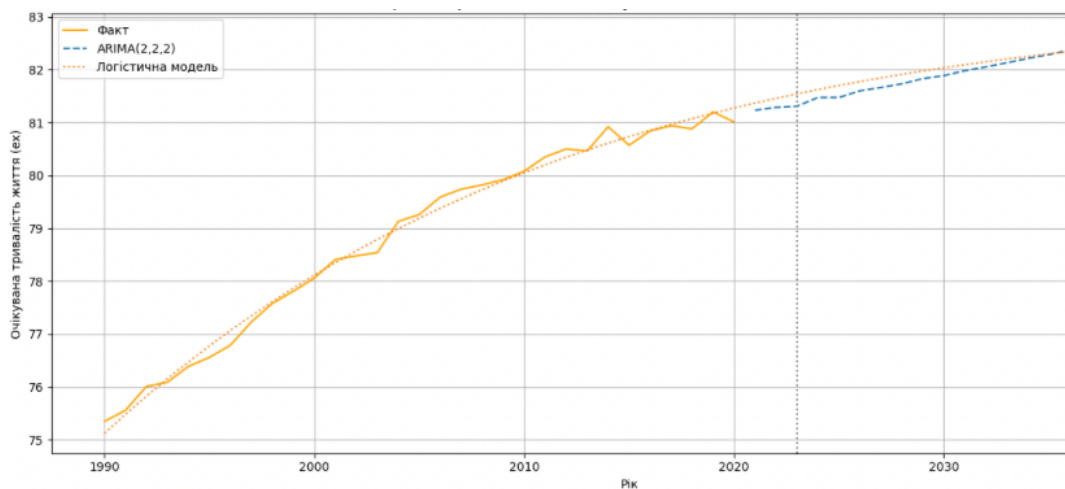


Рисунок 3.12 – Побудований прогноз очікуваної тривалості життя

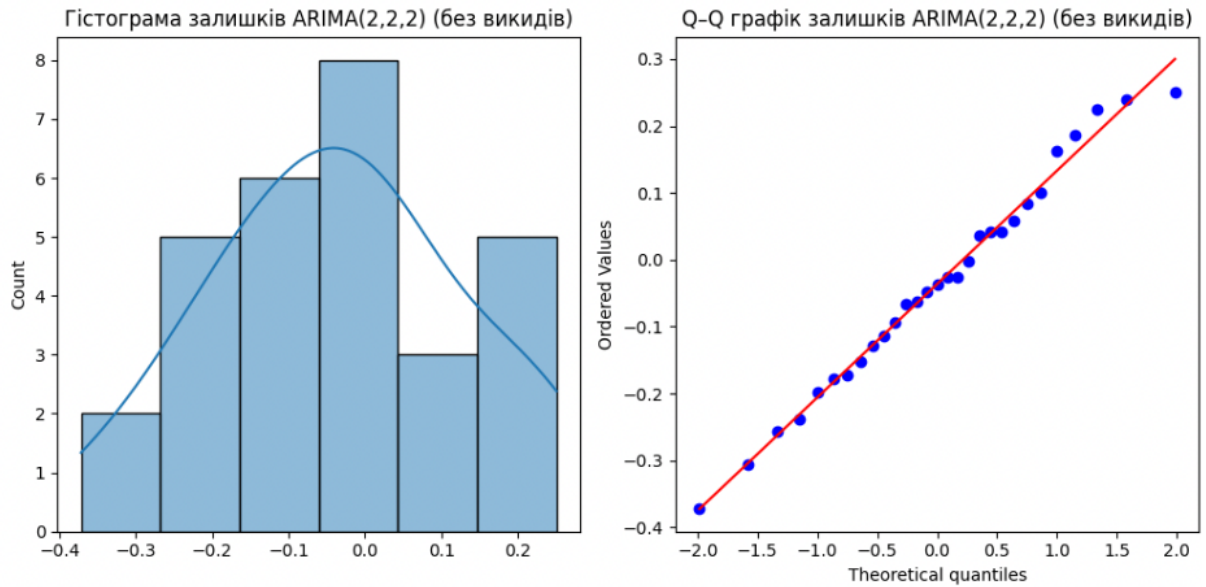
Після побудови моделі ARIMA було виявлено, що початкові залишки мають суттєві викиди. Це є типовим явищем для авторегресивних моделей, зумовленим обмеженою кількістю лагів на перших етапах прогнозу. Залишки всієї моделі ARIMA мали високу асиметрію, що підтверджується результатом тесту Джарка–Бера: статистика становила 444.09, а р-значення було близьким до нуля, що вказує на відхилення від нормального розподілу.

З метою підвищення достовірності моделі було здійснено виключення перших двох залишкових значень, які мали найбільші відхилення. Після цього перевірка на нормальність залишків підтвердила відповідність розподілу: значення статистики Jarque–Bera становило 0.49 при р-значенні 0.78. Отже, модель ARIMA(2,2,2) після корекції може вважатися статистично прийнятною. Такий результат свідчить про те, що припущення про нормальність залишків не порушено, а отже — модель придатна для подальшого використання у прогнозуванні. Це підтверджує коректність застосованого підходу до моделювання та дозволяє вважати отримані прогнозні значення статистично обґрунтованими. Відсутність істотних відхилень від нормального розподілу також підвищує достовірність оцінювання параметрів моделі.

Додатково візуальний аналіз Q–Q графіка та гістограми залишків не виявив суттєвих асиметрій чи перекосів, а автокореляційна функція (ACF) залишків не демонструє значущих кореляцій. Це підтверджує, що модель добре відображає структуру даних і не залишає систематичних помилок у залишках, що могло б свідчити про недоопрацьовану специфікацію моделі.

У паралельному розрахунку було побудовано логістичну модель, що описує довгострокове зростання тривалості життя як S-подібну криву з асимптотичним наближенням до межі. Залишки логістичної моделі були також перевірені на нормальність. За результатами тесту Джарка–Бера ($stat = 0.71$, $p = 0.70$) було встановлено, що вони не відрізняються від нормального

розподілу. Візуальний аналіз гістограми та Q–Q графіку зображений на рисунку 3.13 підтвердив відсутність систематичних відхилень.



Jarque–Bera (ARIMA без викидів): stat = 0.49, p = 7.83e-01

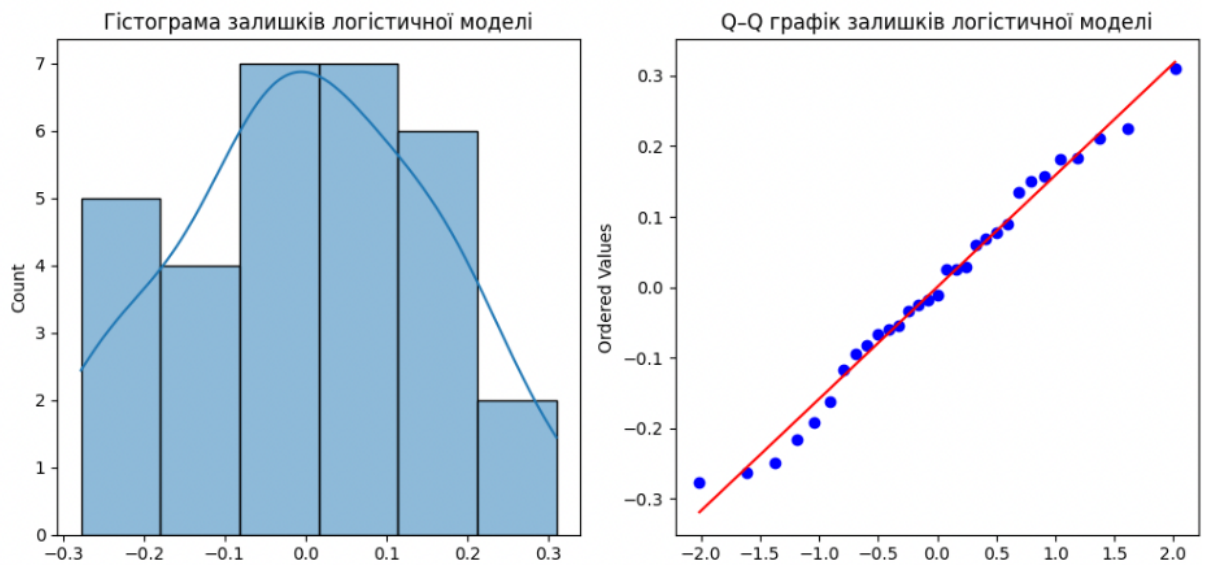


Рисунок 3.13 – Аналіз залишків

3.8 Сценарне прогнозування чисельності населення Німеччини до 2035 року

На завершальному етапі дослідження було здійснено комплексне прогнозування чисельності населення Німеччини до 2035 року з урахуванням

попередньо побудованих прогнозів ключових демографічних компонентів — народжуваності (TFR), смертності (через очікувану тривалість життя) та нетто-міграції. Розрахунки виконувалися шляхом покрокового оновлення чисельності населення на кожен рік прогнозного періоду за формулою (3.5):

$$P_{t+1} = P_t + B_t - D_t + M_t \quad (3.5)$$

де

P_t — чисельність населення у році t ;

B_t — кількість народжених;

D — кількість померлих;

M_t — нетто-міграція.

Початкове значення P_{2023} становило 83.2 млн осіб, відповідно до актуальних даних Федерального статистичного відомства Німеччини.

Кількість народжених B_t розраховувалась за формулою (3.6):

$$B_t = \frac{TFR_t * P_t * s}{1000} \quad (3.6)$$

де

$s=0.23$ — частка жінок репродуктивного віку у структурі населення.

Кількість смертей t визначалася на основі умовного коефіцієнта смертності, що лінійно залежить від прогнозованої тривалості життя за формулою (3.7):

$$\text{mortality_rate} = 0.015 - 0.00005 \cdot ex \quad (3.7)$$

Окремо було побудовано два варіанти смертності: на основі моделі ARIMA (основний сценарій) та на основі логістичної моделі (альтернативний сценарій), що дозволило оцінити чутливість загального прогнозу до моделі тривалості життя. Для уникнення впливу викидів перші два значення прогнозу ex в ARIMA були замінені стабілізованим значенням третьої точки.

На основі цієї моделі було реалізовано шість сценаріїв демографічного розвитку:

1) базовий — використання основних прогнозів TFR, міграції та смертності (ARIMA).

2) оптимістичний — збільшення міграції на 50% порівняно з базовим варіантом.

3) песимістичний — зменшення міграції на 50%.

4) оптимістичний (850 тис.) — фіксація нетто-міграції на рівні 850 тис. осіб щороку.

5) Policy reform — одночасне підвищення TFR (до 1.5) та міграції (+30%).

6) Базовий (логістична ex) — заміна ARIMA-прогнозу тривалості життя на логістичний.

На рисунку 3.14 зображено результати моделювання чисельності населення Німеччини.

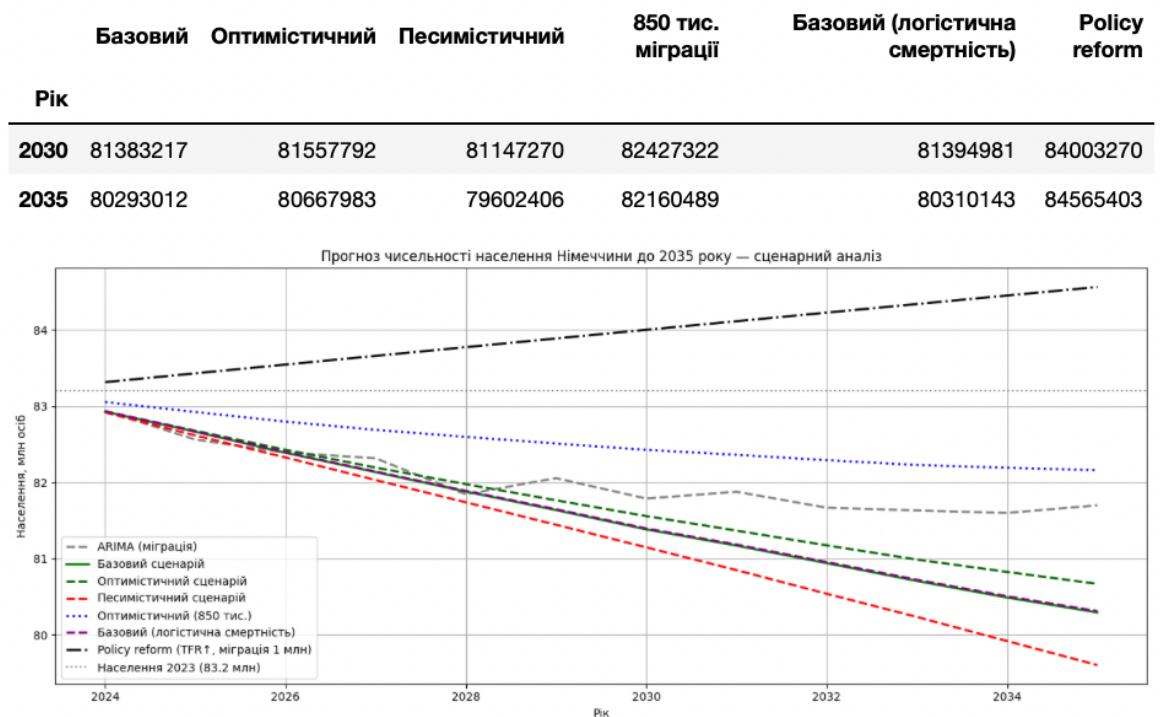


Рисунок 3.14 – Результати прогнозування

Результати моделювання демонструють, що у відсутності активних політичних заходів населення Німеччини зменшуватиметься. У базовому сценарії прогнозується зниження чисельності населення з 83,2 млн у 2023 році до 80,3 млн у 2035 році, тобто приблизно на 2,9 млн осіб. У песимістичному

сценарії скорочення є ще значнішим — до 79,6 млн, що відповідає зменшенню майже на 3,6 млн осіб.

Натомість сценарій Policy reform, який поєднує зростання сумарного коефіцієнта народжуваності до 1,65 та щорічну міграцію на рівні 1 млн осіб, демонструє незначне, але стабільне зростання населення до 84,6 млн у 2035 році. Таким чином, лише цей сценарій дозволяє компенсувати природне скорочення населення за рахунок ефективної демографічної політики.

Сценарії з підвищеною міграцією (до 850 тис. осіб) або з альтернативною моделлю смертності (логістична функція) не забезпечують повної стабілізації, але уповільнюють спад порівняно з базовим сценарієм. Це підтверджує критичну залежність демографічного прогнозу від політичних рішень у сфері народжуваності та імміграції.

3.9 Висновки до розділу 3

У третьому розділі було реалізовано практичну частину дослідження, спрямовану на моделювання та прогнозування демографічних змін у Федеративній Республіці Німеччина з урахуванням впливу міграційних процесів. Проведено статистичний аналіз динаміки основних демографічних показників, таких як сумарний коефіцієнт народжуваності, тривалість життя та нетто-міграція. Для кожного з показників були побудовані прогнози за допомогою відповідних моделей: ARIMA, експоненційного згладжування, логістичної регресії та інших.

Окрему увагу приділено сценарному підходу до прогнозування міграції. Зокрема, було сформовано чотири сценарії розвитку: базовий, оптимістичний, песимістичний та модельний прогноз на основі ARIMA. На основі отриманих прогнозних значень сформовано узагальнену таблицю параметрів

(final_input_forecast), яка стала основою для подальших розрахунків очікуваної чисельності населення Німеччини до 2035 року.

Результати дослідження показали, що міграційний чинник відіграє ключову роль у компенсації природного скорочення населення. У сценаріях зі стабільним або позитивним міграційним сальдо зменшення чисельності населення уповільнюється, натомість за умов низької народжуваності та зменшення міграції спостерігається стійка тенденція до депопуляції.

Таким чином, у межах практичного розділу було продемонстровано, що поєднання сучасних інструментів аналізу часового ряду та сценарного моделювання дає змогу сформулювати обґрунтований прогноз демографічних змін. Отримані результати можуть бути використані для оцінки можливих стратегій демографічної та міграційної політики на національному рівні.

РОЗДІЛ 4 ФУНКЦІОНАЛЬНО-ВАРТІСНИЙ АНАЛІЗ ПРОГРАМНОГО ПРОДУКТУ

У цьому розділі побудована система прогнозування демографічних змін у Німеччині з урахуванням міграційних процесів розглядається як програмний продукт, що поєднує в собі не лише технічну реалізацію, а й потребує економічного обґрунтування. Результати реалізованої моделі можуть бути використані для підтримки прийняття рішень у сфері демографічної політики, наукових досліджень, а також як програмний інструмент для державних і муніципальних установ, що займаються питаннями планування населення.

У даному випадку як метод оцінювання використовується функціонально-вартісний аналіз (ФВА), який дозволяє оцінити ефективність реалізації функцій системи, визначити критичні технічні параметри, оцінити витрати на розробку і знайти оптимальне рішення з огляду на задані обмеження. Методика ФВА широко застосовується для вибору найкращого варіанту реалізації програмного забезпечення з урахуванням співвідношення функціональності, трудомісткості та вартості.

ФВА передбачає аналіз основних функцій продукту, диференціацію варіантів їх реалізації, порівняння з альтернативами та оцінку якості технічних характеристик. Крім того, в процесі аналізу розглядаються ключові параметри (точність моделі, витрати часу, складність реалізації), проводиться експертне оцінювання та економічний аналіз витрат на створення та використання системи.

Таким чином, застосування функціонально-вартісного аналізу дає змогу комплексно оцінити створений продукт не лише з точки зору його математичної точності, а й за критеріями вартості, ефективності та доцільності впровадження.

4.1 Постановка задачі проектування

У роботі застосовується метод ФВА для проведення техніко-економічного аналізу розробки системи прогнозу стійкості фінансових показників. Оскільки рішення стосовно проектування та реалізації компонентів, що розробляється, впливають на всю систему, кожна окрема підсистема має її задовольняти. Тому фактичний аналіз представляє собою аналіз функцій програмного продукту, призначеного для збору, обробки та проведення аналізу даних по компанії.

Технічні вимоги до програмного продукту включають п'ять аспектів.

1. Функціонування на персональних комп'ютерах із стандартним набором компонентів.
2. Зручність та зрозумілість для користувача.
3. Швидкість обробки даних та доступ до інформації в реальному часі.
4. Можливість зручного масштабування та обслуговування.
5. Мінімальні витрати на впровадження програмного продукту.

4.2 Обґрунтування функцій програмного продукту

Головна функція F_0 полягає у створенні програмного продукту, який реалізує прогноз демографічних змін у Німеччині з урахуванням динаміки народжуваності, смертності та міграційних процесів. Такий інструмент має забезпечувати точність прогнозу, адаптивність до зміни вхідних сценаріїв, ефективність реалізації в обчислювальному середовищі та зручність для подальшого аналізу результатів.

У межах цієї функції виділяються три ключові підфункції.

1. F_1 — вибір мови програмування та середовища реалізації;

2.F₂ — вибір методу прогнозування демографічних показників;

3.F₃ — вибір підходу до аналізу та оцінки результатів прогнозу.

Кожна функція має декілька варіантів реалізації, що відповідають її специфіці.

F₁ — Мова програмування та середовище:

1) Python – мова з відкритим кодом, яка активно використовується в наукових дослідженнях;

2) R – статистичне середовище з вбудованими демографічними пакетами;

F₂ — Метод прогнозування:

1) ARIMA – авторегресійна модель, що враховує тренди і лаги (реалізовано у роботі);

2) сценарне прогнозування – побудова варіантів базового, оптимістичного та песимістичного сценаріїв;

3) лінійна регресія – класичний статистичний підхід, часто використовується як базова модель.

F₃ — Підхід до аналізу результатів:

1) порівняння з фактичними демографічними даними за попередні роки;

2) порівняння з офіційними прогнозами міжнародних організацій (UN WPP, Destatis);

3) Порівняння з статистичним трендом

Варіанти реалізації основних функцій наведені у морфологічній карті системи (рис. 4.1).

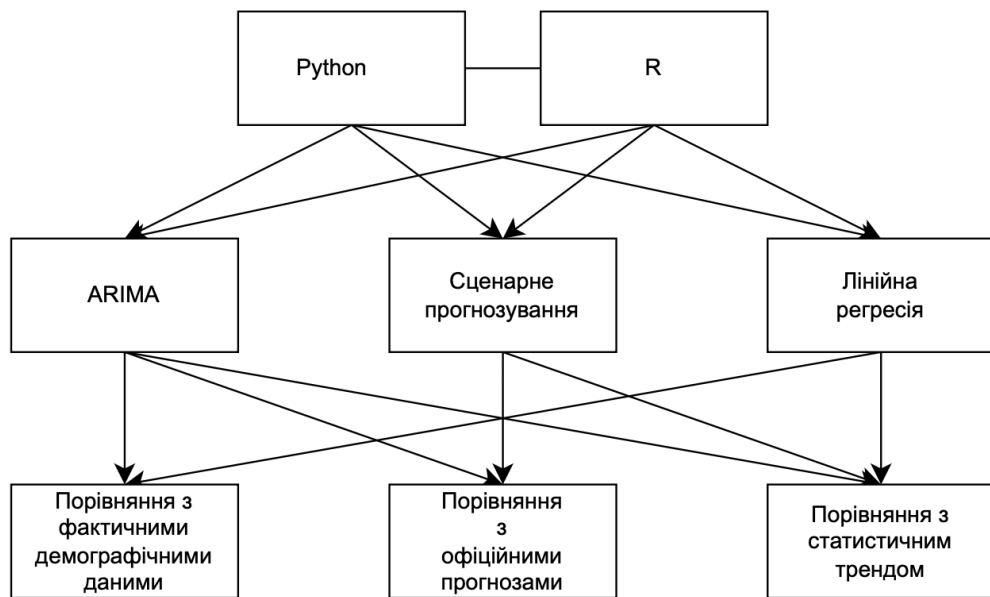


Рисунок 4.1 – Морфологічна карта

За допомогою морфологічної карти є можливість побудувати позитивно-негативну матрицю з відображенням усіх варіантів основних функцій(табл. 4.1).

Таблиця 4.1 – Позитивно-негативна матриця

Функції	Варіанти	Переваги	Недоліки
F_1	<i>A</i>	Багатий вибір бібліотек для часових рядів	Потребує налаштування середовища (Jupyter, залежності);
	<i>B</i>	Має спеціалізовані пакети для демографії сильна статистична основа	Слабша інтеграція з іншими мовами;

Продовження таблиці 4.1

F_2	<i>A</i>	Враховує тренди та автокореляцію; дозволяє робити точні коротко- та середньострокові прогнози;	Потребує перевірки залишків та стаціонарності; не враховує зовнішні впливи
	<i>B</i>	Дозволяє змоделювати різні варіанти розвитку, легко інтерпретується	Не має чіткої математичної моделі; залежить від припущень дослідника
	<i>B</i>	Дуже проста у реалізації; може використовуватись як базова модель або для валідації	Не враховує циклічність, лаги чи зовнішні впливи; слабка прогностична здатність у демографії
F_3	<i>A</i>	Дозволяє об'єктивно оцінити точність прогнозу; та зробити графічне порівняння	Обмежене лише доступними роками;
	<i>B</i>	Дає змогу співставити модель з визнаними джерелами; показує відповідність загальним трендам	Методи розрахунку офіційних прогнозів не завжди відомі; можливі відмінності в базових припущеннях
	<i>B</i>	Швидкий контроль на стабільність прогнозу; зручно візуалізується як додатковий фон	Не показує справжню точність моделі;

Згідно з аналізом позитивно-негативної матриці, частина варіантів реалізації функцій програмного продукту не відповідає практичним потребам моделювання демографічних процесів у Німеччині. Такі варіанти не були обрані для подальшого впровадження.

Функція F1. Серед запропонованих мов програмування перевага надається Python, оскільки він забезпечує широку підтримку демографічного аналізу завдяки бібліотекам для обробки даних, побудови моделей ARIMA, візуалізації.

Функція F2. Найбільш доцільними методами прогнозування є ARIMA (варіант а) та сценарний підхід (варіант б). Вони виконують різні завдання: ARIMA — для формального статистичного прогнозу, а сценарне моделювання — для побудови варіативних траєкторій розвитку. Лінійну регресію (варіант в) було розглянуто як допоміжний інструмент, проте вона не забезпечує достатньої точності.

Функція F3. Обидва підходи до оцінки результатів — порівняння з фактичними даними (а) та з офіційними прогнозами (б) — є доцільними. Проте перевага надається варіанту б, оскільки він дозволяє співставити результати з існуючими довгостроковими прогнозами міжнародних організацій.

Таким чином, розглядаємо два сценарії реалізації програмного продукту.

1. $F_{1a} - F_{2a} - F_{3б}$

2. $F_{1a} - F_{2б} - F_{3б}$

Оцінювання якості згаданих функцій виконується з допомогою параметрів нижче.

4.3 Обґрунтування системи параметрів програмного продукту

Враховуючи специфіку розробки програмного продукту для прогнозування демографічних змін у Німеччині з використанням методів ARIMA, сценарного моделювання та інструментів Python, було визначено низку ключових параметрів, що впливають на ефективність реалізації обраних підходів.

Для порівняння можливих реалізацій (мова програмування, метод прогнозування, тип аналізу результатів) були сформовані критерії, які дозволяють оцінити технічний рівень рішення. До таких критеріїв належать:

- 1.X1 – гіпотетичний об'єм коду програми;
- 2.X2 – час на попередню обробку та моделювання;
- 3.X3 – час, необхідний на виконання повного циклу аналізу;
- 4.X4 – точність моделі.

Параметри поділяються на гірші, середні і кращі та обираються замовником та диктуються умовами використання програмного продукту(табл. 4.2).

Таблиця 4.2 - Основні параметри програмного продукту

Назва параметра	Умовні позначення	Одиниці виміру	Значення параметра		
			гірші	середні	кращі
Гіпотетичний об'єм коду програми	X1	кількість рядків коду	1200	800	500
Час на попередню обробку та моделювання	X2	годин	10	5	2
Час необхідний на виконання циклу аналізу	X3	дні	5	3	1
Точність моделі	X4	відсотки	75	85	95

Дані в таблиці 4.3 дозволяють побудувати графічні характеристики для оцінки та порівняння параметрів (рис. 4.2 – 4.5).

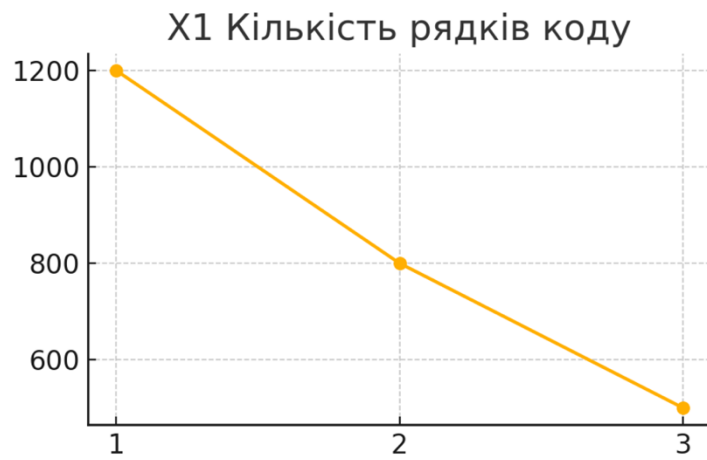


Рисунок 4.2 – X1, Гіпотетичний об'єм коду програми



Рисунок 4.3 – X2, Час на моделювання

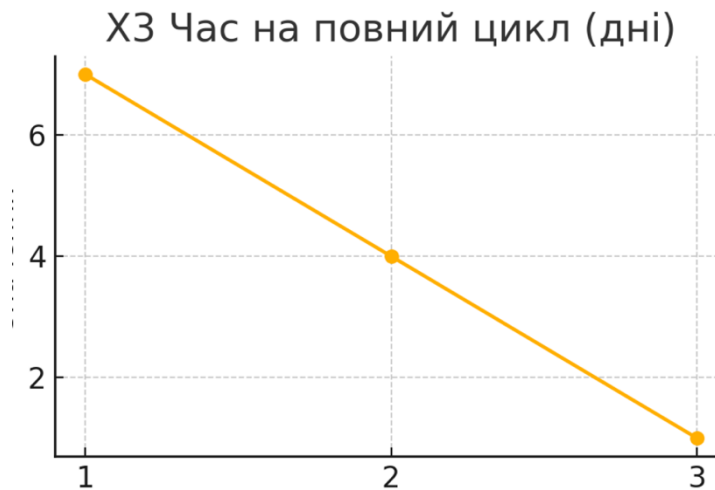


Рисунок 4.4 – X2, Час необхідний на виконання повного циклу аналізу

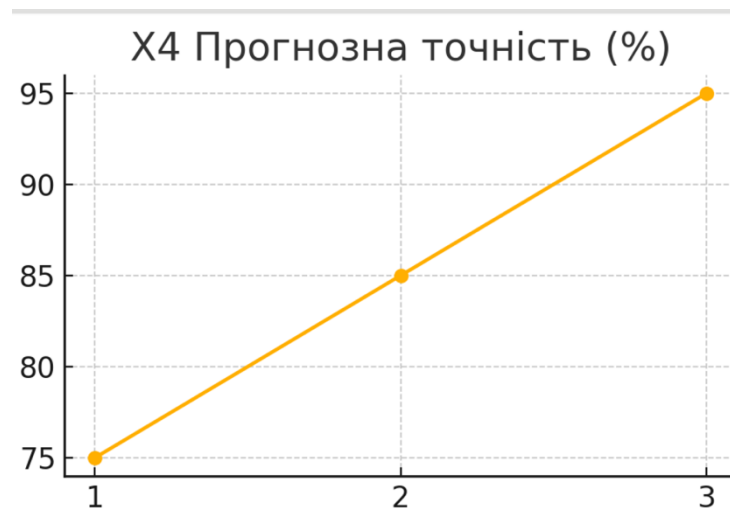


Рисунок 4.5 – X4, Точність моделі

4.4 Аналіз експертного оцінювання параметрів

Провівши ґрунтовне обговорення та комплексний аналіз, кожен експерт ретельно оцінює вагу кожного параметра у контексті поставленої мети - розробки програмного продукту. Цей продукт має забезпечувати максимальну точність при визначенні параметрів моделей адаптивного прогнозування та розрахунку прогнозованих значень.

Значимість кожного параметра визначається методом попарного порівняння. Оцінку проводить експертна комісія із 7 людей. Визначення коефіцієнтів значимості передбачає 4 етапи.

1. Визначення рівня значимості параметра шляхом присвоєння різних рангів.
2. Перевірку придатності експертних оцінок для подальшого використання.
3. Визначення оцінки попарного пріоритету параметрів.
4. Обробку результатів та визначення коефіцієнту значимості.

Результати рангової оцінки експертами наведені у таблиці 4.3.

Таблиця 4.3 - Результати ранжування параметрів

Позначення параметра	Назва параметра	Одиниці виміру	Ранг параметра за оцінкою експерта							Сума рангів R_i	Відхилення Δ_i	Δ_i^2
			1	2	3	4	5	6	7			
$X1$	Потенційний об'єм програмного коду	Кількість рядків	4	4	3	4	3	4	4	26	-8,5	72,25
$X2$	Час на попередню обробку та моделювання	години	3	2	4	3	4	3	3	22	-4,5	20,25
$X3$	Час на повний цикл аналізу	дні	2	3	2	1	1	2	1	12	5,5	30,25
$X4$	Прогнозна точність моделі	відсотки	1	1	1	2	2	1	2	10	7,5	56,25
	Разом		10	10	10	10	10	10	10	70	0	197

Для перевірки степені достовірності експертних оцінок, визначимо наступні параметри:

1) сума рангів кожного з параметрів і загальна сума рангів за формулою (4.1):

$$R_i = \sum_{j=1}^N r_{ij} R_{ij} = \frac{Nn(n+1)}{2} = 70, \quad (4.1)$$

де N – число експертів;

n – кількість параметрів;

2) середня сума рангів за формулою (4.2):

$$T = \frac{1}{n} R_{ij} = 17.5; \quad (4.2)$$

3) відхилення суми рангів кожного параметра від середньої суми рангів за формулою (4.3):

$$\Delta_i = R_i - T. \quad (4.3)$$

Сума відхилень по всіх параметрах повинна дорівнювати 0;

4) загальна сума квадратів відхилення за формулою (4.4):

$$S = \sum_{i=1}^N \Delta_i^2 = 179.0 \quad (4.4)$$

Коефіцієнт узгодженості:

$$W = \frac{12S}{N^2(n^3 - n)} = \frac{12 \cdot 179}{7^2(4^3 - 4)} = 0,731 > W_k = 0,56. \quad (4.5)$$

Отримані результати ранжування можна вважати надійними, адже розрахований коефіцієнт узгодженості перевищує нормативний показник, значення якого дорівнює 0,56.

Таким чином утворимо таблицю 4.4, де попарно порівняємо всі параметри.

Таблиця 4.4 - Попарне порівняння параметрів.

Параметри	Експерти							Кінцева оцінка	Числове значення
	1	2	3	4	5	6	7		
X1 і X2	>	>	<	>	<	>	>	>	1,5
X1 і X3	>	>	>	>	>	>	>	>	1,5
X1 і X4	>	>	>	>	>	>	>	>	1,5
X2 і X3	>	<	>	>	>	>	>	>	1,5
X2 і X4	>	>	>	>	>	>	>	>	1,5
X3 і X4	>	>	>	<	<	>	<	>	1,5

Ступінь переваги i -го параметра над j -тим визначається ячисловим параметром a_{ij} і визначається по формулі:

$$a_{ij} = \begin{cases} 1.5 \text{ при } X_i > X_j \\ 1.0 \text{ при } X_i = X_j \\ 0.5 \text{ при } X_i < X_j \end{cases} \quad (4.6)$$

З отриманих числових оцінок переваги складемо матрицю $A = \|a_{ij}\|$.

Для кожного параметра зробимо розрахунок вагомості $K_{\delta i}$ за наступними формулами:

$$K_{\delta i} = \frac{b_i}{\sum_{i=1}^n b_i} \quad (4.7)$$

$$b_i = \sum_{i=1}^N a_{ij} \quad (4.8)$$

Відносні оцінки розраховуються декілька разів доти, поки наступні значення не будуть незначно відрізнятись від попередніх (менше 2%). На другому і наступних кроках відносні оцінки розраховуються за наступними формулами:

$$K_{Bi} = \frac{b'_i}{\sum_{i=1}^n b'_i}, \quad (4.9)$$

$$b'_i = \sum_{j=1}^N a_{ij} b_j \quad (4.10)$$

Як видно з таблиці 4.5, різниця значень коефіцієнтів вагомості не перевищує 2%, тому більшої кількості ітерацій не потрібно.

Таблиця 4.5 - Розрахунок вагомості параметрів

Параметри x_i	Параметри x_j				Перша ітер.		Друга ітер.	
	1	2	3	4	b_i	K_{Bi}	b_i^1	K_{Bi}^1
X1	1	0,5	1,5	1,5	4,5	0,28125	16,15	0,2740
X2	1,5	1	0,5	1,5	4,5	0,28125	17,09	0,2896
X3	0,5	0,5	1	1,5	3,5	0,21875	13,35	0,2261
X4	0,5	1,5	0,5	1	3,5	0,21875	12,41	0,21031
Всього:					16	1	59	1

4.5 Аналіз рівня якості варіантів реалізації функцій

Визначаємо рівень якості кожного варіанту виконання основних функцій окремо.

Абсолютні значення параметрів $X1$ (Гіпотетичний об'єм коду програми), $X2$ (Час моделювання) та $X4$ (Точність моделі) відповідають технічним вимогам умов функціонування даного ПП.

Абсолютне значення параметра $X3$ (Час витрачений на виконання поставленого завдання) обрано не найгіршим.

Коефіцієнт технічного рівня для кожного варіанта реалізації ПП розраховується так (таблиця 4.6):

$$K_K(j) = \sum_{i=1}^n K_{ei,j} B_{i,j}, \quad (4.11)$$

де n – кількість параметрів;

K_{ei} – коефіцієнт вагомості i -го параметра;

B_i – оцінка i -го параметра в балах.

Таблиця 4.6 - Розрахунок показників рівня якості варіантів реалізації основних функцій ПП

Основні функції	Варіант реалізації функції	Параметри	Абсолютне значення параметра	Бальна оцінка параметра	Коефіцієнт вагомості параметра	Коефіцієнт рівня якості
F1	А	X1	700	26	0,22	5,81
F2	А	X4	95	33	0,37	12,21
F3	Б	X4	80	16	0,37	5,83
F3	А	X3	80	20	0,17	3,33

За даними з таблиці 4.6 за формулою:

$$K_K = K_{TY}[F_{1k}] + K_{TY}[F_{2k}] + \dots + K_{TY}[F_{zk}], \quad (4.12)$$

визначаємо рівень якості кожного з варіантів:

$$K_{K1} = 5,81 + 12,21 + 3,33 = 21,35 ;$$

$$K_{K2} = 5,81 + 5,83 + 3,33 = 14,97$$

Як видно з розрахунків, кращим є 1 варіант, для якого коефіцієнт технічного рівня має найбільше значення.

Як видно з розрахунків, кращим є 1 варіант, для якого коефіцієнт технічного рівня має найбільше значення.

4.6 Економічний аналіз варіантів розробки ПП

Для визначення вартості розробки ПП спочатку проведемо розрахунок трудомісткості.

Всі варіанти включають в себе два окремих завдання:

1. Розробка проекту програмного продукту;
2. Розробка програмної оболонки;

Завдання 1 за ступенем новизни відноситься до групи Б, завдання 2 до групи Б. За складністю алгоритми, які використовуються в завданні 1 належать до групи 1; а в завданні 2 – до групи 2.

Для реалізації завдання 1 використовується довідкова інформація, а завдання 2 використовує інформацію у вигляді даних.

Проведемо розрахунок норм часу на розробку та програмування для кожного з завдань.

Загальна трудомісткість обчислюється як:

$$T_0 = T_P \cdot K_{\Pi} \cdot K_{СК} \cdot K_M \cdot K_{СТ} \cdot K_{СТ.М}, \quad (4.13)$$

де T_P – трудомісткість розробки ПП;

K_{Π} – поправочний коефіцієнт;

$K_{СК}$ – коефіцієнт на складність вхідної інформації;

K_M – коефіцієнт рівня мови програмування;

$K_{СТ}$ – коефіцієнт використання стандартних модулів і прикладних програм;

$K_{СТ.М}$ – коефіцієнт стандартного математичного забезпечення

Для першого завдання, виходячи із норм часу для завдань розрахункового характеру ступеню новизни А та групи складності алгоритму

1, трудомісткість дорівнює: $T_p = 37$ людино-днів. Поправочний коефіцієнт, який враховує вид нормативно-довідкової інформації для першого завдання: $K_{\Pi} = 1.8$. Поправочний коефіцієнт, який враховує складність контролю вхідної та вихідної інформації для всіх завдань рівний 1: $K_{СК} = 1$. Оскільки при розробці першого завдання використовуються стандартні модулі, врахуємо це за допомогою коефіцієнта $K_{СТ} = 0.6$. Тоді загальна трудомісткість програмування першого завдання дорівнює:

$$T_1 = 37 \cdot 1.8 \cdot 1 \cdot 0.6 = 39.96 \text{ людино-днів.}$$

Проведемо аналогічні розрахунки для подальших завдань.

Для другого завдання (використовується алгоритм третьої групи складності, степінь новизни Б), тобто $T_p = 43$ людино-днів, $K_{\Pi} = 1.08$, $K_{СК} = 1$, $K_{СТ} = 0.7$:

$$T_2 = 43 \cdot 1.08 \cdot 1 \cdot 0.7 = 32.6 \text{ людино-днів.}$$

Складаємо трудомісткість відповідних завдань для кожного з обраних варіантів реалізації програми, щоб отримати їх трудомісткість:

$$T_0 = (39.96 + 32.6) \cdot 8 = 582.7 \text{ людино-годин.}$$

В розробці беруть участь один програміст з окладом 40000 грн., один спеціаліст аналітик у сфері демографії з окладом 35000. Визначимо середню зарплату за годину за формулою:

$$C_{\text{ч}} = \frac{M}{T_m \cdot t} \text{ грн.,} \quad (4.14)$$

де M – місячний оклад працівників;

T_m – кількість робочих днів на місяць;

t – кількість робочих годин в день.

$$C_{\text{ч}} = \frac{40000 + 35000}{2 \cdot 18 \cdot 8} = 260,42 \text{ грн.} \quad (4.15)$$

Тоді, розрахуємо заробітну плату за формулою:

$$C_{\text{зп}} = C_{\text{ч}} \cdot T_i \cdot K_d, \quad (4.16)$$

де $C_{\text{ч}}$ – величина погодинної оплати праці програміста;

T_i – трудомісткість відповідного завдання;

K_d – норматив, який враховує додаткову заробітну плату.

Зарплата розробників за варіантами становить:

$$C_{ЗП} = 260.42 \cdot 578.88 \cdot 1.125 = 169542,06 \text{ грн.}$$

Відрахування на єдиний соціальний внесок становить 22%:

$$C_{ВД} = C_{ЗП} \cdot 0.22 = 169542,06 \cdot 0.22 = 37299,25 \text{ грн.}$$

Тепер визначимо витрати на оплату однієї машино-години. (C_M)

Так як одна ЕОМ обслуговує одного програміста з окладом 40000 грн., з коефіцієнтом зайнятості 0,2 то для однієї машини отримаємо:

$$C_G = 12 \cdot M \cdot K_3 = 12 \cdot 40000 \cdot 0,2 = 96000 \text{ грн.}$$

З урахуванням додаткової заробітної плати:

$$C_{ЗП} = C_G \cdot (1 + K_3) = 96000 \cdot (1 + 0.2) = 115200 \text{ грн.}$$

Відрахування на соціальний внесок:

$$C_{ВД} = C_{ЗП} \cdot 0.22 = 115200 \cdot 0,22 = 25344 \text{ грн.}$$

Амортизаційні відрахування розраховуємо при амортизації 25% та вартості ЕОМ – 41500 грн.

$$C_A = K_{TM} \cdot K_A \cdot Ц_{ПР} = 1.23 \cdot 0.25 \cdot 41500 = 12761,25 \text{ грн.,}$$

де K_{TM} – коефіцієнт, який враховує витрати на транспортування та монтаж приладу у користувача;

K_A – річна норма амортизації;

$Ц_{ПР}$ – договірна ціна приладу.

Витрати на ремонт та профілактику розраховуємо як:

$$C_P = K_{TM} \cdot Ц_{ПР} \cdot K_P = 1.23 \cdot 41500 \cdot 0.03 = 1531,35 \text{ грн.,}$$

де K_P – відсоток витрат на поточні ремонти.

Ефективний годинний фонд часу ПК за рік розраховуємо за формулою:

$$\begin{aligned} T_{ЕФ} &= (D_K - D_B - D_C - D_P) \cdot t_3 \cdot K_B = (365 - 105 - 21 - 16) \cdot 8 \cdot 0.87 = \\ &= 1552,08 \text{ години,} \end{aligned}$$

де D_K – календарна кількість днів у році;

D_B, D_C – відповідно кількість вихідних та святкових днів;

D_P – кількість днів планових ремонтів устаткування;

t – кількість робочих годин в день;

K_B – коефіцієнт використання приладу у часі протягом зміни.

Витрати на оплату електроенергії розраховуємо за формулою:

$$C_{\text{ЕЛ}} = T_{\text{ЕФ}} \cdot N_{\text{С}} \cdot K_3 \cdot C_{\text{ЕН}} = 1552,08 \cdot 0,3 \cdot 0,5 \cdot 9,43 = 2195,58 \text{ грн.},$$

де $N_{\text{С}}$ – середньо-споживча потужність приладу;

K_3 – коефіцієнтом зайнятості приладу;

$C_{\text{ЕН}}$ – тариф за 1 кВт-годин електроенергії.

Накладні витрати розраховуємо за формулою:

$$C_{\text{Н}} = C_{\text{ПР}} \cdot 0,67 = 169542,06 \cdot 0,67 = 113596,18 \text{ грн.}$$

Тоді, річні експлуатаційні витрати будуть:

$$C_{\text{ЕКС}} = C_{\text{ЗП}} + C_{\text{ВІД}} + C_{\text{А}} + C_{\text{Р}} + C_{\text{ЕЛ}} + C_{\text{Н}}, \quad (4.17)$$

$$C_{\text{ЕКС}} = 169542,06 + 37299,25 + 12759,38 + 1531,35 + 2195,58 + 113596,18 = 336923,80 \text{ грн}$$

Собівартість однієї машино-години ЕОМ дорівнюватиме:

$$C_{\text{М-Г}} = C_{\text{ЕКС}} / T_{\text{ЕФ}} = 336923,80 / 1552,08 = 217,07 \text{ грн/год.}$$

Оскільки в даному випадку всі роботи, які пов'язані з розробкою програмного продукту ведуться на ЕОМ, витрати на оплату машинного часу, в залежності від обраного варіанта реалізації, складає:

$$C_{\text{М}} = C_{\text{М-Г}} \cdot T, \quad (4.18)$$

$$C_{\text{М}} = 217,07 \cdot 578,88 = 125624,83 \text{ грн.}$$

Накладні витрати складають 67% від заробітної плати:

$$C_{\text{Н}} = C_{\text{ЗП}} \cdot 0,67, \quad (4.19)$$

$$\text{I. } C_{\text{Н}} = 169542,06 \cdot 0,67 = 113596,18 \text{ грн.}$$

$$\text{II. } C_{\text{Н}} = 151189,5 \cdot 0,67 = 101296,965 \text{ грн.}$$

Отже, вартість розробки ПП за варіантами становить:

$$C_{\text{ПП}} = C_{\text{ЗП}} + C_{\text{ВІД}} + C_{\text{М}} + C_{\text{Н}}, \quad (4.20)$$

$$\text{I. } C_{\text{ПП}} = 169542,06 + 37299,25 + 125624,83 + 113596,18 = 445062,32 \text{ грн}$$

$$\text{II. } C_{\text{ПП}} = 169542,06 + 37299,25 + 125624,83 + 101296,965 = 433061,22 \text{ грн.}$$

4.7 Вибір кращого варіанту ПП техніко-економічного рівня

Розрахуємо коефіцієнт техніко-економічного рівня за формулою:

$$K_{\text{ТЕР}j} = K_{\text{К}j} / C_{\text{Ф}j}, \quad (4.21)$$

$$K_{\text{ТЕР}1} = 21,35 / 445062,32 = 4,796 \cdot 10^{-5},$$

На основі проведеного аналізу вартості розробки програмного продукту було визначено трудомісткість для двох ключових завдань: розробки проекту програмного продукту та програмної оболонки. Для першого завдання трудомісткість склала 65.94 людино-днів, з урахуванням використання довідкової інформації та стандартних модулів, що враховано через поправочний коефіцієнт 1.021 та коефіцієнт використання стандартних модулів 0.6. Друге завдання має трудомісткість 20.42 людино-дні та передбачає використання алгоритмів третьої групи складності з поправочним коефіцієнтом 1.08.

Сумарна трудомісткість обох завдань у людино-годинах становить 578.88, що дозволяє оцінити необхідні ресурси та тривалість реалізації проекту.

Загальна вартість розробки програмного продукту становить 445 062.32 грн, що включає заробітну плату, соціальні внески, витрати на машинний час та накладні витрати.

Значення коефіцієнта техніко-економічного рівня становить: $4,796 \cdot 10^{-5}$.

Це свідчить про помірну ефективність використання ресурсів і підкреслює необхідність подальшої оптимізації витрат або обґрунтування очікуваної користі від розробки.

4.8. Висновки до розділу 4

У даному розділі було здійснено повний функціонально-вартісний аналіз програмного продукту, призначеного для прогнозування демографічних змін у Німеччині. Проведено декомпозицію основної цілі на окремі функції, побудовано морфологічну карту та сформовано позитивно-негативну матрицю для варіантів реалізації.

Проведено оцінювання параметрів за допомогою експертного ранжування та обчислено вагомості функцій. На основі бальної оцінки та технічних характеристик було розраховано коефіцієнт технічного рівня. Крім того, виконано економічне обґрунтування вартості розробки, включаючи розрахунок заробітної плати, ЄСВ, вартості машино-години, накладних витрат та інших складових.

У результаті визначено найбільш доцільний варіант реалізації системи з точки зору техніко-економічної ефективності, що дозволяє раціонально використовувати ресурси в межах проекту.

ВИСНОВКИ

Дослідження, виконане у межах дипломної роботи, було зосереджене на кількісному аналізі демографічної ситуації в Німеччині з урахуванням сучасних викликів, таких як старіння населення, зниження народжуваності та залежність від міграційних потоків.

У рамках роботи було досліджено часові ряди ключових показників (TFR, смертність, нетто-міграція) та реалізовано підхід прогнозування на основі ARIMA-моделей. Для кожного з процесів проведено перевірку стаціонарності, обрано фактори, оцінено залишки, а результати проаналізовано з точки зору якості. Додатково було розроблено сценарний блок, який дозволяє моделювати декілька варіантів розвитку ситуації залежно від змін у міграційній динаміці.

Розрахунки та візуалізації реалізовано в середовищі Python.

Аналіз результатів показав, що за відсутності достатньої міграційної підтримки чисельність населення Німеччини в середньостроковій перспективі буде скорочуватись, навіть за умов помірнього зростання народжуваності. Це підкреслює ключову роль міграційної політики у демографічній стабільності країни.

Робота має прикладне значення: отримані моделі можуть бути використані як інструмент у плануванні політики у сфері соціального захисту, інфраструктури, охорони здоров'я та ринку праці.

У подальших дослідженнях доцільно розширити модель, включивши такі параметри, як доходи населення, рівень зайнятості, освітній рівень, а також додатково протестувати комбіновані методи прогнозування. Актуальним також є вивчення впливу кризових подій, зокрема війни в Україні, на демографічні потоки до країн ЄС, включно з Німеччиною.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. World Population Prospects 2024: Summary of Results. United Nations Department of Economic and Social Affairs, Population Division. URL: https://population.un.org/wpp/assets/Files/WPP2024_Summary-of-Results.pdf (last accessed: 28.05.2025).
2. Lutz W., Amran G., Belanger A. Demographic scenarios for the EU: Demography and human capital scenarios for the 21st century. URL: https://knowledge4policy.ec.europa.eu/publication/demographic-scenarios-eu_en (last accessed: 28.05.2025).
3. Bevölkerung Deutschlands nach Altersgruppen. Statistisches Bundesamt (Destatis). URL: <https://www.destatis.de> (last accessed: 20.05.2025).
4. Rowland D. T. *Demographic methods and concepts*. Oxford: Oxford University Press, 2003. 550 p.
5. Developments in demographic forecasting / N. Keilman, S. Mazzuco (eds.). Cham: Springer, 2021. 323 p.
6. Box G. E. P., Cox D. R. An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*. 1964. Vol. 26, No. 2. P. 211–252.
7. Yusuf F., Martins J. M., Swanson D. A. *Demographic methods and concepts*. Oxford: Oxford University Press, 2014. 324 p.
8. Keilman N., Mazzuco S. On the relationship between the total fertility rate and tempo-adjusted fertility. *Demographic Research*. 2020. Vol. 42. P. 429–458.
9. Alkema L., Raftery A. E., Gerland P., Clark S. J., Pelletier F. Probabilistic projections of the total fertility rate for all countries. *Demography*. 2011. Vol. 48, No. 3. P. 815–839.

10. Bevölkerung und Erwerbstätigkeit. Sterbefälle und Lebenserwartung. Statistisches Bundesamt (Destatis). URL: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Sterbefaelle-Lebenserwartung/_inhalt.html (last accessed: 28.05.2025).
11. Migration und Integration: Bericht 2016. Bundesamt für Migration und Flüchtlinge (BAMF). URL: <https://www.bamf.de> (last accessed: 21.05.2025).
12. International Migration Outlook 2023. OECD. URL: <https://doi.org/10.1787/b0f40584-en> (last accessed: 20.05.2025).
13. Net migration statistics. World Bank. URL: <https://data.worldbank.org/indicator/SM.POP.NETM> (last accessed: 22.05.2025).
14. World Development Indicators: Total Fertility Rate (TFR). World Bank. URL: <https://data.worldbank.org/indicator/SP.DYN.TFRT.IN> (last accessed: 29.05.2025).
15. Germany: Period Life Tables 1x1. Human Mortality Database. URL: <https://www.mortality.org> (last accessed: 29.05.2025).
16. International Migration 2020 Highlights. United Nations Department of Economic and Social Affairs. URL: https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/undes_a_pd_2020_international_migration_highlights.pdf (last accessed: 26.05.2025).
17. Booth H., Hyndman R. J., Tickle L., de Jong P. Lee–Carter mortality forecasting: a multi-country comparison of variants and extensions. *Demographic Research*. 2006. Vol. 15, No. 9. P. 289–310. URL: <https://doi.org/10.4054/DemRes.2006.15.9> (last accessed: 29.05.2025).
18. Guerrero V. M. Time-series analysis supported by power transformations. *Journal of Forecasting*. 1993. Vol. 12, No. 1. P. 37–48.

ДОДАТОК А (ілюстративний матеріал доповіді)

Моделі та прогнози демографічних змін у Німеччині з урахуванням міграційних процесів

Виконала: студентка 4-го курсу
Групи КА-15
Лоїк Ірина Валеріївна

Керівник: Старший викладач кафедри ММСА, PhD
Левенчук Людмила Борисівна

Актуальність

Демографічна ситуація в Німеччині характеризується тривалим зниженням народжуваності, старінням населення та дефіцитом трудових ресурсів. Ці процеси створюють навантаження на соціальну сферу, пенсійну систему та економіку загалом.

- Побудова прогнозних моделей демографічних змін з урахуванням міграційних процесів є важливим інструментом для підтримки прийняття рішень у сфері державної політики. Такий підхід дозволяє досліджувати динаміку складної соціально-демографічної системи в умовах невизначеності.

Об'єкт дослідження:

Демографічні процеси в Німеччині.

Предмет дослідження:

Моделі прогнозування чисельності населення з урахуванням впливу міграційних потоків.

Мета дослідження:

Розробити модель прогнозування чисельності населення Німеччини до 2035 року з урахуванням впливу міграційних процесів та показників народжуваності, застосовуючи методи системного аналізу, статистичного моделювання (ARIMA) та сценарного прогнозування.

3

Постановка задачі

1	Провести огляд сучасних підходів до моделювання та прогнозування демографічних процесів, зокрема з урахуванням міграції.
2	Зібрати та проаналізувати основні демографічні показники Німеччини: народжуваність, смертність та міграцію.
3	Побудувати прогнозні моделі для показників TFR, міграції та тривалості життя з використанням статистичних методів.
4	Розрахувати чисельність населення Німеччини до 2035 року з урахуванням прогнозованих демографічних змін.
5	Реалізувати модель прогнозування у середовищі Python з подальшою візуалізацією результатів.
6	Провести оцінку якості прогнозу та зробити висновки щодо ймовірного розвитку демографічної ситуації

4

Аналіз і прогноз TFR (сумарного коефіцієнта народжуваності)

- Проведено тест Дікі–Фуллера: підтверджено нестационарність ряду ($p \approx 0.21$)
- Перебір моделей ARIMA(p,1,q)
- Обрано ARIMA(2,1,1) за критеріями AIC, MAE, DW, Ljung–Box

Модель	AIC	BIC	DW	MAE	Ljung–Box (p-value)
ARIMA(2,1,1)	-104.69	-98.71	1.200	0.0866	1.0
ARIMA(1,1,5)	-105.81	-95.33	1.192	0.0890	1.0
ARIMA(2,1,3)	-104.71	-95.74	1.194	0.0885	1.0

Таблиця 1. Моделі кандидати для прогнозування TFR

Остаточний вибір — ARIMA(2,1,1) як найкраща за простотою та стабільністю залишків.

5

Прогноз TFR

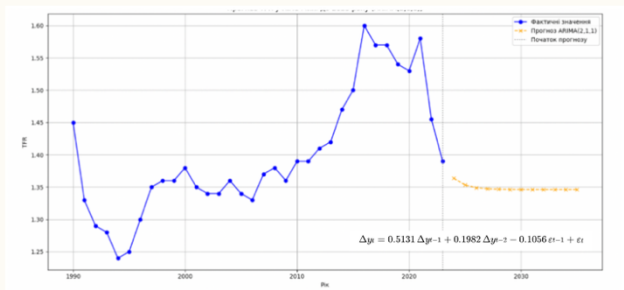
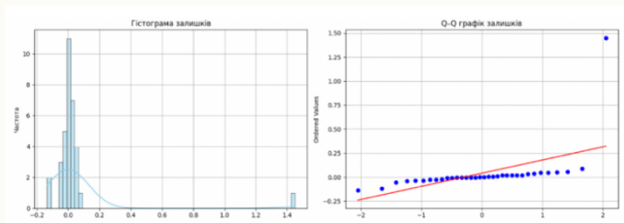


Рисунок 1. Побудований прогноз ARIMA(2,1,1)



Рік	Прогноз TFR
2024	1.364
2025	1.353
2026	1.349
2027	1.347
2028	1.347
2029	1.347
2030	1.346
2031	1.346
2032	1.346
2033	1.346
2034	1.346
2035	1.346

Таблиця 2. Значення прогнозованого показника народжуваності

6

Прогноз показника нетто-міграції в Німеччині

Застосовано комбіновану модель:

- лінійна регресія для тренду (2010–2023)

-ARIMA(1,0,1) для моделювання залишків

Попередньо ряд перевірено на стаціонарність (ADF $p = 0.0219 \rightarrow d=0$)

Показник	Значення
Durbin-Watson	2.27
Ljung-Box (p)	0.6125
D'Agostino (p)	0.8089
MAE	≈ 64 тис.
RMSE	≈ 84 тис.

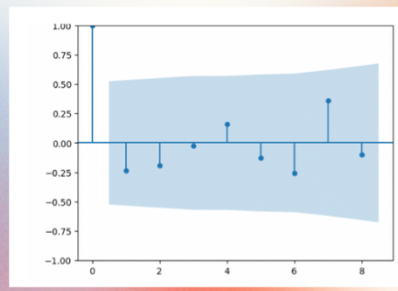
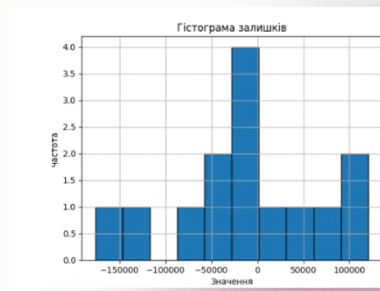
Таблиця 3. Статистичні показники комбінованої моделі



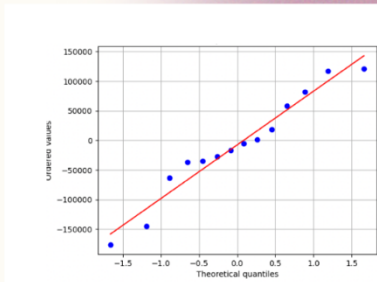
Рисунок 2. Побудований прогноз нетто-міграції

7

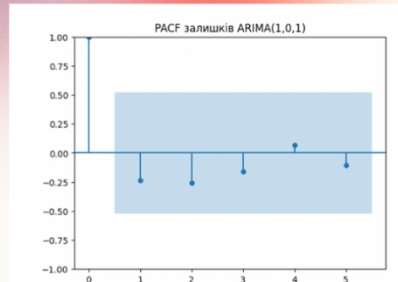
Прогноз показника нетто-міграції в Німеччині



ACF залишків



Q-Q графік залишків



PACF залишків ARIMA(1,0,1)

Рік	Прогноз нетто-м
2024	604 512
2025	484 323
2026	922 065
2027	685 560
2028	517 044
2029	960 071
2030	447 921
2031	917 588
2032	764 294
2033	927 506
2034	858 477
2035	842 262

Таблиця 5. Прогнозовані значення показника нетто-міграції

Рисунок 4. Аналіз залишків

8

Сценарне прогнозування нетто-міграції в Німеччині

Використано експоненційну модель з доданим шумом:

$$M_t = M_{t-1} \cdot (1 + g)^t + \varepsilon_t$$



Рисунок 5. Побудований прогноз нетто-міграції

9

Прогноз тривалості життя (expectancy at birth) у Німеччині

За ARIMA

- Проведено тест Дікі-Фуллера: підтверджено нестационарність ряду
- Серед моделей кандидатів обрано ARIMA(2,2,2) за критеріями AIC, MAE, DW, Ljung-Box

Модель	AIC	BIC	DW	MAE	Ljung-P
ARIMA(2,2,2)	-12.43	-5.60	1.998	3.76	0.98
ARIMA(1,2,1)	-13.81	-9.07	1.192	3.77	0.98
ARIMA(0,2,0)	16.71	17.90	1.980	3.96	0.95

Таблиця 6. Найкращі протестовані моделі ARIMA.
Остаточний вибір — ARIMA(2,2,2)

10

Прогноз тривалості життя (expectancy at birth) у Німеччині

Також була протестована логістична модель

$$e_0(t) = \frac{L}{1 + e^{-k(t-m)}}$$

Модель	MAE (роки)	MAPE	RMSE	R2	DW	Jarque-Bera p
ARIMA(1,2,1)	0.14	0.18	0.24	-	1.99	0.78
Логістична	0.12	0.16	0.15	0.99	-	0.70

Таблиця 7. Порівняння основних статистичних показників прогнозування за ARIMA та Logistic regression

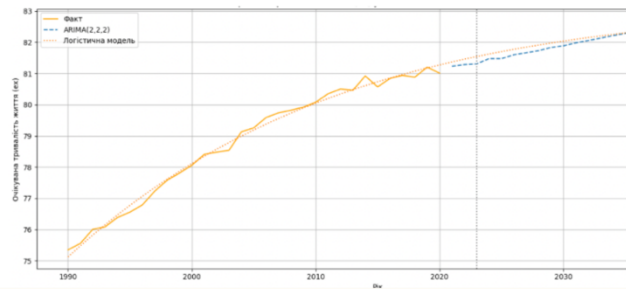


Рисунок 7. Побудовані прогнози показника тривалості життя)

Прогноз тривалості життя (expectancy at birth) у Німеччині

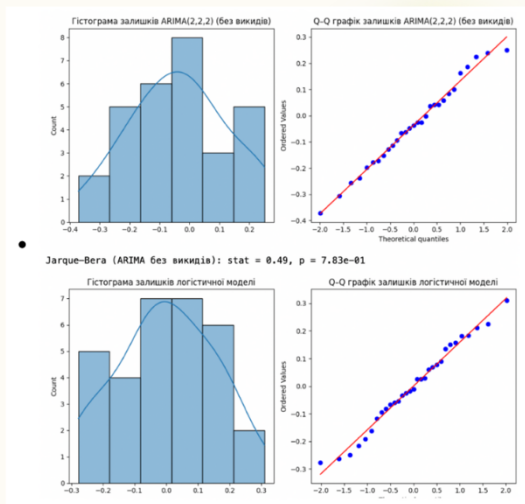


Рисунок 8. Аналіз залишків

Рік	Прогноз (логістика)
2024	81.77
2025	81.77
2026	81.77
2027	81.84
2028	81.91
2029	81.97
2030	82.03
2031	82.09
2032	82.15
2033	82.20
2034	82.25
2035	82.30

Таблиця 8. Прогнозовані значення показника тривалості життя

Прогноз чисельності населення Німеччини (2024–2035)

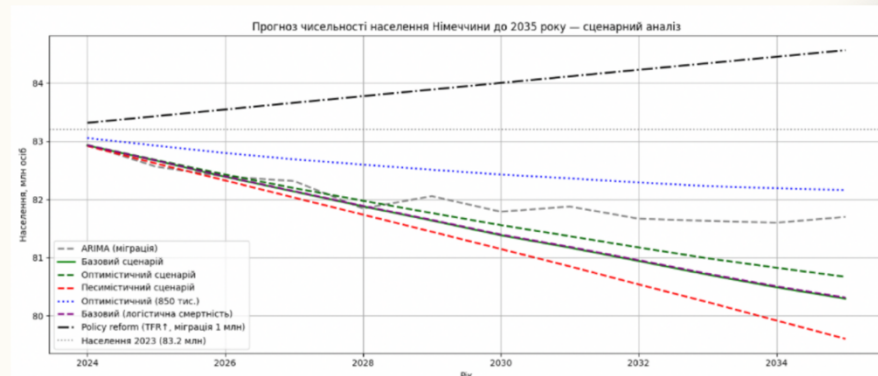


Рисунок 9. Побудовані прогнози чисельності населення Німеччини)

Сценарій	2030	2035
Базовий	81,383,217	80,293,012
Оптимістичний	81,557,792	80,667,983
Песимістичний	81,147,270	79,602,406
850 тис. міграції щороку	82,427,322	82,160,489
Базовий (логістична смертність)	81,394,981	80,310,143
Policy reform (TFR ↑ , міграція 1 млн)	84,003,270	84,565,403

13

Висновки

1. Побудовано прогнози моделі для трьох ключових демографічних компонент: народжуваності (TFR), смертності (e_0) та нетто-міграції.
2. Для кожної змінної протестовано кілька підходів (ARIMA, логістична модель, гібридні схеми) обрано найадекватніші з точки зору статистичних критеріїв.
3. Проведено сценарний аналіз чисельності населення Німеччини до 2035 року з урахуванням різних варіантів міграційної та соціальної політики.
4. Результати показали: міграція є критичним компенсатором демографічного спаду.
5. За умови збереження поточних тенденцій населення зменшується, тоді як при реформеному сценарії можливе зростання.

14

Наукова новизна

Запропоновано гібридний підхід до моделювання міграції

який поєднує побудову лінійного тренду з окремим моделюванням залишкових флуктуацій за допомогою ARIMA.

Проведено порівняльний аналіз моделей прогнозування смертності:

логістичної функції та ARIMA-моделі.

Реалізовано сценарне прогнозування з урахуванням випадкових флуктуацій

білий шум $\varepsilon_t \sim N(0, \sigma^2)$, що наближає модель до реальних демографічних процесів

15

Подальші дослідження

1. Розширення горизонту прогнозування до 2050 року з деталізацією змін у довгостроковій перспективі
2. Побудова когортно-компонентної моделі з урахуванням вікової та статевої структури населення
3. Аналіз впливу економічних показників (ВВП, рівень зайнятості) на демографічні процеси
4. Врахування зовнішньополітичних факторів у сценарному моделюванні (наприклад, міграційні кризи)
5. Розробка простого програмного інструменту для демонстрації та візуалізації сценаріїв зміни населення

16

Дякую за увагу!

ДОДАТОК Б (статистичні дані)

Основним джерелом даних для аналізу сумарного коефіцієнта народжуваності (TFR) у Німеччині був використаний офіційний портал: <https://data.worldbank.org/indicator/SP.DYN.TFRT.IN>.

Значення були зведені до таблиці Б.1:

Таблиця Б.1 – Динаміка TFR у Німеччині (1990–2024)

Рік	Сумарний коефіцієнт народжуваності (TFR)
1990	1.45
1991	1.33
1992	1.29
1993	1.28
1994	1.24
1995	1.25
1996	1.3
1997	1.35
1998	1.36
1999	1.36
2000	1.38
2001	1.35
2002	1.34
2003	1.34
2004	1.36
2005	1.34
2006	1.33
2007	1.37
2008	1.38
2009	1.36
2010	1.39
2011	1.39
2012	1.41
2013	1.42
2014	1.47
2015	1.5
2016	1.6
2017	1.57
2018	1.57
2019	1.54
2020	1.53
2021	1.58
2022	1.455
2023	1.39

Основним джерелом даних щодо нетто-міграції виступила база *World Bank*
<https://data.worldbank.org/indicator/SM.POP.NETM>

Отримані щорічні значення міграційного приросту для Німеччини зведено в таблицю Б.2:

Таблиця Б.2 – Нетто-міграція до Німеччини (1990–2024)

Рік	Нетто-міграція
1990	679702
1991	602638
1992	811606
1993	462096
1994	315585
1995	398077
1996	281280
1997	93281
1998	-754469
1999	219280
2000	174594
2001	272875
2002	218795
2003	142009
2004	81735
2005	81351
2006	-489056
2007	-24772
2008	-23341
2009	46533
2010	130116
2011	295282
2012	199143
2013	462384
2014	592338
2015	1175283
2016	517480
2017	452988
2018	403908
2019	319082
2020	203468
2021	310264
2022	981552
2023	609553
2024	36954

Дані очікуваної тривалості життя (ex) отримані з *Human Mortality Database*

<https://www.mortality.org/Country/Country?cntr=DEUTNP>

Отримані щорічні значення очікуваної тривалості життя для Німеччини зведено в таблицю Б.3:

Таблиця Б.3 – Очікувана тривалість життя в Німеччині (1990–2024)

Рік	Очікувана тривалість життя (ex)
1990	75.35
1991	76.01
1992	76.09
1993	76.23
1994	76.37
1995	76.56
1996	76.69
1997	77.31
1998	77.58
1999	77.81
2000	78.3
2001	78.4
2002	78.6
2003	78.7
2004	78.8
2005	79.0
2006	79.1
2007	79.3
2008	79.4
2009	79.6
2010	79.7
2011	79.9
2012	80.0
2013	80.2
2014	80.4
2015	80.5
2016	80.7
2017	80.8
2018	81.0
2019	81.1
2020	81.3
2021	81.4
2022	81.2
2023	81.0

ДОДАТОК В (код программного продукта)

```

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from statsmodels.tsa.arima.model import ARIMA
from statsmodels.tsa.stattools import adfuller
from statsmodels.stats.diagnostic import acorr_ljungbox
from statsmodels.stats.stattools import durbin_watson
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_absolute_error,
mean_squared_error
from scipy.signal import savgol_filter
from scipy.optimize import curve_fit
from scipy.stats import jarque_bera, probplot, normaltest
from statsmodels.graphics.tsaplots import plot_acf, plot_pacf
import warnings
warnings.filterwarnings("ignore")

# === TFR: ARIMA(2,1,1) ===
df =
pd.read_csv("API_SP.DYN.TFRT.IN_DS2_en_csv_v2_19650.csv",
skiprows=4)
df = df[df["Country Name"] == "Germany"]
df = df.melt(id_vars=["Country Name", "Country Code",
"Indicator Name", "Indicator Code"], var_name="Year",
value_name="TFR")
df = df.dropna(subset=["TFR"])
df["Year"] = df["Year"].astype(int)
df["TFR"] = pd.to_numeric(df["TFR"], errors="coerce")
df = df.set_index("Year").loc[1990:2023]
tfr_series = df["TFR"].copy()
tfr_series.index = range(len(tfr_series))

```

```

adfuller(tfr_series)
df_diff = tfr_series.diff().dropna()
adfuller(df_diff)

model_211 = ARIMA(tfr_series, order=(2, 1, 1)).fit()
forecast_12 =
model_211.get_forecast(steps=12).predicted_mean
forecast_12.index = list(range(2024, 2036))

plt.figure(figsize=(12, 6))
plt.plot(df.index, df["TFR"], label="Факт", marker='o')
plt.plot(forecast_12.index, forecast_12.values,
label="Прогноз ARIMA(2,1,1)", linestyle="--", marker='x')
plt.axvline(x=2023, color='gray', linestyle=':')
plt.title("Прогноз TFR у Німеччині до 2035 року")
plt.xlabel("Рік")
plt.ylabel("TFR")
plt.grid(True)
plt.legend()
plt.tight_layout()
plt.show()

residuals_211 = model_211.resid
jarque_bera(residuals_211)
acorr_ljungbox(residuals_211, lags=[10], return_df=True)

fig, axs = plt.subplots(1, 2, figsize=(14, 5))
sns.histplot(residuals_211, kde=True, ax=axs[0])
axs[0].set_title("Гістограма залишків")
probplot(residuals_211, dist="norm", plot=axs[1])
axs[1].set_title("Q-Q графік залишків")
plt.tight_layout()
plt.show()

# === Нetto-міграція: ARIMA(1,0,1) ===

```

```

df = pd.read_csv("API_SM.POP.NETM_DS2_en_csv_v2_19300.csv",
skiprows=4)
df = df[df["Country Name"] == "Germany"]
df = df.melt(id_vars=["Country Name", "Country Code",
"Indicator Name", "Indicator Code"], var_name="Year",
value_name="NetMigration")
df = df.dropna(subset=["NetMigration"])
df["Year"] = df["Year"].astype(int)
df = df.set_index("Year").loc[1990:2023]
series = df["NetMigration"]

recent_data = series.loc[2010:2023]
reg =
LinearRegression().fit(recent_data.index.values.reshape(-1, 1),
recent_data.values.reshape(-1, 1))
future_years = np.arange(2024, 2036).reshape(-1, 1)
trend_forecast = reg.predict(future_years).flatten()
trend_smooth = pd.Series(savgol_filter(recent_data,
window_length=5, polyorder=2), index=recent_data.index)
residuals = recent_data - trend_smooth

model_mig = ARIMA(residuals.dropna(), order=(1, 0, 1)).fit()
resid = model_mig.resid

plot_acf(resid.dropna(), lags=8)
plt.title("ACF залишків ARIMA(1,0,1)")
plt.show()

plot_pacf(resid.dropna(), lags=5)
plt.title("PACF залишків ARIMA(1,0,1)")
plt.show()

probplot(resid.dropna(), dist="norm", plot=plt)
plt.title("Q-Q plot залишків")
plt.grid(True)
plt.show()

```

```

plt.hist(resid.dropna(), bins=10, edgecolor="black")
plt.title("Гістограма залишків")
plt.xlabel("Значення")
plt.ylabel("Частота")
plt.grid(True)
plt.show()

simulated_resid =
model_mig.simulate(nsimulations=len(future_years), anchor='end')
final_forecast = trend_forecast + simulated_resid.to_numpy()

# === Нето-міграція: сценарії ===
last_value = series.loc[2023]
t_range = np.arange(1, len(future_years) + 1)
trend_base = last_value * (1 + 0.003) ** t_range
trend_high = last_value * (1 + 0.01) ** t_range
trend_low = last_value * (1 - 0.01) ** t_range
np.random.seed(555)
base_noise = np.random.normal(0, 12000, len(future_years))
high_noise = np.abs(np.random.normal(0, 8000,
len(future_years)))
low_noise = -np.abs(np.random.normal(0, 8000,
len(future_years)))
forecast_base = trend_base + base_noise
forecast_high = trend_high + high_noise
forecast_low = trend_low + low_noise

# === Смертність (ex): ARIMA(2,2,2) та логістична ===
df = pd.read_csv("bltper_1x1.txt", sep=r"\s+", skiprows=2)
df = df[df["Age"] == "0"]
df["Year"] = df["Year"].astype(int)
df["ex"] = pd.to_numeric(df["ex"], errors="coerce")
df = df.dropna(subset=["ex"])
df = df[df["Year"] >= 1990].reset_index(drop=True)

```

```

model_ex = ARIMA(df["ex"], order=(2, 2, 2)).fit()
forecast_ex = model_ex.forecast(steps=12)

def logistic(x, L, k, x0): return L / (1 + np.exp(-k * (x -
x0)))

popt, _ = curve_fit(logistic, df["Year"].values,
df["ex"].values, p0=[90, 0.05, 2000])
logistic_ex_forecast = logistic(np.arange(2024, 2036), *popt)
logistic_ex_forecast[:2] = logistic_ex_forecast[2]

resid_arima = model_ex.resid
resid_arima_filtered = resid_arima[2:]

plt.figure(figsize=(10, 5))
plt.subplot(1, 2, 1)
sns.histplot(resid_arima_filtered, kde=True)
plt.title("Гістограма залишків ARIMA(2,2,2)")
plt.subplot(1, 2, 2)
probplot(resid_arima_filtered, dist="norm", plot=plt)
plt.title("Q-Q графік залишків")
plt.tight_layout()
plt.show()

resid_logistic = df["ex"].values -
logistic(df["Year"].values, *popt)
plt.figure(figsize=(10, 5))
plt.subplot(1, 2, 1)
sns.histplot(resid_logistic, kde=True)
plt.title("Гістограма залишків логістичної моделі")
plt.subplot(1, 2, 2)
probplot(resid_logistic, dist="norm", plot=plt)
plt.title("Q-Q графік залишків логістичної моделі")
plt.tight_layout()
plt.show()

# === Фінальний прогноз населення ===

```

```

df_forecast = pd.DataFrame({
    "Year": np.arange(2024, 2036),
    "TFR": forecast_12.values,
    "net_migration_arima": final_forecast,
    "net_migration_scenario_base": forecast_base,
    "NetMigration_Optimistic": forecast_high,
    "NetMigration_Pessimistic": forecast_low,
    "Mortality_ex": forecast_ex.values,
    "Mortality_ex_logistic": logistic_ex_forecast
})

df_forecast["NetMigration_Optimistic_850"] =
df_forecast["NetMigration_Optimistic"] * 1.2
df_forecast["NetMigration_Policy"] = 1_000_000
df_forecast["TFR_Policy"] =
np.linspace(df_forecast["TFR"].iloc[0], 1.65, len(df_forecast))
df_forecast.loc[:, "Mortality_ex"] =
df_forecast["Mortality_ex"].iloc[2]

mortality_rate = 0.015 - 0.00005 *
df_forecast["Mortality_ex"]
mortality_rate_logistic = 0.015 - 0.00005 *
df_forecast["Mortality_ex_logistic"]
p_start = 83_200_000
fertile_female_share = 0.23

def forecast_population(tfr, mig, mort_rate):
    P = [p_start]
    for i in range(len(tfr)):
        B = tfr[i] * (P[-1] * fertile_female_share) / 1000
        D = P[-1] * mort_rate[i]
        M = mig[i]
        P.append(P[-1] + B - D + M)
    return np.array(P[1:])

```

```

    pop_arima      =      forecast_population(df_forecast["TFR"],
df_forecast["net_migration_arima"], mortality_rate)
    pop_base       =      forecast_population(df_forecast["TFR"],
df_forecast["net_migration_scenario_base"], mortality_rate)
    pop_high       =      forecast_population(df_forecast["TFR"],
df_forecast["NetMigration_Optimistic"], mortality_rate)
    pop_low        =      forecast_population(df_forecast["TFR"],
df_forecast["NetMigration_Pessimistic"], mortality_rate)
    pop_850        =      forecast_population(df_forecast["TFR"],
df_forecast["NetMigration_Optimistic_850"], mortality_rate)
    pop_policy     =      forecast_population(df_forecast["TFR_Policy"],
df_forecast["NetMigration_Policy"], mortality_rate)
    pop_base_logistic = forecast_population(df_forecast["TFR"],
df_forecast["net_migration_scenario_base"],
mortality_rate_logistic)

    plt.figure(figsize=(14, 6))
    plt.plot(df_forecast["Year"], pop_arima / 1e6, label="ARIMA
(міграція)", linestyle="--", color="gray", linewidth=2)
    plt.plot(df_forecast["Year"], pop_base / 1e6, label="Базовий
сценарій", color="green", linewidth=2)
    plt.plot(df_forecast["Year"],      pop_high      /      1e6,
label="Оптимістичний      сценарій",      linestyle="--",
color="darkgreen", linewidth=2)
    plt.plot(df_forecast["Year"],      pop_low      /      1e6,
label="Песимістичний      сценарій",      linestyle="--",      color="red",
linewidth=2)
    plt.plot(df_forecast["Year"],      pop_850      /      1e6,
label="Оптимістичний      (850 тис.)",      linestyle=":",      color="blue",
linewidth=2)
    plt.plot(df_forecast["Year"],      pop_base_logistic / 1e6,
label="Базовий      (логістична      смертність)",      linestyle="--",
color="purple", linewidth=2)
    plt.plot(df_forecast["Year"], pop_policy / 1e6, label="Policy
reform (TFR↑, міграція 1 млн)", linestyle="-.", color="black",
linewidth=2)

```

```
plt.axhline(y=p_start / 1e6, color="gray", linestyle=":",  
label="Населення 2023 (83.2 млн)")  
plt.title("Прогноз чисельності населення Німеччини до 2035  
року – сценарний аналіз")  
plt.xlabel("Рік")  
plt.ylabel("Населення, млн осіб")  
plt.grid(True)  
plt.legend()  
plt.tight_layout()  
plt.show()
```