

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

**Факультет інформатики та обчислювальної техніки
Автоматизованих систем обробки інформації і управління**

«На правах рукопису»
УДК 004.946 .

«До захисту допущено»
Завідувач кафедри
_____ О.А.Павлов
«__» _____ 2018 р.

Магістерська дисертація

на здобуття ступеня магістра

зі спеціальності 121 Інженерія програмного забезпечення

**на тему: «Методи та програмні засоби прогнозування часового ряду на
основі методу аналізу ієрархій»**

Виконав (-ла):

студент (-ка) VI курсу, групи ПП-61м

Грачова Олександра Андріївна _____

Керівник:

доцент, к.т.н., доцент

Фіногенов О.Д. _____

Рецензент:

Засвідчую, що у цій магістерській
дисертації немає запозичень з праць
інших авторів без відповідних посилань.
Студент (-ка) _____

**Методи та програмні засоби прогнозування часового ряду на основі методу
аналізу ієрархій**

РЕФЕРАТ

Звіт про НДР: с., рис., табл., додатки, джерела.

Об'єкт дослідження – часові ряди показників погодних умов (температури, вологості, атмосферного тиску, хмарності і т.д.)

Мета роботи – створення системи аналізу та прогнозування часового ряду на основі методу аналізу ієрархій, що дозволить обирати найкращий ресурс погоди.

Результатом дослідження є розроблений математичний апарат буде здатний надавати достовірні погодні дані на основі вибірки з п'яти популярних сайтів погоди та пропонувати найкраще достовірне джерело даних. Завдяки універсальності методу аналізу ієрархій, математичний апарат може бути використаний для дослідження та прогнозування будь-якого часового ряду.

Зв'язок з іншими роботами: Грачова О.А, Фіногенов О.Д. «Аналіз даних сайтів погоди на основі методу аналізу ієрархій»: Міжнародна науково-технічна конференція SAIT 2018, Київ; Грачова О.А, Фіногенов О.Д «МЕТОД АНАЛІЗУ ІЄРАРХІЙ ПРИ АНАЛІЗІ ДАНИХ, ЗАЛЕЖНИХ У ЧАСІ»: Інформатика і обчислювальна техніка ІОТ-2018.

МЕТОД АНАЛІЗУ ІЄРАРХІЙ, АНАЛІХ ДОСТОВІРНОСТІ ДАНИХ, АРХІТЕКТУРА СИСТЕМИ.

ЗМІСТ

Реферат	2
Перелік умовних позначень, символів, одиниць, скорочень.....	6
Вступ.....	7
1 Огляд методів прийняття рішень в умовах багатокритеріальності. Особливості використання для аналізу часових рядів	9
1.1 Поняття задачі вибору.....	9
1.2 Класифікація задач вибору	10
1.3 Постановка задачі багатокритеріального вибору.....	11
1.4 Теорія прийняття рішень.....	12
1.4.1 Теорія багатовимірної корисності	13
1.4.2 Теорія нечітких множин	15
1.4.3 Метод аналізу ієрархій.....	16
1.5 Порівняльний аналіз методів прийняття рішень.....	17
1.6 Поняття часового ряду	19
1.7 Класифікація часових рядів.....	19
1.7.1 Класифікація за часом.....	20
1.7.2 Класифікація за відстанню між датами.....	20
1.7.3 Класифікація за змістом показників	20
1.8 Характеристики часових рядів	20
1.9 Компоненти часових рядів.....	23
2 Засоби покращення методу аналізу ієрархій для аналізу та прогнозування часових рядів.....	24
2.1 Забезпечення можливості порівняння рівнів часового ряду.....	24

2.2	Ієрархія тенденцій та коливань	25
2.2.1	Періодизація часового ряду	27
2.2.2	Показники, що характеризують динаміку часового ряду	28
2.3	Метод моделювання і прогнозування часових рядів з довгою пам'яттю	29
2.4	Методи класифікації часового ряду.....	30
3	Дослідження ефективності системи	33
3.1	Загальний опис системи	33
3.2	Загальний опис прикладу	33
3.3	Структура вхідних даних	34
3.3.1	Стовбець Name.....	34
3.3.2	Стовбець Date.....	35
3.3.3	Стовбець Time.....	35
3.3.4	Стовбці Temperature_X	36
3.3.5	Стовбці Wind_X.....	36
3.3.6	Стовбці Pressure_X	37
3.3.7	Стовбці Humidity_X	37
3.3.8	Стовбці Preciptation_X	38
3.3.9	Стовбці Cloudness_X.....	38
3.3.10	Стовбці Feels_Like_X.....	39
3.4	Побудова та оцінка ієрархії	39
3.5	Аналіз результатів тестового прикладу на достовірність	42
3.5.1	Аналіз отриманих даних на прикладі показників з числовими значеннями.....	43

3.5.2	Аналіз отриманих даних на прикладі показників з текстовими значеннями.....	46
3.6	Розрахунки.....	48
3.6.1	Розрахунок власного вектору.....	48
3.6.2	Лінійна згортка.....	49
4	Порівняння побудованої системи із системою expert choice.....	51
4.1	Внесення даних	51
4.2	Побудова ієрархії та оцінка критеріїв	51
4.3	Оцінка альтернатив.....	55
4.4	Вибір найкращої альтернативи	57
	Висновки	61
	Перелік посилань.....	62

**ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ,
СКОРОЧЕНЬ**

МАІ – метод аналізу ієрархій;

СПР – система прийняття рішень;

ОПР – особа, що приймає рішення

ТПР – теорія прийняття рішень;

ТК – теорія користності;

ТНМ – теорія нечітких множин.

ВСТУП

У сучасному світі зі збільшенням негативних факторів впливу все важливіше місце займає дослідження здоров'я людини. Першим кроком до усунення негативних ефектів є відповідні дії людини на погодні умови, прогнозування яких є вкрай трудомісткою та в багатьох випадках нечіткою задачею. На даний момент, все більше людей отримують дані погоди завдяки мережі інтернет, але на багатьох сайтах, що інформують про погодні умови, є відмінності як у показниках, так і в наведених даних.

Вибір одного з декількох джерел інформації про погоду можна представити у вигляді багатокритеріальної задачі, одним з методів вирішення якої є метод аналізу ієрархій (МАІ), що містить аналіз не лише кількісних, але і якісних показників. Вибір людиною погодних сайтів також обумовлений великою інерцією – звичне налаштування або інтуїтивно зрозуміле подання інформації може бути більш впливовим фактором, а ніж достовірність.

Нажаль, система Expert Choice, яка є однією з найпоширеніших систем прийняття рішень (СПР) не надає можливості вносити та аналізувати показники альтернатив. Таким чином, доцільним стає питання дослідження ефективності МАІ у порівнянні з іншими методами та теоріями прийняття рішень, оцінка універсальності та можливостей застосування системи, яка б підтримувала такий підхід до аналізу та прогнозування.

В ході дослідження буде розглянуто існуючі методи прийняття рішень, проведено їх аналіз та зроблено висновки про їх універсальність.

Об'єкт дослідження – часові ряди показників погодних умов (температури, вологості, атмосферного тиску, хмарності і т.д.)

Мета роботи – створення системи аналізу та прогнозування часового ряду на основі методу аналізу ієрархій, що дозволить обирати найкращий ресурс погоди.

Результатом дослідження є розроблений математичний апарат буде здатний надавати достовірні погодні дані на основі вибірки з п'яти популярних сайтів погоди та пропонувати найкраще достовірне джерело даних. Завдяки універсальності методу аналізу ієрархій, математичний апарат може бути використаний для дослідження та прогнозування будь-якого часового ряду.

Зв'язок з іншими роботами: Грачова О.А, Фіногенов О.Д. «Аналіз даних сайтів погоди на основі методу аналізу ієрархій»: Міжнародна науково-технічна конференція SAIT 2018, Київ; Грачова О.А, Фіногенов О.Д «МЕТОД АНАЛІЗУ ІЄРАРХІЙ ПРИ АНАЛІЗІ ДАНИХ, ЗАЛЕЖНИХ У ЧАСІ»: Інформатика і обчислювальна техніка ІОТ-2018.

МЕТОД АНАЛІЗУ ІЄРАРХІЙ, АНАЛІХ ДОСТОВІРНОСТІ ДАНИХ, АРХІТЕКТУРА СИСТЕМИ.

1 ОГЛЯД МЕТОДІВ ПРИЙНЯТТЯ РІШЕНЬ В УМОВАХ БАГАТОКРИТЕРІЛЬНОСТІ. ОСОБЛИВОСТІ ВИКОРИСТАННЯ ДЛЯ АНАЛІЗУ ЧАСОВИХ РЯДІВ

1.1 Поняття задачі вибору

У сучасному світі людина часто стикається із ситуаціями, що вимагають від неї прийняття рішення. Наприклад, це може бути вибір найкращого вищого навчального закладу для здобуття професії або ж вибір смаку чаю до сніданку. Задачі вибору також зустрічаються і в електронних системах, наприклад, навігатор вирішує задачу вибору, коли будує найкращий маршрут з пункту А в пункт Б. Сфер, у яких зустрічається цей клас задач, надзвичайна кількість, проте набір характеристик, необхідних для розв'язання таких задач, є спільним для всіх.

По-перше, для задачі вибору необхідно мати деяку кількість варіантів вирішення, тобто таких варіантів має бути більше, аніж два. Ці варіанти називаються *множиною імовірних розв'язків*. Зрозуміло, що обмежити зверху таку множину у деяких випадках неможливо, тому вона може бути як скінченою, так і нескінченою.

По-друге, у задачах вибору важливо мати особу, що приймає рішення. Таку людину (або групу людей) називають *експертом* або ж *особою, що приймає рішення* (ОПР). У випадках, коли група людей має узгоджену точку зору на проблему і їх рішення однакові, прийнято вважати їх єдиним ОПР.

По-третє, задача вибору зазвичай не має єдиного рішення через те, що для різних ОПР важливими виявляються різні критерії. *Оптимальним* прийнято вважати рішення, яке у повній мірі відповідає інтересам, намірам або цілям особи, що приймає рішення. Для того, щоб вирішити проблему узгодження цілей ОПР використовують функцію, що є відображеннями цілей ОПР, вираженими через мінімізацію або ж максимізацію деякої математичної функції на множині імовірних розв'язків. Якщо таких функцій декілька, то така задача вибору стає *задачею багатокритеріального вибору* [1].

1.2 Класифікація задач вибору

У загальному випадку, задачі прийняття рішень можуть бути представлені наступним набором інформації: $\langle T, A, K, X, F, G, D \rangle$, де T – постановка задачі, A – множина допустимих альтернатив, K – множина критеріїв, X – множина методів виміру уподобань, F – відображення множини допустимих альтернатив на множину критеріальних оцінок, G – система уподобань ОПР, D – вирішальне правило, що відображує систему уподобань. Зазвичай, задачі вибору класифікуються на основі виду відображення F оскільки відображення A у K може бути як детермінованим, так і недетермінованим, бути імовірнісним або ж невизначеним; потужності множини K або ж на основі типу системи G , оскільки уподобання можуть формуватися колективом або ж однією особою [2].

Першим типом задач є задачі прийняття рішень в умовах визначеності. Для вирішення задач прийняття рішень в умовах ризику інформація, що надається, є достатньою, кількісною та достовірною. Через це доцільно використовувати математичні методи, що дозволяють знаходити оптимальні рішення через математичну модель реального об'єкту. Для цього необхідно, щоб:

- задача була формалізованою, тобто існувала адекватна модель реального об'єкту;
- існував критерій оптимізації, що дозволяє оцінити якість альтернатив;
- цільова функція могла бути кількісно оцінюваною;
- задача мала ресурси оптимізації.

Наступним типом задач є задачі прийняття рішення в умовах невизначеності. У таких задачах інформація, що надається, є неточною, неповною або не кількісною, а формальні моделі або не існують, або занадто складні для побудови. У таких випадках необхідно залучати ОПР, адже за

допомогою їх уподобань можна сформулювати кількісні дані, на основі яких можна робити вибір альтернативи.

Ще одним типом задач є задачі прийняття рішення в умовах ризику. Для цих задач імовірні результати можуть бути описані як деякий імовірнісний розподіл. Для цього необхідно використати ОПР або ж зібрати статистичні дані для побудови розподілу. Для таких задач характерно використовувати теорію багатовимірної корисності для знаходження рішення [2].

1.3 Постановка задачі багатокритеріального вибору

У складних випадках доводиться працювати не з однією, а з багатьма функціями такого роду. Якщо ж явище розвивається в динаміці послідовно, тоді для оцінювання кожного з етапів необхідно вводити окрему функцію та враховувати декілька функціональних показників.

Задачу прийняття рішення в умовах багатокритеріальності можна сформулювати наступним чином: необхідно знайти рішення задачі, яка описується наступною функцією:

$$(f_1(x), f_2(x), \dots, f_n(x)) \xrightarrow{x \in D} \min, \quad (1.1)$$

$$D = \{x \in E^m \mid g_j(x) \leq 0, j = \overline{1, J}\},$$

де $f_i(x), i = \overline{1, n}$ – критерії, що оптимізуються, D – множина допустимих значень x ; $g_j(x), j = \overline{1, J}$ – функції обмежень; m – розмірність простору параметрів [1].

Окремо варто відзначити, що у рамках задачі багатокритеріального вибору множина імовірних рішень може мати дуже різну природу, що може ускладнити побудову моделі прийняття рішень. У таких випадках для вирішення задачі необхідно використовувати декілька ОПР, які є експертами у сферах, що стосуються задачі багатокритеріального вибору. Це дозволяє більш

точно побудувати множину імовірних рішень та обрати критерії, за якими альтернативи будуть оцінюватися [2, 3].

1.4 Теорія прийняття рішень

Проблема вибору принципу оптимальності є ключовою для *теорії прийняття рішень* (ТПР), оскільки вона пов'язана з визначенням властивостей оптимального рішення і дослідженням, за якими критеріями оптимальне рішення є кращим за інші. Проблема врахування критеріїв з'являється у той час, коли критерії є нерівнозначними, що призводить до потреби оцінки рівня їх впливу на вирішення задачі, тобто для визначення пріоритетності критеріїв. Це призводить до іншої проблеми – проблеми нормалізації критеріїв, адже зазвичай критерії мають різні одиниць вимірювань, що майже унеможлиблює їх порівняння між собою [4].

У ТПР розділяють два типи цілей, що стоять перед ОПР: якісні цілі та кількісні цілі. Якісні цілі – це цілі, результат яких або повністю задовольняє, або повністю не задовольняє такі цілі. У випадку, коли якісна ціль досягнута, нема потреби детально аналізувати, яким чином вона була досягнута. Математично, якісну ціль можна описати як деяку підмножину A множини очікуваних результатів. Будь-який результат $a \in A$ буде задовольняти ціль, а результат $a \notin A$ не буде задовольняти ціль. Тоді, кожному результату a буде відповідати деяка міра виконання цілі $\mu(a)$, $0 \leq \mu(a) \leq 1$. Якщо $\mu(a) = 1$ і $a \in A$ або ж $\mu(a) = 0$ при $a \notin A$, тоді ціль є чіткою.

Кількісними є цілі, що є результатом вибору на множині кількісних результатів, які описуються деякою дійсною функцією $f : A \rightarrow E^1$. У цьому випадку задача прийняття вибору трансформується в задачу знаходження оптимуму функції f на множині A . Якісна чітка ціль може бути зведена до кількісної цілі, якщо $f(a) = 1, a \in A$ або ж $f(a) = 0, a \notin A$. Якісні нечіткі ж цілі потребують збору додаткової інформації [5].

У сучасному світі великої популярності набули такі методи як теорія корисності, методи нечітких множин, методи векторної стратифікації та метод аналізу ієрархій Сааті. Це зумовлено тим, що ці методи є теоретично обґрунтованими та загалом відповідають умові універсальності, що є необхідним у вирішенні задач багатокритеріального вибору.

1.4.1 Теорія багатовимірної корисності

Теорія багатовимірної корисності дозволяє отримати функцію багатовимірної корисності для задач багатокритеріального вибору. Максимальним значенням такої функції є найкращим варіантом. Зазвичай функція багатовимірної корисності є адитивним або ж мультиплікативним поєднанням одновимірних функцій, що будуються за результатами опитувань експертів, що дозволяє зробити ранжування імовірних результатів без порівняння альтернатив. Також для функції корисності вважається, що критерії є взаємозалежними один від одного.

Аналіз рішення, отриманого завдяки теорії багатовимірної корисності, можна проводити за допомогою декількох послідовних кроків. Першим таким кроком буде попередній аналіз, що дозволяє вивести основу проблему, для якої треба знайти рішення і сформуванати множину імовірних рішень проблеми. Наступним проводиться структурний аналіз, завдяки якому проблема розглядається на якісному рівні та може бути структурована таким чином, що дозволяє розпочати побудову дерева рішень.

Надалі, проводиться аналіз невизначеності, задачею якого є задати імовірності для гілок дерева рішень. Необхідною частиною аналізу невизначеності є перевірка імовірностей на узгодженість. Після цього настає етап аналізу корисності, на якому визначаються оцінки корисності альтернатив завдяки побудові функції корисності.

Для формування функції корисності необхідно попередньо узгодити із ОПР вигляд даних, які вони надаватимуть для відображення своїх уподобань.

Перевагою такого методу є можливість оцінювання будь-якої кількості альтернатив з використанням отриманої функції, але це досить складно і при умові, що вхідна інформація є нестійкою, використання методів теорії корисності є малоефективним [6].

1.4.2 Теорія нечітких множин

Теорія нечітких множин (ТНМ) дозволяє обрати найкращу з альтернатив за критеріями за допомогою нечіткої множини. Формування нечіткої множини є простішим, аніж побудова функцій корисності. При використанні теорії нечітких множин, множина альтернатив являє собою множину U , а $\mu_a : U \rightarrow [0,1]$. У цьому випадку нечіткою множиною A в U буде називатися графік відображення μ_a , тобто множина вигляду $\{(u, \mu_a(u)) : u \in U\}$, при цьому значення $\mu_a(u)$ – це степінь приналежності елементу множини u до критеріїв A . Також теорія нечітких множин дозволяє використовувати вагові коефіцієнти та робити нечіткий логічний висновок на правилах визначення кращої альтернативи [6].

Основою підходу формалізації поняття нечіткої множини є узагальнення поняття приналежності. Нехай U – це універсальна множина, з елементів якого формуються множини, що розглядаються у задачі прийняття рішень. Тоді характеристична функція множини $A \subseteq U$ – це функція, значення якої визначають, чи належить $x \in U$ множині A . Значення цієї функції мають бінарний характер. Для того, щоб не обмежувати оцінку формулюваннями « x приймає значення приблизно від A до B », оцінка була розширена до необмежено багатозначної оцінки через заміну характеристичної функції, описаною вище, на функцію приналежності. Це дозволило вважати, що елемент x може не належати множині A , може бути у значному степені елементом A , може більш або менш належати множині A або ж бути елементом A [6].

Основною проблемою багатокритеріального вибору із застосуванням теорії нечітких множин є надання інформації про зв'язки між критеріями і моделі обчислення інтегральних оцінок [7].

1.4.3 Метод аналізу ієрархій

Метод аналізу ієрархій (МАІ) – це замкнута логічна конструкція, яка виконує аналіз складних проблем за допомогою простих правил. Метод заснований на парних порівняннях альтернативних варіантів за різними критеріями з використанням бальної шкали (як правило, від 1 до 9 балів) і надалі ранжируванні набору альтернатив за всіма критеріями і цілях. Зв'язки між критеріями враховуються шляхом побудови ієрархії критеріїв і застосування парних порівнянь для визначення важливості критеріїв і підкритеріїв.

Ідея методу полягає у групуванні елементів задачі за основними властивостями. Такий підхід дозволяє проводити цей процес таким чином, що угруповання або ж їх сумісні властивості формували наступний рівень системи. Надалі, ці елементи також можуть бути згруповані на основі їх подоби, що утворюватиме новий рівень системи. Цей процес може повторюватися до того моменту, поки не буде досягнута мета процесу прийняття рішень. Отримана в результаті угруповань систем називається ієрархією [8].

Нерівномірність впливу ознак на досягнення кінцевої мети призводить до необхідності визначення інтенсивності впливу ознак на кінцеву оцінку альтернативи. Визначення пріоритетів ознак щодо мети зводиться до послідовності задач визначення пріоритетів ознак для кожного рівня, а кожна така задача - до послідовності попарних порівнянь.

Якість альтернативи у загальному випадку визначається ієрархічною системою векторів:

$$y^{(j-1)} = \{y_i^{(j-1)}\}_{i=1}^{n^{(j-1)}}, j \in [2, m] \quad (1.2)$$

де $y^{(j-1)}$ – вектор критеріїв на $(j-1)$ -ому рівні ієрархії, за компонентами якого оцінюється якість ознак альтернативи на j -ому рівні, m – кількість рівней ієрархії, $n^{(j-1)}$ – кількість оцінюваних ознак $(j-1)$ -го рівня ієрархії. Чисельні значення n критеріїв $y^{(1)} = n$ першого рівня ієрархії задані для даної альтернативи, що дозволяє стверджувати, що $n^{(1)} = n$ и $n^{(m)} = 1$.

Важливість, або ж значимість кожного компонента критерію $(j-1)$ -го рівня при оцінці k -тої ознаки j -го рівня може бути охарактеризована коефіцієнтом пріоритету, сукупність яких являє собою систему векторів пріоритету:

$$p^{(j-1)} = \{p_k^{(j-1)}\}_{k=1}^{n^{(j-1)}}, j \in [2, m] \quad (1.3)$$

Виходячи з цього, постановка задачі багатокритеріального вибору вибору може бути сформульована наступним чином: необхідно знайти аналітичну оцінку $y^{(m)}$ и якісну оцінку ефективності кожного альтернативного рішення у вигляді загального критерію прийняття рішення і обрати найкращу [8, 9].

Недоліками МАІ є те, що збір даних для підтримки прийняття рішення здійснюється головним чином за допомогою процедури парних порівнянь, результати якої можуть бути суперечливими, внаслідок чого виникає необхідність в повторному перегляді даних для мінімізації протиріч та що у рамках методу аналізу ієрархій немає коштів для перевірки достовірності даних

1.5 Порівняльний аналіз методів прийняття рішень

Описані вище методи являють собою методи, що дозволяють враховувати багатокритеріальність і невизначеність, що робить їх цікавими для дослідження. Серед таких методів перевага зазвичай надається методам, що дозволяють декомпозицію та використовують методи теорії очікуваної корисності. Це пов'язане із тим, що саме такі методи є універсальними, враховують багатокритеріальність вибору в умовах невизначеності з

дискретною чи неперервною множиною альтернатив та простоти підготовки і аналізу інформації від ОПР.

На основі опису методів прийняття рішень можна виділити ряд критеріїв, за якими можна оцінити ефективність та результативність кожного з методів. До таких критеріїв доцільно віднести точність, простоту реалізації, врахування пріоритетності критеріїв, врахування оцінок експертів та можливість врахування прогнозування. Результати порівнянь наведені у таблиці 1.1

Таблиця 1.1 Результати порівнянь методів прийняття рішень

Назва методу	Точність	Простота	Врахування пріоритетності	Врахування оцінок експертів	Можливість врахування прогнозування
Теорія багатовимірної корисності	ні	ні	ні	так	ні
Теорія нечітких множин	так	ні	ні	так	так
Метод аналізу ієрархій	так	так	так	ні	ні

Було прийнято рішення розробити інформаційну технологію шляхом модифікації методу Сааті, яка б дозволила врахувати не тільки результати прогнозування часових рядів, які описують явище або об'єкт, але і в експертних оцінках особи, що приймає рішення.

До того ж, метод аналізу ієрархій є перспективним методом дослідження та прогнозування часових рядів та може претендувати на універсальність, оскільки він є досить простим для реалізації, враховує експертні оцінки та пріоритетність критеріїв и надає можливість врахування прогнозування.

Недоліком методу аналізу ієрархій є точність оскільки метод не передбачає перевірки достовірності даних, що може бути виправлено завдяки проведенню попереднього аналізу часового ряду для оцінки його типу та дослідження його показників.

1.6 Поняття часового ряду

Часовий ряд, або ж ряд динаміки – це послідовність упорядкованих у часі чисельних показників, які характеризують рівень стану та змінюваності явища, що є предметом досліджень. Таким чином, задачею аналізу часового ряду є дослідження та виявлення природи точок даних для розуміння, через що вони виникли та спроба побудови прогнозу. Прогнозування часового ряду, в свою чергу, передбачає побудову моделі, що дозволяє передбачати майбутні події до того, як дані буде вимірено, базуючись на інформації про минулі та сучасні події.

Зазвичай, часовий ряд складається з наступних показників: *тренду*, що є загальною довготривалою тенденцією зміни часового ряду, що полягає в основі його динаміки, *сезонної варіації*, яка є короткостроковим та регулярно виникаючим коливанням значень часового ряду навколо тренду та *циклічних коливань*, що характеризують цикл активності ряд.

1.7 Класифікація часових рядів

Часові ряди можуть бути класифіковані згідно своїх ознак. Таким чином, розрізняють класифікації за часом, за відстанню між датами та по змісту показників [10].

1.7.1 Класифікація за часом

При класифікації часових рядів за часом, виділяють *інтервальні* та *моментні* ряди.

Інтервальний ряд – це послідовність, в якій рівень явища відносять до результату, накопиченого або знову триманого за певний інтервал часу. Якщо ж рівень ряду характеризує досліджуване явище в конкретний момент часу, то сукупність рівнів утворює моментний ряд.

Важлива відмінність моментних рядів від інтервальних полягає в тому, що сума рівнів інтервального ряду дає цілком реальний показник - загальний випуск продукції за рік (п'ятиріччя, десятиріччя), загальні витрати робочого часу, загальний обсяг продажів акцій і т.д., сума ж рівнів моментного ряду іноді і підраховується, але реального змісту, як правило, не має [10].

1.7.2 Класифікація за відстанню між датами

За відстанню між датами показників розрізняють *повні* та *неповні* часові ряди. Повні ряди мають місце, коли дати реєстрації або закінчення періодів слідує один за одним з рівними інтервалами, неповні - коли принцип рівних інтервалів не дотримується [10].

1.7.3 Класифікація за змістом показників

За змістом показників часові ряди розділяються на ряди *окремих* та *агрегованих* показників.

Окремі показники описують явище ізольовано. Агреговані ж показники засновані на окремих показниках, що дозволяє вивчати явище комплексно [10].

1.8 Характеристики часових рядів

При вивченні часових рядів доцільно використовувати імовірностатистичні моделі. При цьому, часовий ряд y_t розглядається як дискретний випадковий процес, а отже, у нього можуть бути наступні характеристики:

математичне сподівання $y_t : E[y_t]$ дисперсія $y_t : D[y_t] = \sigma_t^2$ та автокореляційна функція.

Усі випадкові процеси можна умовно розбити на два види: *стаціонарні* та *нестаціонарні*.

Стаціонарний процес - процес, імовірнісні характеристики якого залишаються незмінними в часі. Під стаціонарними рядами розуміються випадкові процеси y_t , характеристики яких не змінюються з часом t , тобто інваріантні щодо тимчасових зрушень $t \rightarrow t + T$, $y_t \rightarrow y_{t+1}$ при будь-якому фіксованому T . Умовою стаціонарності у вузькому значенні є інваріантність вимірної щільності ймовірності щодо тимчасового зсуву τ .

Формально, випадковий процес з фіксованим середнім значенням та дисперсією називають стаціонарним, якщо для усіх t та $t - s$, ($k = 0, 1, 2, \dots$) виконуються наступні умови:

- математичне очікування не залежить від часу.

$$E[y_t] = E[y_{t-s}] = \mu = \text{const}; \quad (1.5)$$

- дисперсія залишається незмінною протягом усього часового інтервалу, на якому розглядається процес:

$$E\{[y_t - \mu]^2\} = E\{[y_{t-s} - \mu]^2\} = \sigma_y^2 = \text{const}; \quad (1.6)$$

- автоковаріація залишається незмінною у часі для усього часового проміжку і залежить лише від обраних моментів часу t та $t - s$

$$E\{[y_t - \mu][y_{t-s} - \mu]\} = E\{[y_{t-j} - \mu][y_{t-j-s} - \mu]\} = \gamma_s = \text{const} \quad (1.7)$$

Процес вважається нестаціонарним, якщо не виконується хоча б одне з умов (1.5) - (1.7). Далеко не завжди значення часового ряду формуються тільки під впливом будь-яких чинників. Особливий інтерес викликають процеси, які знаходяться в так званому «перехідному» режимі, тобто процеси, які, по суті, стаціонарні, але на досліджуваному проміжку часу проявляють властивості нестаціонарного часового ряду, що, в свою чергу, пояснюється нечіткими для

стаціонарного режиму початковими умовами. У ситуаціях, коли тимчасової ряд формується під впливом деякого набору випадкових і невідповідних факторів, аналіз окремих часових рядів має величезне значення [11].

Це необхідно для правильної ідентифікації моделей, які будуються по інформації про досліджуваних процесах (векторні авторегресії, моделі корекції похибок, динамічні моделі з розподіленими запізненнями і т.п.). Беручи до уваги складні умови протікання часових рядів, значну сукупність зовнішніх і внутрішніх факторів, які на нього впливають, з широкого класу нестационарних часових рядів можна виділити так звані випадкові процеси з довгою пам'яттю, для яких лише в кінці ХХ століття були запропоновані адекватні математичні моделі [11].

Необхідно відзначити, що поняття довгою пам'яті проміжне між поняттями пам'яті короткою (стаціонарні ряди) і пам'яті нескінченної (нестационарні ряди). Це зауваження набуває особливого значення, оскільки дуже часто важко провести жорстке розмежування між цими типами поведінки часових рядів, особливо в випадку розгляду економічних даних або інформаційних потоків в технічних системах.

Для ряду з короткою пам'яттю, «ефект шоку» не впливає на поведінку аналізованого ряду в довгостроковому періоді. І навпаки, для ряду, який характеризується нескінченною пам'яттю, ефект шоку позначається на всіх без винятку майбутніх його значеннях. У проміжному випадку наявність ефекту довгою пам'яті призводить до вкрай тривалим, але не перманентним наслідків шоку. Ці наслідки не є постійними в тому сенсі, що рано чи пізно ряд повернеться до свого рівноважного стану (такий стан часто називають «нормальним» або «природним»), тоді як ряд з нескінченною пам'яттю ніколи не повернеться до рівноважного стану після шоку [11].

1.9 Компоненти часових рядів

Пізнання закономірностей змін у часі - складна і трудомістка процедура дослідження, так як будь-який досліджуване явище формує безліч факторів, що діють в різних напрямках. За характером безпосереднього впливу ці фактори можуть бути розділені на дві групи. До першої групи належать фактори, що визначають основну тенденцію динаміки (зростання або зниження рівнів).

Друга група факторів, що викликає випадкові коливання, відхиляє рівні від тенденції то в одному, то в іншому напрямку. наприклад, тенденція динаміки врожайності пов'язана з прогресом агротехніки, з зміцненням економіки даної сукупності господарств, вдосконаленням організації та управління виробництвом. коливання врожайності викликана чергуванням сприятливих по погоді і несприятливих років, циклами сонячної активності, коливаннями у розвитку шкідливих комах і хвороб рослин.

При статистичному вивченні динаміки необхідно чітко розділяти два основних її елементи - тенденцію і коливання, щоб дати кожному з них кількісну характеристику за допомогою спеціальних показників.

Основною тенденцією, або трендом, називається характеристика процесу зміни явища за тривалий час, звільнена від випадкових коливань, створюваних другою групою факторів. На відміну від варіації явищ в просторової сукупності, вимірюваної за відхиленнями рівнів для окремих одиниць сукупності від їх середньої величини, коливанням слід називати відхилення рівнів окремих періодів часу від тенденції динаміки (тренду).

2 ЗАСОБИ ПОКРАЩЕННЯ МЕТОДУ АНАЛІЗУ ІЄРАРХІЙ ДЛЯ АНАЛІЗУ ТА ПРОГНОЗУВАННЯ ЧАСОВИХ РЯДІВ

2.1 Забезпечення можливості порівняння рівнів часового ряду

Однією з умов правильного формування часових рядів є порівнянність рівнів, що утворюють ряд. Рівні ряду, що підлягають вивченню, повинні бути однорідні за економічним змістом і враховувати істота досліджуваного явища і мета дослідження. Статистичні дані, представлені у вигляді часових рядів, повинні бути порівняні з територією, колу охоплених об'єктів, одиницям вимірювання, моменту реєстрації, методикою розрахунку, цінами, достовірності.

Несумісність по території виникає в результаті змін кордонів країн, регіонів, господарств тощо. Для приведення даних до порівнянного виду проводиться перерахунок колишніх даних з урахуванням нових кордонів.

Повнота охоплення різних частин явища - найважливіша умова порівнянності рівнів ряду. Вимога однаковою повноти охоплення різних частин досліджуваного об'єкта означає, що рівні ряду за окремі періоди повинні характеризувати розмір того чи іншого явища по одному і тому ж колу входять до складу його частин. Наприклад, при характеристиці динаміки врожайності овочевих культур в регіоні по роках не можна в одні роки враховувати тільки сільськогосподарські підприємства, а в інші - всі категорії господарств, в тому числі присадибні ділянки сільських жителів і сади, городи городян.

При визначенні порівнюваних рівнів ряду необхідно використовувати єдину методику їх розрахунку. Особливо часто ця проблема виникає при міжнародних порівняннях. Несумісність показників, що виникає в силу неоднаковості застосовуваних одиниць виміру, сама по собі очевидна. З різницею застосовуваних одиниць виміру доводиться зустрічатися при вивченні динаміки: виробничих ресурсів, коли вони представляються то в вартісному, то

в трудовому обчисленні; енергетичних потужностей (кВт-рік, к.с.); атмосферного тиску і т.д.

Труднощі при порівнянні даних по моменту реєстрації виникають через сезонних явищ. Чисельність худоби в домашніх господарствах через економічної доцільності різна взимку і влітку, тому рівні при порівняно повинні ставитися до певної дати щорічно. При аналізі показників у вартісному вираженні слід враховувати, що з плином часу відбувається безперервна зміна цін. Причин у цього процесу безліч - інфляція, зростання витрат, ринкові умови (попит і пропозиція) і т.д. У зв'язку з цим при характеристиці вартісних показників обсягу продукції в часі має бути усунуто вплив зміни цін.

Для вирішення цього завдання кількість продукції, виробленої в різні періоди, оцінюють в цінах одного періоду, які називають фіксованими або в певних статистичних органах - порівнянними цінами. Широке використання в статистичних дослідженнях вибіркового методу вимагає враховувати достовірність кількісних і якісних характеристик досліджуваних явищ в динаміці. різна репрезентативність вибірки за періодами внесе суттєві похибки в величини рівнів ряду.

Однією з умов порівнянності рівнів інтервального ряду, крім рівності періодів, за які наводяться дані, є однорідність етапів, в межах яких показник підпорядковується одному закону розвитку. В цих випадках проводять періодизацію часових рядів, типологічну угруповання в часі. Всі вищеназвані обставини слід враховувати при підготовці інформації для аналізу змін явищ у часі (динаміці).

2.2 Ієрархія тенденцій та коливань

У попередніх розділах були розглянуті дві основні компоненти часового ряду - його тенденція і коливання окремих рівнів. При вирішенні конкретних завдань статистичного дослідження ці компоненти слід розділяти, вимірювати кожен з них окремо. У той же час при розгляді складних процесів на великих

інтервалах часу ми спостерігаємо ієрархію тенденцій і коливань: то, що для часу високого порядку, наприклад століття, виступає як коливання, на інтервалі часу нижчого порядку, наприклад трьох-п'яти років, може виступати як тенденція.

Наприклад, існує 10-11-річна циклічна коливання сонячної активності, одним з показників якої служать числа Вольфа W (число груп сонячних плям, помножене на 10, плюс число окремих плям). За 100 років відбувається в середньому 9 або 10 циклів коливань. Але якщо розглядати помісячні дані про числах Вольфа за 2-3 роки фази зниження активності Сонця, то саме це зниження можна вважати вже не частиною коливання, а тенденцією, на тлі якої відбуваються хаотичні, випадкові коливання.

Є «вищий» рівень динаміки температур - її тенденція до підвищення, в основному в результаті антропогенного впливу – зростання викидів продуктів спалювання палива в атмосферу. Це повільна тенденція зростання середньорічних температур приблизно на 0,03 градуса за рік.

На тлі цієї тенденції середньорічні температури окремих років коливаються в середньому на 2 - 3 градуси. Усередині кожного року на середніх широтах відбуваються коливання середніх температур місяців – циклічні сезонні коливання, які, однак, для температури в окремі дні виступають як тенденція зниження температури восени та її зростання навесні.

Близько цих тенденцій середньодобові температури коливаються в основному хаотично, з огляду на зміни холодних і теплих повітряних мас, тобто циклонічної і антициклонічною динаміки атмосфери. Але якщо спуститися на нижній рівень часу і розглядати температуру повітря в окремі години доби, то ми побачимо нові, дрібномасштабні циклічні коливання часових температур: з ранку і до 13-14ч температура має тенденцію зростання, а до вечора - тенденцію зниження зважаючи денного нагрівання повітря сонячним світлом і охолодження вночі. Але в окремі годинки температура коливається близько

цих «тенденцій» в залежності від хмарності, вітру, дощу і т.д., іноді за півгодини температура повітря може змінитися на 5 - 10 градусів.

У зв'язку з цим знання статистичних методів і вивчення тенденцій і коливань для економіста-менеджера, для статистика-аналітика мають величезне значення.

2.2.1 Періодизація часового ряду

Періодизація розвитку, тобто розчленування періоду розвитку у часу на однорідні етапи, в межах яких показник підпорядковується одному закону розвитку, це, по суті, типологічна угруповання на часу. Періодизація може здійснюватися кількома методами.

Історичний метод. Періодизація здійснюється на основі «Узаконеної» структури динаміки. При цьому звертають увагу на значущі дати і події, а саме: час прийняття управлінських рішень по даному показнику, зміну господарського механізму, зміну керівництва, війни і т.п. Недоліком цього методу є те, що точні часові межі періодів шляхом теоретичного аналізу вдається напівчирити вкрай рідко.

Метод паралельної періодизації. Ідея цього методу полягає в наступному. Нехай y - аналізований показник, розгорнутий в динамічний ряд $\{y_t\}$, де y_t - значення рівня ряду в момент (інтервал) часу t . Можливо, існує показник x , якому відповідав би динамічний ряд $y_t = \{Xt\}$, який визначає поведінку досліджуваного показника, тоді в ролі «однорякісності» періодів розвитку потрібно взяти періоди x .

Методи багатовимірного статистичного аналізу. Часто потрібно виділити однорякісні періоди в розвитку явищ чи процесів, отримати адекватне відображення яких за допомогою одного лише показника важко. До таких належать, зокрема, здоров'я населення, розвиток сільськогосподарського виробництва та ін. Очевидно, що навіть такі комплексні показники, як смертність, тривалість життя, захворюваність, недостатні для еквівалентного

опису такого складного, інтегрованого явища, як здоров'я. Необхідна система показників, при якій:

- враховується різноманіття аспектів явища;
- амортизується спотворення, недостовірних та неточних статистичних даних;
- наявність великої кількості показників підвищує обґрунтованість статистичних висновків, т. е. забезпечується надійність їх екстраполяції.

Ідеальним виходом є використання безлічі, що включає всі характеристики процесу. Однак це не завжди можливо за різними причин, найчастіше внаслідок недоступності статистичної інформації. На основі комплексних динамічних рядів (системи показників) періодизація реалізується методом багатовимірної середньої та методами факторного та кластерного аналізу.

Одноякісність рівнів часового ряду означає, що в межах усього досліджуваного періоду, до якого відносяться рівні, повинна бути проведена типологічна угруповання. Після виділення однорідних груп можуть використовуватися і аналізуватися рівні ряду. Ця вимога може бути сформульовано як забезпечення порівнянності по структурі сукупності, для чого зазвичай застосовується стандартна, нормативна структура.

2.2.2 Показники, що характеризують динаміку часового ряду

Щоб побудувати систему показників, що характеризують тенденцію динаміки, потрібно відповісти на питання: які риси, властивості цієї тенденції потрібно виміряти і виразити в статистичних показниках. Очевидно, що необхідно визначати величину змін рівня як в абсолютному, так і у відносному вираженні. Далі важливо встановити, чи є зміна рівномірним або нерівномірним, прискореним (уповільненим). Нарешті, нас цікавить вираз

тенденції в формі деякого досить простого рівняння, найкращим чином апроксимує фактичну тенденцію динаміки.

Абсолютна зміна рівнів - в даному випадку його можна назвати абсолютним приростом - це різниця між порівнюваним рівнем і рівнем більш раннього періоду, прийнятим за базу порівняння. Якщо ця база - безпосередньо попередній рівень, показник називають ланцюговим, якщо за базу взято, наприклад, початковий рівень, показник називають базисним.

Формули абсолютного зміни рівня:

$$\Delta_{\text{ц}} = y_n - y_{n-1} \quad (2.1)$$

$$\Delta_0 = y_n - y_1 \quad (2.2)$$

Якщо абсолютна зміна рівнів негативна, тоді зміну рівня прийнято вважати абсолютним скороченням.

Прискорення – це різниця між абсолютною зміною за даний період та абсолютною зміною за попередній період однакової довготи.

$$\Delta_i = \Delta_i - \Delta_{i-1} \quad (2.1)$$

Показник прискорення використовується лише у ціпному варіані, але не в базисному.

2.3 Метод моделювання і прогнозування часових рядів з довгою пам'яттю

Довга пам'ять, або довгострокова залежність - це властивість, яке описує кореляційну структуру високого порядку тимчасового ряду. В разі, якщо ряд характеризується довгою пам'яттю, то залежність існує навіть між далеко віддаленими один від одного в часі спостереженнями. оскільки довга пам'ять створює в моделі для середнього рівня ряду нелінійну залежність в перші моменти розподілу, то вона генерує в динаміці ряду потенційно корисний для прогнозування компонент. Довга пам'ять означає наявність значущих автокорреляцій між віддаленими спостереженнями, тому її присутність

призводить до того, що минулі результати можуть допомогти в прогнозі майбутньої поведінки ряду [12].

Необхідно відзначити, що поняття довгою пам'яті проміжне між поняттями пам'яті короткою і пам'яті нескінченною. Це зауваження набуває особливу важливість, оскільки часто складно провести жорстке розмежування між цими типами поведінки часових рядів, особливо в разі розгляду економічних даних або інформаційних потоків в технічних системах. В випадку ряду з короткою пам'яттю ефект шоку не впливає на поведінку аналізованого ряду в довгостроковому періоді. І навпаки, для ряду з нескінченною пам'яттю ефект шоку позначається на всіх без винятку майбутніх його значеннях. У проміжному випадку наявність ефекту довгострокової пам'яті призводить до вкрай тривалим, але не перманентним наслідками шоку. Ці наслідки не постійні в тому сенсі, що рано чи пізно ряд повернеться до свого рівноважного рівня (такий рівень часто називають «нормальним» або «природним»), тоді як ряд з нескінченною пам'яттю ніколи не повертається до рівноважного рівня після того, що сталося шоку [14].

У зв'язку з тим, що довга пам'ять є особливою формою нелінійної динаміки, вона ставить під сумнів лінійне моделювання і вимагає розробки на теоретичному рівні нових нелінійних моделей оцінки поведінки тимчасових рядів, у яких ідентифікується наявність довгої пам'яті.

2.4 Методи класифікації часового ряду

Поширений підхід при класифікації об'єктів будь-якої природи, в тому числі і тимчасових рядів, полягає у виділенні деякого ряду інформативних ознак [15, 16]. Проведений аналіз показав, що інформативні ознаки, які використовуються при класифікації часових рядів, відрізняються різноманітністю і визначаються поставленої автором метою класифікації. У теорії випадкових процесів найбільш загальною класифікацією є класифікація

«за часом» і «за станом» [16]. Тому ознаки можна виділити чотири класи процесів:

- процеси з дискретними станами та дискретним часом;
- процеси з дискретними станами і безперервним часом;
- процеси з безперервними станами і дискретним часом;
- процеси з безперервними станами і безперервним часом.

На практиці, як правило, розглядають одновимірний або двовимірний закони розподілу випадкового процесу, оскільки вони містять достатній обсяг інформації про властивості випадкового процесу, а приріст кількості інформації при використанні імовірнісних характеристик вищого порядку виявиться незначним.

Також в більшості випадків кореляційна функція є досить повною характеристикою процесу, тому найчастіше обмежуються виявленням стаціонарності процесу в широкому сенсі. Структуру випадкового процесу можна встановити по кореляційній функції або за відомою щільності розподілу ймовірностей. В залежності від типу законів розподілу можна виділити нормальні, рівномірні, пуассоновским і інші випадкові процеси. Відхилення від класичної форми розподілу говорить про нестаціонарності процесу. За однією реалізації обмеженої довжини важко з достатньою точністю судити про закон розподілу часового ряду, і в більшості прикладних випадків аналізу дослідник не має в своєму розпорядженні інформації про вид функції розподілу [16].

Для нестаціонарних процесів характерна зміна в часі їх статистичних характеристик, тому при виконанні класифікації це можна врахувати. З точки зору такого підходу, як звичайно виділяють процеси, які мають змінне в часі середнє значення; змінне в часі середнє значення квадрата; змінні в часі середню та середню значення квадрата; змінне в часі частотну структуру. Подібна класифікація відображає зміну в часі оцінок імовірнісних характеристик. Можна ще виділити кілька підходів до класифікації процесів.

Найбільш загальний підхід до класифікації як стаціонарних, так і нестаціонарних процесів пов'язаний з них безперервним або дискретним уявленнями [16]. У прикладних випадках враховується специфіка завдань, вирішення яких повинна передувати класифікація спостережуваних процесів.

3 ДОСЛІДЖЕННЯ ЕФЕКТИВНОСТІ СИСТЕМИ

3.1 Загальний опис системи

Система складається з 2 основних компонент: архіву даних та програмної частини, що забезпечує зчитування даних та виконує обчислення.

В якості формату даних було обрано csv формат, оскільки він є одним з найпопулярніших структурованих форматів для збереження даних та є досить легким для зчитування програмою.

Програмна частина написана на мові програмування Java. Завдяки цьому програма може виконуватися майже на будь-якому пристрої, де встановлена віртуальна машина Java. Для роботи з csv форматом використовується бібліотека univocity, що дозволяє швидко та ефективно розібрати дані та організувати їх у клас.

3.2 Загальний опис прикладу

Для проведення оцінки роботи запропонованої системи буде описано тестовий приклад і обчислено оцінки діяльності математичного апарату.

Перш за все заповнюється таблиця даних. Дані отримуються з п'яти сайтів погоди. Через те, що деякі сайти змінюють показники погоди перед збереженням в архів, доцільно записувати дані зранку та ввечері, щоб фіксувати імовірні зміни показників.

Для перевірки отриманих даних на достовірність, вводиться шосте джерело - міжнародна метеостанція і дані кожного з джерел порівнюються із шостим. Результати порівняння

Приклад даних, отриманих за один день:

```
Gismeteo,01.01.2018,Morning,0,1,5,4,0,2S,4S,4S,744,745,745,744,1.00,1.00,0.84,0.84,,,,,
Fog,Fog,Cloudy,Cloudy,0,0,3,2,
Meta,01.01.2018,Morning,0,4,3,3,2.2S,3.8S,4.3S,4.8S,761,760,760,760,0.98,0.93,0.89,0.87
,0.03,0.02,0.03,0.04,Cloudy,Cloudy,Cloudy,Cloudy,-2,0,-1,-1,
Meteo,01.01.2018,Morning,0,1,4,4,1.8S,2.1S,3.8S,4.3S,743,743,743,743,0.96,0.97,0.83,0.8
4,0.02,0.02,0.02,0.02,Cloudy,Cloudy,Cloudy,Cloudy,-2,-2,1,0,
```

Rp5,01.01.2018,Morning,0,1,3,3,2SE,3S,4S,5S,743,761,742,742,0.98,1.00,0.89,0.93,,,,,Clo
 udy,Cloudy,Cloudy,Cloudy,,, -1, -1,
 Sinoptik,01.01.2018,Morning,0,2,4,4,0,4S,4S,4.5S,744,744,745,744,1.00,1.00,0.89,0.86,0.
 00,0.00,0.02,0.02,Cloudy,Cloudy,Cloudy,Cloudy,1,2,4,4,
 ACTUAL,01.01.2018,,1,1,4,4,2.9SE,3S,5.1SW,5.8SE,758,760,759,758,0.91,0.94,0.87,0.89,,,,,
 ,Cloudy,Cloudy,Cloudy,Cloudy,,,,,
 Gismeteo,01.01.2018,Evening,0,1,3,3,0,2S,4S,4S,744,745,746,745,1.00,1.00,0.91,0.91,,,,,
 Fog,Fog,Cloudy,Cloudy,0,0,1,1,
 Meta,01.01.2018,Evening,0,1,3,3,2.2S,3.8S,3.9S,4.8S,761,760,760,760,0.98,0.98,0.92,0.89
 ,0.03,0.02,0.02,0.04,Cloudy,Cloudy,Cloudy,Cloudy,-2,0,0,-1,
 Meteo,01.01.2018,Evening,0,1,4,4,1.8S,2.1S,4.1S,4.1S,743,743,743,743,0.96,1.00,0.95,0.9
 2,0.02,0.02,0.02,0.02,Cloudy,Cloudy,Cloudy,Cloudy,-2,-2,1,1,
 Rp5,01.01.2018,Evening,0,1,3,3,2SE,3S,4S,5S,743,761,742,742,0.98,1.00,0.89,0.93,,,,,Clo
 udy,Cloudy,Cloudy,Cloudy,,, -1, -1,
 Sinoptik,01.01.2018,Evening,0,2,3,4,0,3SW,3S,4.4S,744,745,744,744,1.00,1.00,0.91,0.90,,
 ,, ,Cloudy,Cloudy,Cloudy,Cloudy,1,-1,-1,1,
 ACTUAL,01.01.2018,,1,1,4,4,2.9SE,3S,5.1SW,5.8SE,758,760,759,758,0.91,0.94,0.87,0.89,,,,,
 ,Cloudy,Cloudy,Cloudy,Cloudy,,,,,

Для дослідження системи зручно детально проаналізувати її поведінку спершу на даних, що є числовими, оскільки їх найпростіше візуалізувати та порівнювати.

3.3 Структура вхідних даних

Вхідні дані являють собою csv файл, у якому зберігаються кількісні та текстуальні дані про показники доби.

3.3.1 Стобець Name

Стовбець Name містить інформацію про назви альтернатив, що порівнюються. Він може приймати наступні значення:

- *Gismeteo*, що відповідає значенням критеріїв для першої альтернативи
- *Meta*, що відповідає значенням критеріїв для другої альтернативи
- *Meteo*, що відповідає значенням критеріїв для третьої альтернативи
- *Rp5*, що відповідає значенням критеріїв для четвертої альтернативи
- *Sinoptik*, що відповідає значенням критеріїв для п'ятої альтернативи

- *ACTUAL*, що відповідає значенням критеріїв для даних з міжнародної метеостанції

Таблиця 3.1–Приклад даних стовбця Name

Name
Gismeteo
Meta
Meteo
Rp5
Sinoptik

3.3.2 Стовбець Date

Стовбець Date містить дані про дати, коли знімалися виміри для кожної з альтернатив.

Таблиця 3.2 – Приклад даних стовбця Date

Date
01.01.2018
01.01.2018
01.01.2018
02.01.2018
02.01.2018

3.3.3 Стовбець Time

Стовбець Time містить дані про період доби, коли знімалися виміри для кожної з альтернатив. Він може приймати наступні значення:

- *Morning*, що відповідає ранку
- *Evening*, що відповідає вечору
- *NaN*, що є невідомим

Варто відзначити, що третє значення використовується лише для показників, для яких Name = *ACTUAL*.

Таблиця 3.3 – Приклад даних стовбця Time

Date
Morning
Morning
NaN
Evening
Evening

3.3.4 Стовбці Temperature_X

Стовбці Temperature_X містять інформацію про показники температури у різні періоди доби. Літера X відповідає за період доби. Одиницею виміру даних є градуси Цельсія.

Таблиця 3.4 – Приклад даних стовбців Temperature_X

Temperature_N	Temperature_M	Temperature_D	Temperature_E
0	1	5	4
0	4	3	3
0	1	4	4
0	1	3	3
0	2	4	4

3.3.5 Стовбці Wind_X

Стовбці Wind_X містять інформацію про показники вітру у різні періоди доби. Літера X відповідає за період доби. Одиницею виміру даних є метр за секунду.

Таблиця 3.5 – Приклад даних стовбців Wind_X

Wind_N	Wind_M	Wind_D	Wind_N
0	2S	4S	4S
2.2S	3.8S	4.3S	4.8S
1.8S	2.1S	3.8S	4.3S
2SE	3S	4S	5S
0	4S	4S	4.5S

3.3.6 Стовбці Pressure_X

Стовбці Pressure_X містять інформацію про показники вітру у різні періоди доби. Літера X відповідає за період доби. Одиницею виміру даних є одиниці ртутного стовпчика.

Таблиця 3.6 – Приклад даних стовбців Pressure_X

Pressure_N	Pressure_M	Pressure_D	Pressure_E
744	745	745	744
761	760	760	760
743	743	743	743
743	761	742	742
744	744	745	744

3.3.7 Стовбці Humidity_X

Стовбці Humidity_X містять інформацію про показники вологості у різні періоди доби. Літера X відповідає за період доби. Дані можуть приймати значення від 0 до 1.

Таблиця 3.7 – Приклад даних стовбців Humidity_X

Humidity_N	Humidity_M	Humidity_D	Humidity_E
1.00	1.00	0.84	0.84
0.98	0.93	0.89	0.87
0.96	0.97	0.83	0.84
0.98	1.00	0.89	0.93
1.00	1.00	0.89	0.86

3.3.8 Стовбці Precipitation_X

Стовбці Precipitation_X містять інформацію про показники імовірності опадів у різні періоди доби. Літера X відповідає за період доби. Дані можуть приймати значення від 0 до 1, або *NaN*, якщо альтернатива не надає інформації про виміри.

Таблиця 3.8 – Приклад даних стовбців Precipitation_X

Precipitation_N	Precipitation_M	Precipitation_D	Precipitation_E
NaN	NaN	NaN	NaN
0.03	0.02	0.03	0.04
0.02	0.02	0.02	0.02
NaN	NaN	NaN	NaN
0.00	0.00	0.02	0.02

3.3.9 Стовбці Cloudness_X

Стовбці Cloudness_X містять інформацію про показники хмарності у різні періоди доби. Літера X відповідає за період доби. Дані можуть приймати наступні значення:

- *Fog*, що означає «туман»
- *Cloudy*, що означає «хмарно»
- *Rain*, що означає «дощ»
- *Snow*, що означає «сніг»

- *Clear*, що означає «ясно»

Таблиця 3.9 – Приклад даних стовбців Cloudiness_X

Cloudness_N	Cloudness_M	Cloudness_D	Cloudness_E
Fog	Fog	Cloudy	Cloudy
Cloudy	Cloudy	Cloudy	Cloudy
Cloudy	Cloudy	Cloudy	Cloudy
Cloudy	Cloudy	Cloudy	Cloudy
Cloudy	Cloudy	Cloudy	Cloudy

3.3.10 Стовбці Feels_Like_X

Стовбці Feels_Like_X містять інформацію про показники температури, яка відчувається у різні періоди доби. Літера X відповідає за період доби. Одиницею виміру даних є градуси Цельсія. Дані можуть також приймати значення *NaN* у випадку, коли йдеться мова про показники міжнародної станції.

Таблиця 3.10 - Приклад даних стовбців Feels_Like_X

Feels_Like_N	Feels_Like_M	Feels_Like_D	Feels_Like_E
0	0	3	2
-2	0	-1	-1
-2	-2	1	0
1	-1	-1	-1
1	2	4	4

3.4 Побудова та оцінка ієрархії

Згідно обраного набору критеріїв для оцінки критеріїв, у загальному випадку кожна альтернатива має 28 критеріїв, що складають перший рівень ієрархії:

- показники нічної температури;
- показники ранкової температури;
- показники денної температури;

- показники вечірньої температури;
- показники нічної швидкості та направленості вітру;
- показники ранкової швидкості та направленості вітру;
- показники денної швидкості та направленості вітру;
- показники вечірньої швидкості та направленості вітру;
- показники нічного тиску;
- показники ранкового тиску;
- показники денного тиску;
- показники вечірнього тиску;
- показники нічної вологості;
- показники ранкової вологості;
- показники денної вологості;
- показники вечірньої вологості;
- імовірність нічних опадів;
- імовірність ранкових опадів;
- імовірність денних опадів;
- імовірність вечірніх опадів;
- показники нічної хмарності;
- показники ранкової хмарності;
- показники денної нічної хмарності;
- показники вечірньої хмарності;
- показники нічної температури що відчувається;
- показники ранкової температури що відчувається;
- показники денної температури що відчувається;
- показники вечірньої температури що відчувається;

Однак, варто відзначити що не всі з альтернатив, що було обрано для дослідження, надають усі з вищезначених критеріїв. Так, ресурси Gismeteo та Rp5 не надають інформації про імовірність опадів, а міжнародна метеостанція,

що була обрана для перевірки даних, не надає інформації про імовірність опадів та про температуру, що відчувається.

Для спрощення процедури прийняття рішення, доцільно сформувати другий рівень ієрархії, скориставшись ознаками подібності критеріїв. Так, другий рівень ієрархії буде мати 7 критеріїв:

- показники температури;
- показники швидкості та направленості вітру;
- показники тиску;
- показники вологості;
- імовірність опадів;
- показники хмарності;
- показники температури що відчувається;

Візуальне відображення отриманої ієрархії наведено на рисунку 3.1. Оскільки не всі альтернативи надають інформацію за критеріями, побудована ієрархія є неповнозв'язною, що візуально виглядає як порожній проміжок між лініями, що поєднують альтернативи з критеріями першого рівню ієрархії.

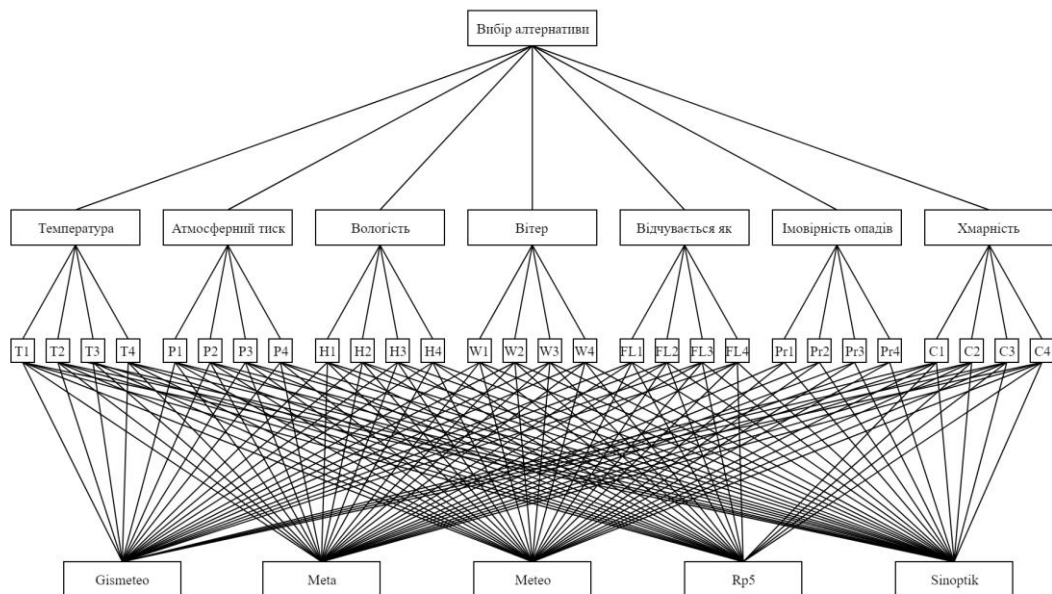


Рисунок 3.2 – Графік зміни показників ранкової температури сайту Gismeteo протягом січня

Надалі, використання методу аналіз ієрархій, спочатку необхідно сформулювати таблицю співвідношення критеріїв. У загальному випадку, критерії оцінюються відносно один одного за шкалою від одного до дев'яти. Оскільки у ієрархії, що була побудована, присутня неповнозв'язність та через той факт, що міжнародна метеорологічна станція не надає інформації про температуру, що відчувається, зрозуміло, що ці вплив цих двох критеріїв необхідно мінімізувати.

Мінімізацію впливу критеріїв на вибір найкращої альтернативи можливо провести вже на етапі оцінки важливості критеріїв відносно один одного. Для цього шкалу оцінки було модифіковано таким чином, що п'ять з семи критеріїв оцінюються за шкалою від 1 до 7, а критерії, вплив яких мінімізується, використовується оцінки 1/8 та 1/9. Отриманий результат наведений у таблиці 3.11.

Таблиця 3.11 – Результати оцінки важливості критеріїв

Критерії	Температура	Атмосферний тиск	Вологість	Вітер	Імовірність опадів	Хмарність	Відчувається як
Температура	1	1/4	2	1/4	8	1/5	9
Тиск	4	1	1/3	1/5	8	1/5	9
Вологість	1/2	3	1	1/6	8	1/3	9
Вітер	4	5	6	1	8	1/6	9
Імовірність опадів	1/8	1/8	1/8	1/8	1	7	9
Хмарність	5	5	3	6	1/7	1	9
Відчувається як	1/9	1/9	1/9	1/9	1/9	1/9	1

3.5 Аналіз результатів тестового прикладу на достовірність

За для прикладу виберемо дані за січень 2018. Задля візуалізації порівняння даних доцільно будувати графіки та діаграми. Кожен з графіків бути містити 3 типи значень – показник, що був знятий зранку (синя лінія), показник, що був отриманий ввечері (помаранчева лінія) та показник, який був наданий міжнародною станцією (зелена лінія).

Для загального опису процедури аналізу отриманих даних для подальшого розрахунку, доцільно візуально продемонструвати показання

погодних ресурсів, що знімалися вранці та ввечері та проаналізувати, наскільки отримані дані є достовірними.

3.5.1 Аналіз отриманих даних на прикладі показників з числовими значеннями

Для прикладу показника з числовими показниками, наведено результати показників температури повітря сайту Gismeteo за січень (рисунки 3.2 – 3.5). Температура вимірювалась у градусах Цельсія.

За допомогою візуальної оцінки можна відмітити, що графіки показників, що були отримані ввечері більш наближені до даних, які надає міжнародна метеостанція та що дані, що знімалися ввечері відрізняються у більшості випадків. Це зумовлено тим, що дані коригувалися перед збереженням у архів.

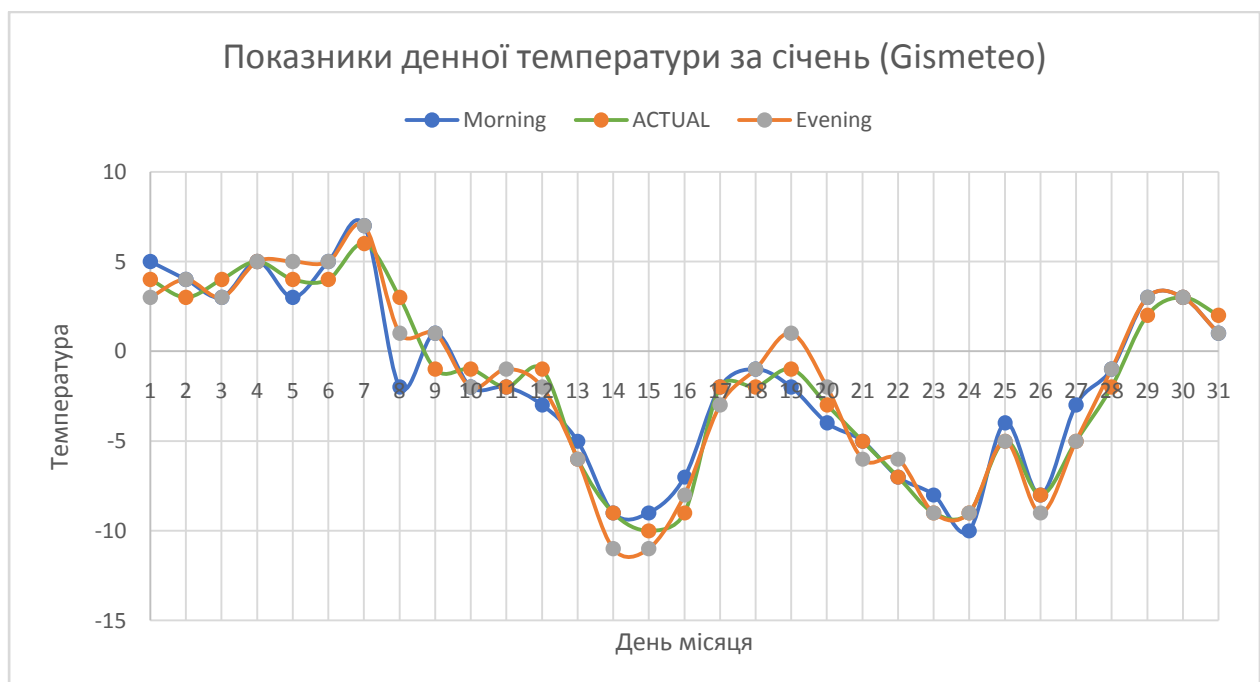


Рисунок 3.2 – Графік зміни показників нічної температури сайту Gismeteo протягом січня

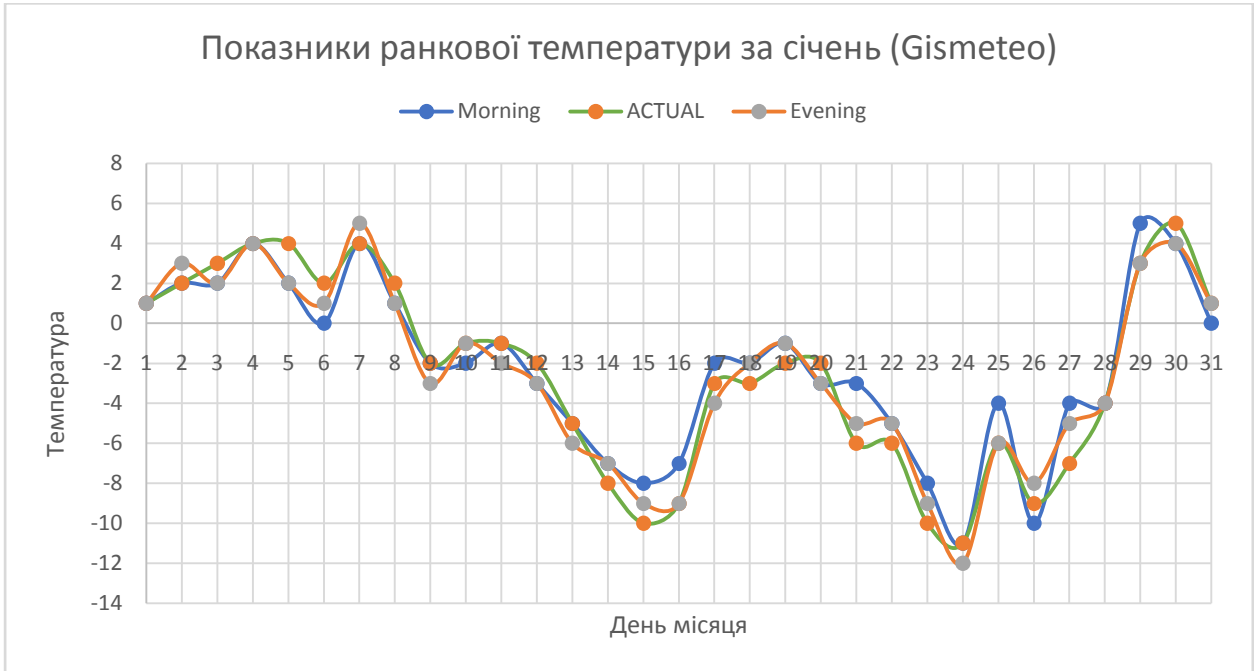


Рисунок 3.3 – Графік зміни показників ранкової температури сайту Gismeteo протягом січня

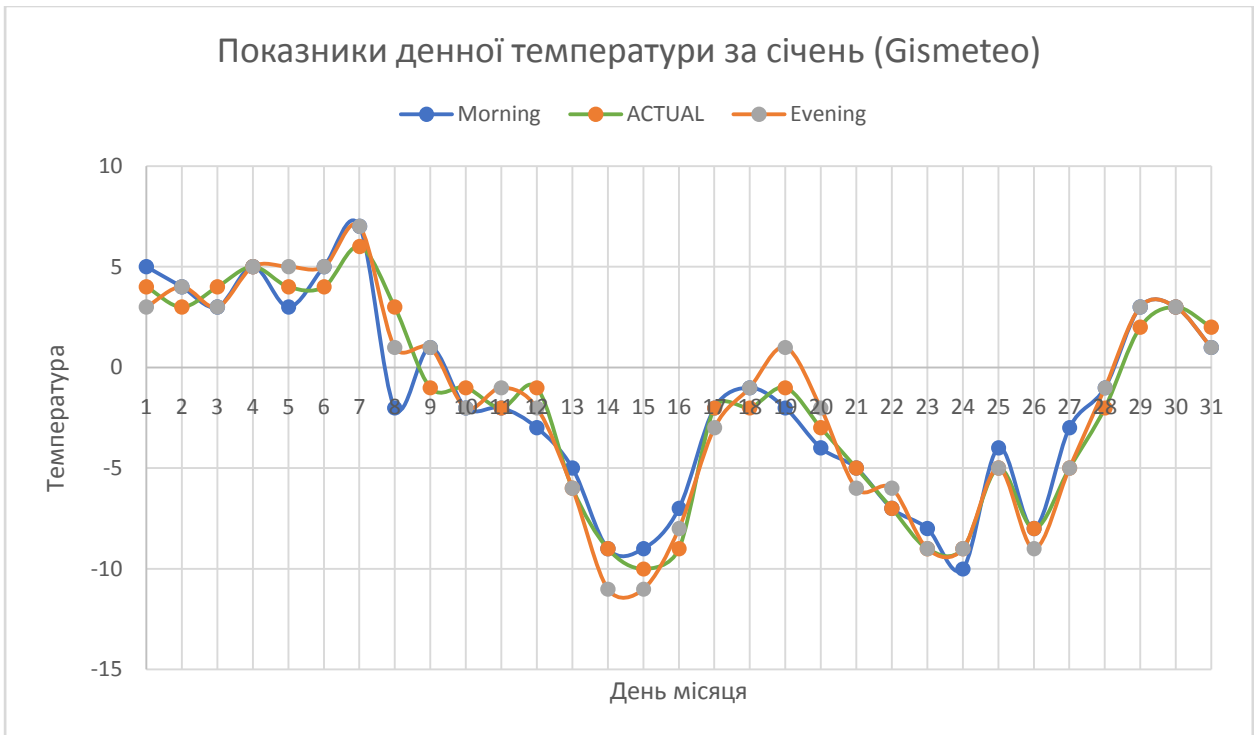


Рисунок 3.4 – Графік зміни показників ранкової температури сайту Gismeteo протягом січня

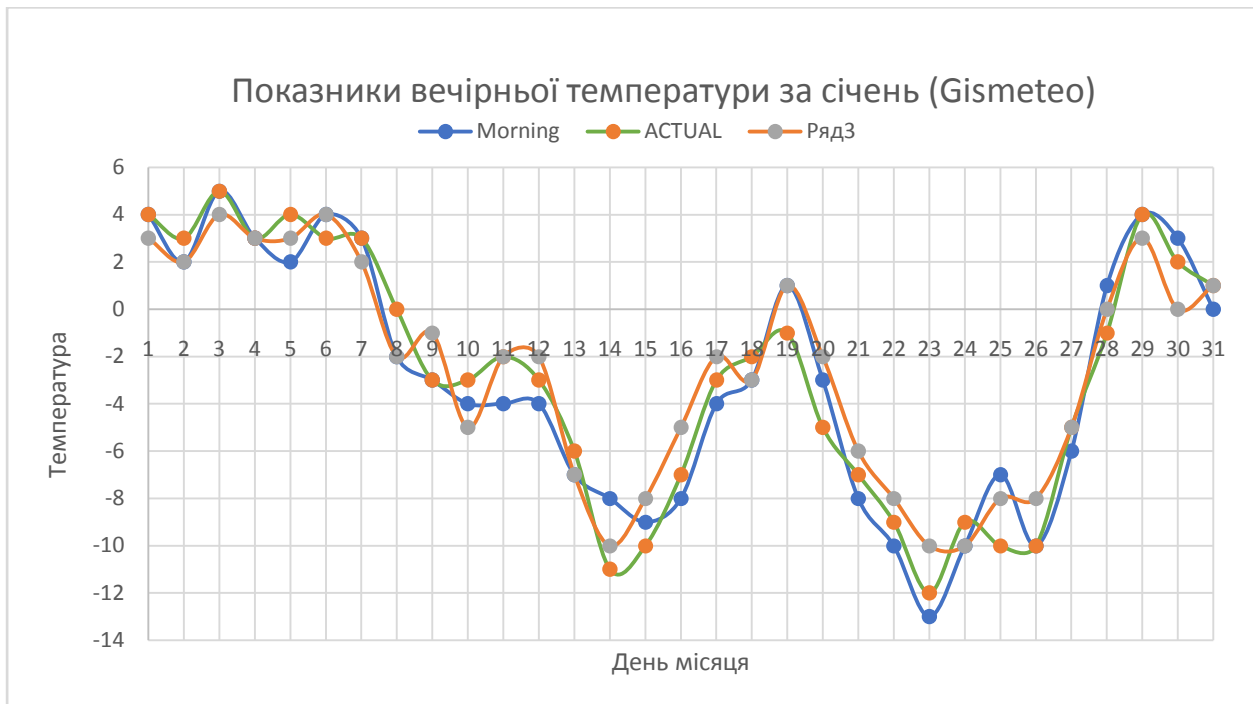


Рисунок 3.5 – Графік зміни показників ранкової температури сайту Gismeteo протягом січня

У той же час, варто відмітити що тенденція зміни показників не є трендом для усіх альтернатив. Цікаво, що ресурси Meta, Meteo та Rp5 не змінюють показників перед занесенням в архів, а ресурс Sinoptik змінює лише ранові, денні та вечірні показники.

Для подальшого аналізу, необхідно розрахувати абсолютні відхилення показників, що дозволить оцінити степінь наближеності показників сайту до показників, які надала міжнародна станція (таблиця 3.12).

Таблиця 3.12 – Результати розрахунку середніх абсолютних відхилень показників температури

Період доби	Середнє абсолютне відхилення
Показник нічної температури	1.113
Показник ранкової температури	0.936
Показник денної температури	0.919
Показник вечірньої температури	0.968

Згідно отриманих даних, найбільш достовірно ресурс Gismeteo надає інформацію про показники денної температури. Таким чином, середнє абсолютне відхилення другого рівня ієрархії буде складати 0.917.

3.5.2 Аналіз отриманих даних на прикладі показників з текстовими значеннями

Для прикладу показника з текстовими показниками обрано показники хмарності, надані сайтом Gismeteo у січні 2018 року. Для графічного відображення розподілу даних використовуються кругові діаграми. Приклад такої діаграми наведено на рисунку 3.6

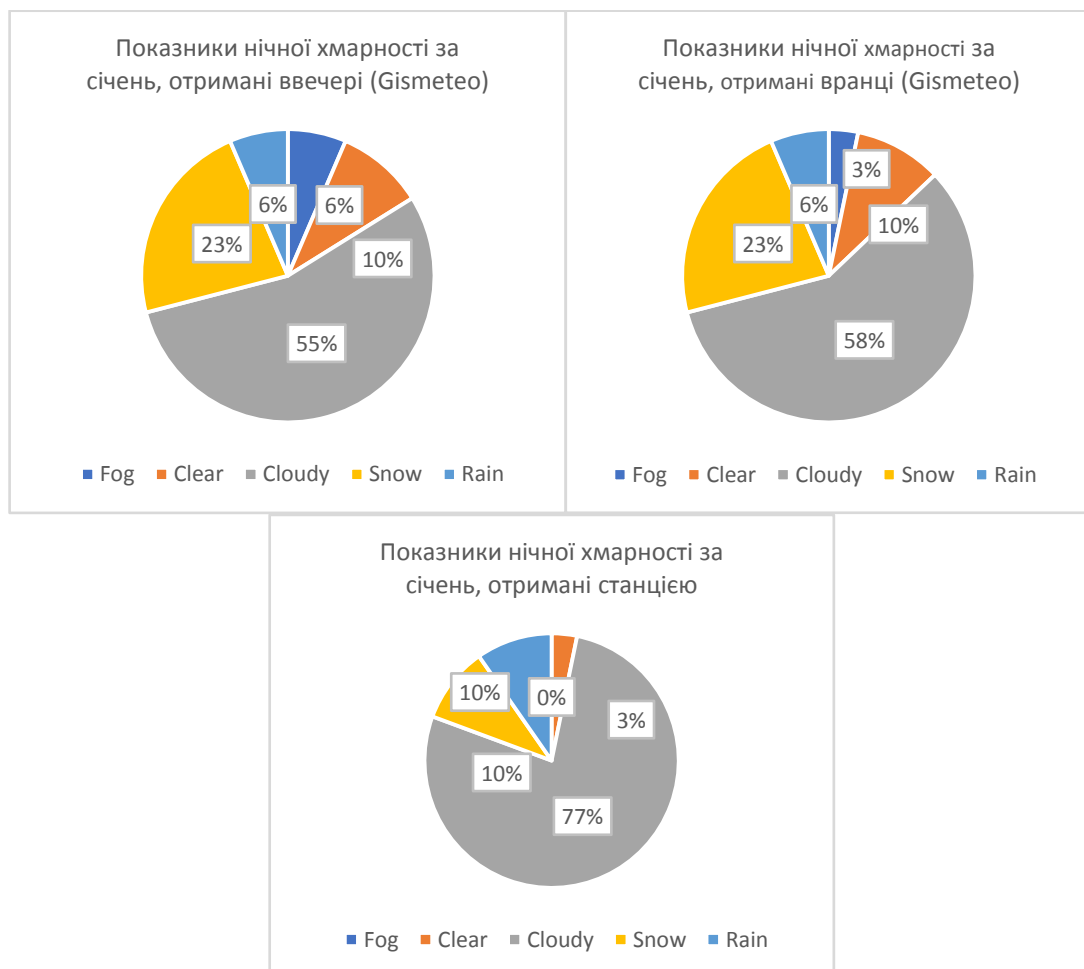


Рисунок 3.6 – Кругові діаграми показників нічної погоди, отримані з сайту Gismeteo протягом січня

Для оцінки таких даних використовувати середнє абсолютне відхилення неможливо, адже дані не є чисельними. Тому, для розрахунку вводяться коефіцієнти оцінки показників погоди відносно один одного. Для спрощення обрахунку, якщо показник хмарності співпадає, значення буде 0, якщо показники полярно різні, то значення стане 1. Коефіцієнти наведені у таблиці 3.13.

Таблиця 3.13 – Коефіцієнти оцінки показників хмарності

Хмарність	Ясно	Туман	Хмарно	Дощ	Сніг
Ясно	0	0.25	0.5	0.75	1
Туман	0.25	0	0.5	0.75	1
Хмарно	0.5	0.25	0	0.75	1
Дощ	1	0.75	0.5	0	0.25
Сніг	1	0.75	0.5	0.25	0

Таким чином, стає можливим розрахувати показники достовірності як середні абсолютні відхилення коефіцієнтів оцінки даних, адже текстові значення було приведено до числового типу. Результати розрахунків наведені у таблиці 3.14.

Таблиця 3.14 – Результати розрахунку середніх абсолютних відхилень коефіцієнтів хмарності

Показник	Середнє абсолютне відхилення
Нічна хмарність	0.294
Ранкова хмарність	0.278
Денна хмарність	0.25
Вечірня хмарність	0.209

Таким чином, на другому рівні ієрархії середнє абсолютне відхилення показника хмарності ресурсу Gismeteo складає 0.258.

3.6 Розрахунки

Після того, як була побудована ієрархія, виконана оцінка важливості критеріїв та розраховані середні абсолютні відхилення, можна розрахувати власний вектор матриці оцінки критеріїв.

3.6.1 Розрахунок власного вектору

Після того, як була побудована ієрархія, виконана оцінка важливості критеріїв та розраховані середні абсолютні відхилення, можна розрахувати власний вектор матриці оцінки критеріїв. Для цього можна використати ітеративний алгоритм. В загальному випадку, ітеративний алгоритм виглядає наступним чином:

while (i) do

$x = m * eigenvector$

$a = sum(x)$

$eigenvector = x / a$

де i – максимальна кількість ітерацій, необхідна для того, щоб a почало сходитися до якогось числа, m – матриця оцінок критеріїв, $eigenvector$ – вектор розмірності N , заповнений значеннями $1 / N$, а N – кількість критеріїв.

Для отриманої ієрархії, дані для роботи алгоритму будуть виглядати наступним чином:

$N = 7$

$m = (1, 0.25, 2, 0.25, 8, 0.2, 9,$

$4, 1, 0.33, 0.2, 8, 0.2, 9,$

0.5, 3, 1, 0.167, 8, 0.2, 9,
 4, 5, 6, 1, 8, 0.167, 9,
 0.125, 0.125, 0.125, 0.125, 1, 7, 9,
 0.11, 0.11, 0.11, 0.11, 0.11, 0.11, 1)

eigenvector = (0.143, 0.143, 0.143, 0.143, 0.143, 0.143, 0.143)

Після першого запуску алгоритму отримуємо наступний власний вектор:

eigenvector = (0.088, 0.101, 0.096, 0.182, 0.068, 0.148, 0.004)

При цьому a приймає наступні значення:

$a = 2.960$

$a = 6.210$

$a = 9.356$

$a = 14.098$

$a = 16.601$

$a = 16.838$

$a = 16.838$

Нескладно помітити, що останні два значення a співпадають, що означає, що a зійшлося до значення 16.838. Таким чином, у даному випадку однієї ітерації достатньо для того, щоб отримати власний вектор матриці оцінки критеріїв.

Згідно отриманого вектору, найважливішим критерієм є вітер, а найменш важливим – температура, що відчувається.

3.6.2 Лінійна згортка

Для упорядкування альтернатив, доцільно провести лінійну згортку критеріїв. У загальному випадку, лінійна згортка описується наступною формулою:

$$s_i = x_{i_1} w_1 + x_{i_2} w_2 + \dots + x_{i_m} w_m \quad (3.1)$$

де s – отриманий результат для кожної альтернативи, w – вектор вагових коефіцієнтів, i – кількість альтернатив, а j – кількість критеріїв.

Оскільки ресурс Gismeteo не надає дані про імовірність опадів, то показники згортки для цього параметру будуть дорівнювати 0.0

Проміжкові результати згортки за тиждень:

1.142 0.008 373.914 0.897 0.0 0.004 0.571
 0.799 0.007 374.165 0.931 0.0 0.004 0.228
 1.256 0.004 373.663 0.838 0.0 0.0 0.685
 1.713 0.005 373.789 0.848 0.0 0.0 0.685
 1.370 0.003 372.533 0.928 0.0 0.004 0.799
 1.256 0.007 372.408 0.948 0.0 0.004 1.027
 1.941 0.005 371.278 0.906 0.0 0.004 1.713
 1.941 0.0 371.152 0.904 0.0 0.006 1.484
 1.027 0.0 372.282 0.936 0.0 0.002 0.685
 1.370 0.005 372.282 0.945 0.0 0.004 0.799
 1.256 0.004 374.417 0.814 0.0 0.0 1.142
 1.370 0.004 374.668 0.816 0.0 0.0 0.456
 1.941 0.002 376.677 0.887 0.0 0.002 1.484

Остаточний результат згортки для сайтів:

Таблиця 3.15 – Результати лінійної згортки

Ресурс	Результат згортки
Gismeteo	6.073
Meta	6.184
Meteo	6.044
Rp5	6.0679
Sinoptik	6.0823

4 ПОРІВНЯННЯ ПОБУДОВАНОЇ СИСТЕМИ ІЗ СИСТЕМОЮ EXPERT CHOICE

Однією з популярних систем для прийняття рішень є система Expert Choice. Expert Choice - це система, що реалізує метод аналізу ієрархій для розв'язання задач багатокритеріального вибору. В рамках магістерської дисертації доцільно порівняти отриману систему з існуючим рішенням.

4.1 Внесення даних

У системі Expert Choice не передбачено завантаження вибірки даних для подальшого використання, що робить її непридатною для аналізу даних, їх дослідження та оцінки на адекватність. Це значно звужує спектр застосування системи, оскільки достовірність даних відіграє важливу роль у багатьох випадках та може впливати на вибір найкращого результату. До того ж, достовірність даних може бути важливим аспектом для експертів, адже зазвичай чим більше достовірність інформації, тим вище рівень довіри, а отже і експертна оцінка. У цьому контексті розроблена у магістерській роботі система є більш універсальною, адже вона дозволяє більш точно оцінити альтернативи.

4.2 Побудова ієрархії та оцінка критеріїв

У системі Expert Choice є можливість вводу багаторівневих ієрархій, що дозволяє будувати складні моделі. Однак, у той же час оцінка альтернатив відбувається для найменших рівнів ієрархії, що не є доцільним при вирішенні задачі, що наведена у магістерській дисертації, адже на першому рівні побудованої ієрархії знаходяться 28 критеріїв та оцінка альтернатив за кожним з них є важкою процедурою через той факт, що для оцінки у програмі Expert Choice буде використовуватися експертна оцінка, що змушує ОПР формувати 140 оцінок.

Побудована у магістерській роботі система навпаки має під собою ідею мінімізації впливу людського фактору на вибір альтернативи. Так, перший

рівень ієрархії не потребує експертної оцінки, але необхідний для того, щоб оцінити достовірність даних. На першому рівні ієрархії дані, що були занесені у систему, порівнюються з достовірним джерелом, що дозволяє сформувати середні абсолютні відхилення, які потім матимуть вплив на остаточний результат розв'язку задачі багатокритеріального вибору.

Для подальшого порівняння двох систем, необхідно сформувати перший рівень ієрархії. Занесення критеріїв до системи наведено на рисунках 4.1 та 4.2.

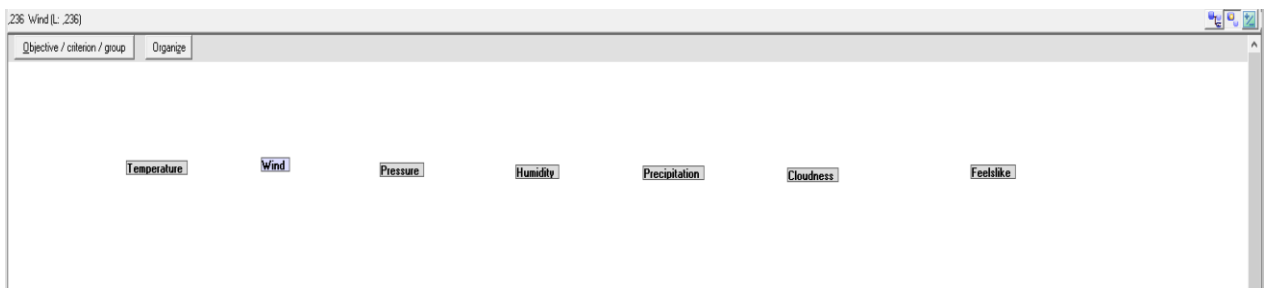


Рисунок 4.1 – Критерії оцінки альтернатив у програмі Expert Choice

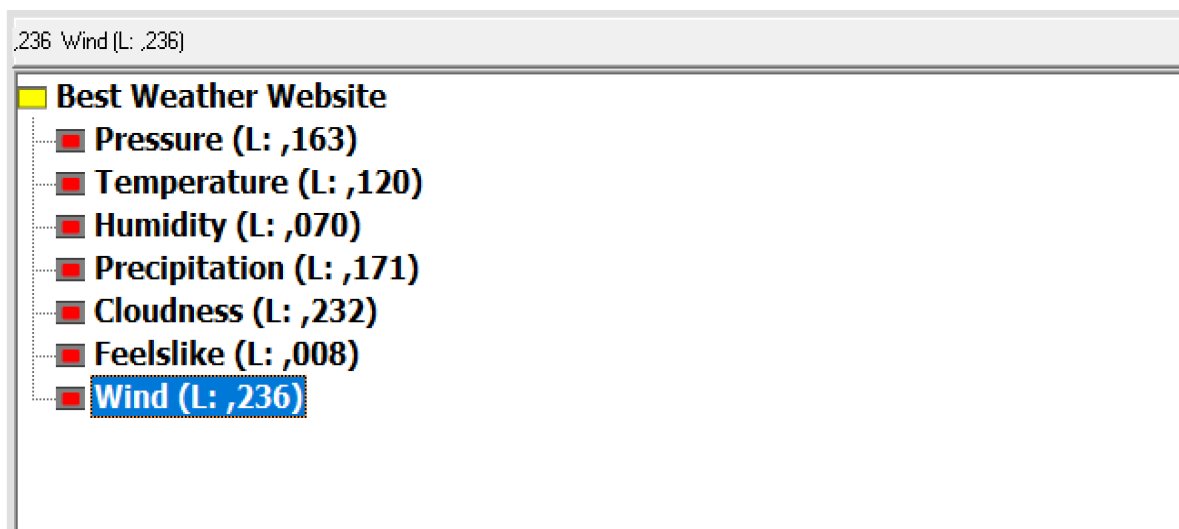


Рисунок 4.2 – Структура моделі у програмі Expert Choice

Через той факт, що у програмі Expert Choice порівняння альтернатив виконується на найменшому рівні, ієрархія у цій системі буде однозв'язною. Графічний вигляд отриманих ієрархій в обох системах наведено на рисунку 4.3 та 4.4.

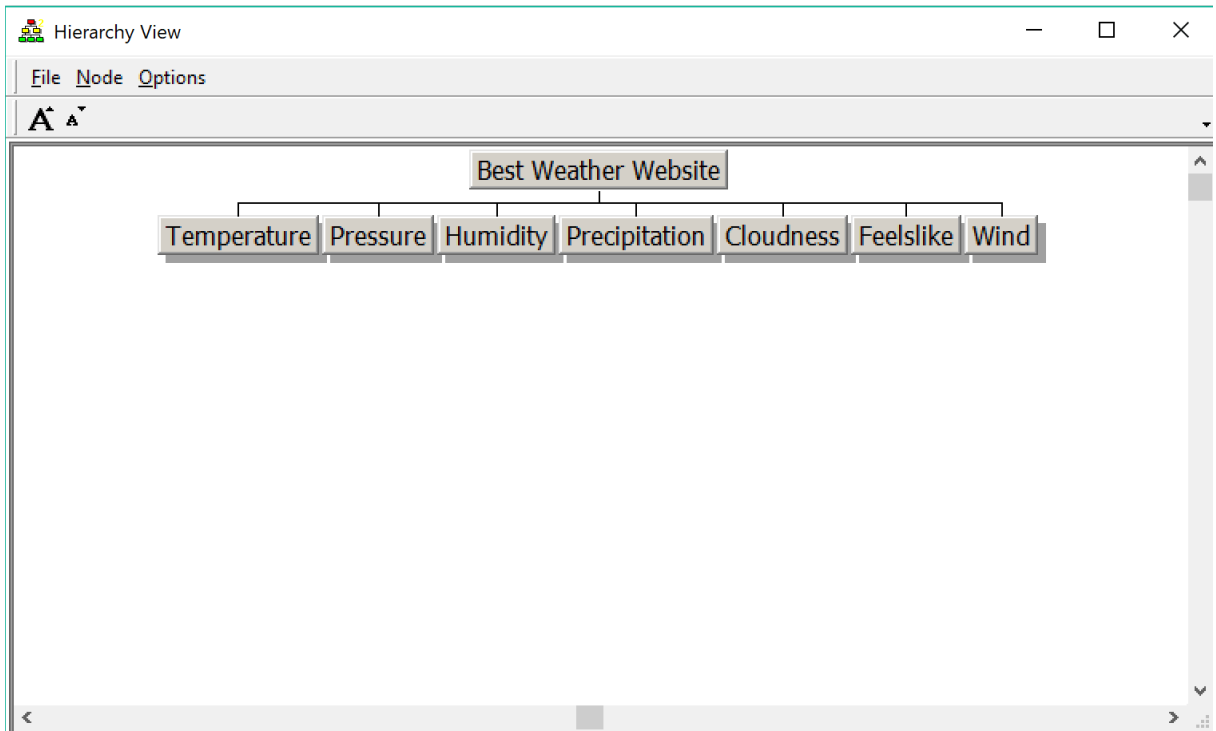


Рисунок 4.3 – Вигляд ієрархії у програмі Expert Choice

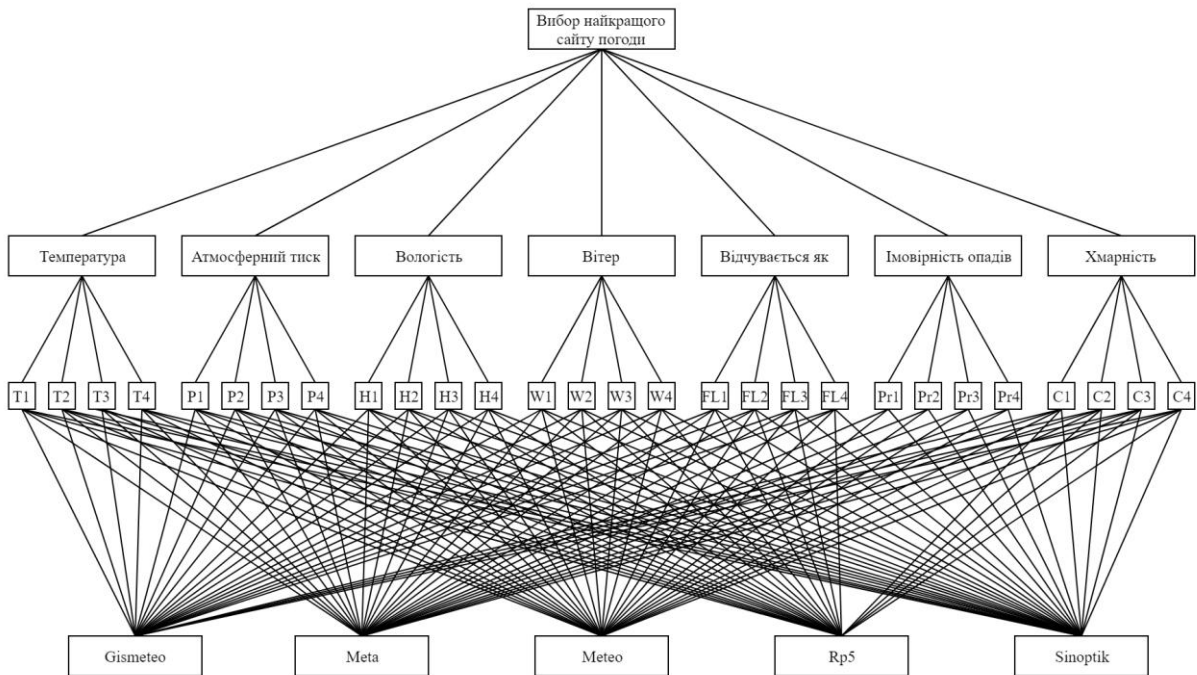


Рисунок 4.4 – Вигляд ієрархії у розробленій програмі

Надалі, необхідно оцінити критерії відносно один одного. У програмі Expert Choice є декілька графічних інтерфейсів для вводу даних у цифровому вигляді або ж для оцінки за допомогою повзунків. Однак, візуально дані

записуються в однаковому форматі, незважаючи на те, який з критеріїв кращий, що не є візуально очевидним навіть не зважаючи на те, що відображаються оцінки різним кольором. У отриманій системі використовуються дроби, що є зручнішим. Таблиці оцінки наведено на рисунка 4.4 та 4.5.

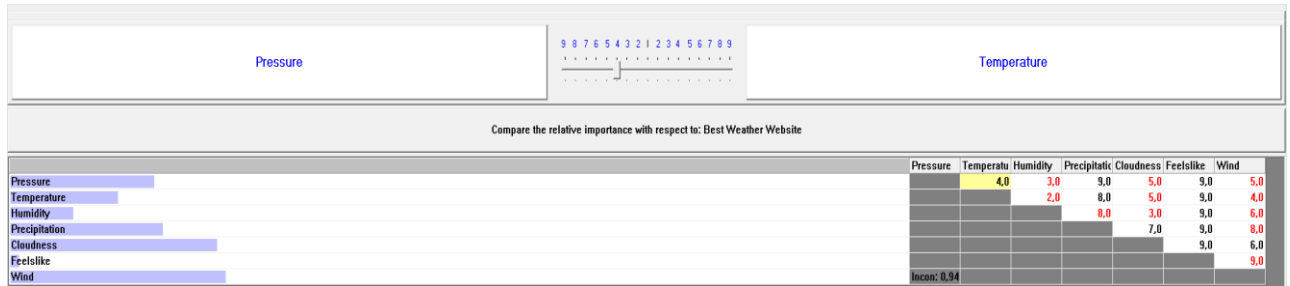


Рисунок 4.4 – Вигляд таблиці оцінки критеріїв у програмі Expert Choice

	Temperature	Pressure	Humidity	Wind	Precipitation	Cloudness	Feels Like
Temperature	1	1/4	2	1/4	1/6	1/5	7
Pressure	4	1	1/3	1/5	1/6	1/5	1/2
Humidity	1/2	3	1	1/6	1/6	1/3	1/6
Wind	4	5	6	1	1/4	1/6	7
Precipitation	6	6	6	4	1	7	1/5
Cloudness	5	5	3	6	1/7	1	1/4
Feels Like	1/7	2	6	1/7	5	4	1

Рисунок 4.5 – Вигляд таблиці оцінки критеріїв у розробленій програмі

Надалі, виконується розрахунок власного вектору матриці оцінки критеріїв. Програма Expert Choice одразу наводить візуалізацію отриманих результатів, що зручно. Результати обчислення пріоритетів наведено на рисунках 4.5 та 4.6 .



Рисунок 4.5 – Пріоритети критеріїв у програмі Expert Choice

```

Temperature: 0.08756758384145678
Pressure: 0.1008353588358979
Humidity: 0.09594815020363624
Wind: 0.18224805852931447
Precipitation: 0.06835041758077615
Cloudness: 0.14824175268083686
Feels like: 0.0038164521837857116

```

Рисунок 4.6 – Пріоритети критеріїв у розроблений програмі

Обидві системи показали, що найбільш пріоритетним критерієм є вітер, а найменш пріоритетним – температура, що відчувається.

4.3 Оцінка альтернатив

На етапі оцінки альтернатив методи, якими користуються системи значно розрізняються. У програмі Expert Choice оцінку альтернатив проводить ОПР, що означає, що на вибір найкращої альтернативи матиме вплив людський фактор. Оскільки багатьом людям зазвичай притаманно обирати якийсь один ресурс для того, щоб дивитися погоду, експертні оцінки матимуть під собою не оцінку достовірності даних, а оцінку безпосередньо сайту, що з точки зору користувача може бути важливим, але з точки зору обраної задачі не є критерієм оцінки.



Рисунок 4.7 – Оцінка альтернатив за тиском

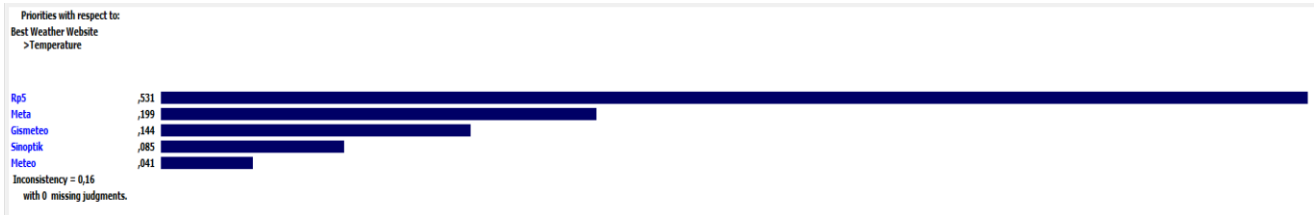


Рисунок 4.8 – Оцінка альтернатив за температурою

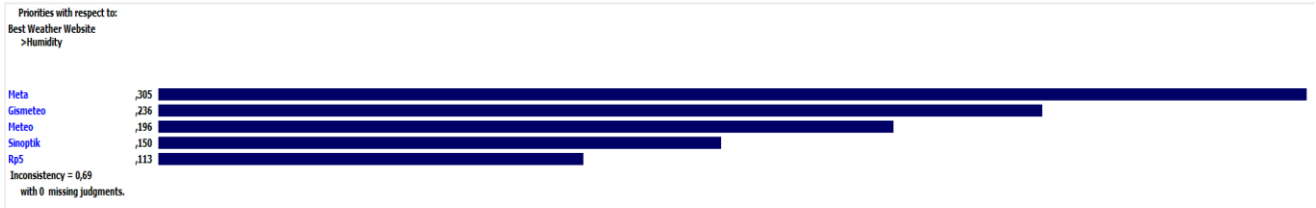


Рисунок 4.9 – Оцінка альтернатив за вологістю

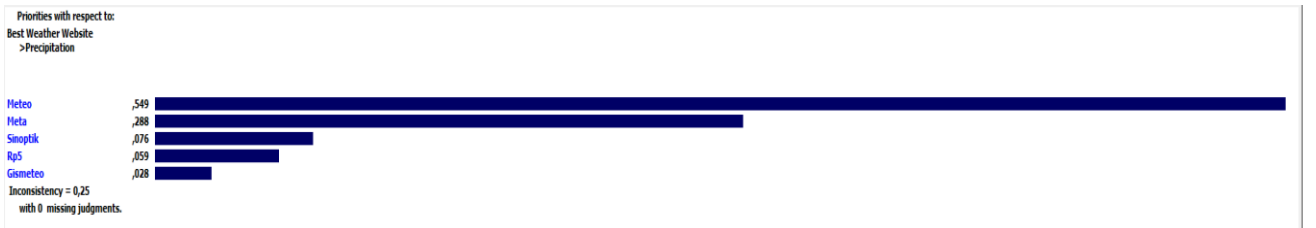


Рисунок 4.10 – Оцінка альтернатив за імовірністю опадів

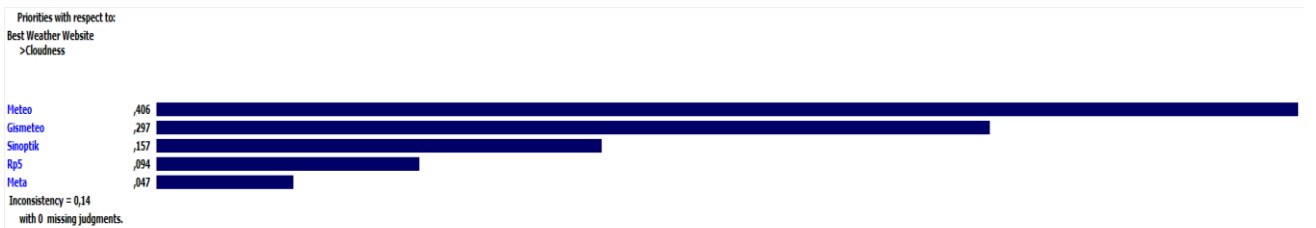


Рисунок 4.11 – Оцінка альтернатив за хмарністю



Рисунок 4.12 – Оцінка альтернатив за температурою, що відчувається



Рисунок 4.13 – Оцінка альтернатив за температурою, що відчувається

Таким чином, експертом було виставлено 35 оцінок щодо того, як кожна з альтернатив надає показники кожного з критеріїв відносна одна одної.

У побудованій системі навпаки, експертна оцінка для оцінки альтернатив не потрібна, адже її роль виконує розрахування середніх абсолютних відхилень. Це дозволяє абстрагуватися від альтернатив як веб-ресурсів та робити висновки виключно на даних, які вони надають. Результати розрахунків середніх абсолютних відхилень наведено у таблиці 4.1.

Таблиця 4.1 – Результати розрахунку середніх абсолютних відхилень

Альтернатива	Температура	Хмарність	Атмосферний тиск	Вологість	Вітер
Gismeteo	0.214	0.025	33.785	0.003	0.753
Meta	0.123	0.012	23.857	0.014	0.635
Meteo	0.356	0.022	40.571	0.022	0.537
Rp5	0.857	0.013	26.285	0.087	0.522
Sinoptik	0.785	0.042	33.214	0.009	0.321

4.4 Вибір найкращої альтернативи

Після проведення усіх необхідних розрахунків, можна робити висновок, яка з альтернатив є найбільш оптимальним рішенням. Програма Expert Choice вибирає найкращу альтернативу з точки зору найважливішого критерія. У випадку побудованої моделі таким критерієм є вітер. На рисунку 4.14 наведено зведений результат роботи програми.

Якщо взяти до уваги розгорнутий результат (рисунок 4.15), то можна побачити, як виглядає розподіл результатів для кожної з альтернатив відносно окремих критеріїв. Це надає змогу оцінити, за умови пріоритетності якого з критеріїв найбільш оптимальна альтернатива зміниться.

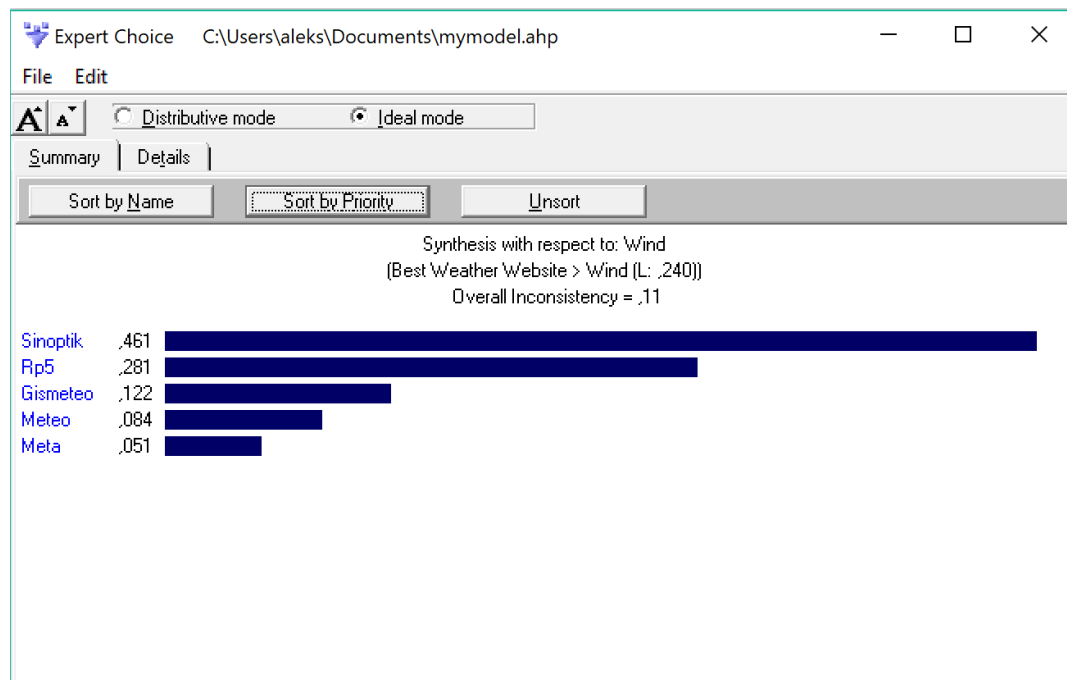


Рисунок 4.14 – Зведений результат роботи програми Expert Choice

Візуально, графік чутливості альтернатив до критеріїв наведено на рисунку 4.16. Так, якщо найбільш пріоритетним критерієм стане температура, то оптимальною альтернативою стане Rp5, якщо буде обрано вологість – Meta, у випадку пріоритетності хмарності – Meteo.

Однак з точки зору достовірності даних цей результат не надає відповіді про те, наскільки альтернатива відповідає очікуванням користувачів сайтів погоди. До того ж, різні комбінації ОПР призведуть до різних результатів вибору найбільш оптимальної альтернативи, оскільки у ОПР буду уподобання щодо якогось з ресурсів і це призведе до того, що результат буде кращим з точки зору окремої групи експертів, але не за точки зору достовірності даних.

Level 1	Alt	Prp
Percent Pressure (L: .188)		18.8
	Gismeteo	.058
	Meta	.013
	Meteo	.011
	Rp5	.063
	Sinoptik	.023
Pressure (L: .188)		
Percent Temperature (L: .143)		14.3
	Gismeteo	.021
	Meta	.028
	Meteo	.006
	Rp5	.076
	Sinoptik	.012
Temperature (L: .143)		
Percent Humidity (L: .109)		10.8
	Gismeteo	.026
	Meta	.033
	Meteo	.021
	Rp5	.012
	Sinoptik	.016
Humidity (L: .109)		
Percent Precipitation (L: .115)		11.5
	Gismeteo	.003
	Meta	.033
	Meteo	.063
	Rp5	.007
	Sinoptik	.009
Precipitation (L: .115)		
Percent Cloudness (L: .196)		19.6
	Gismeteo	.058
	Meta	.009
	Meteo	.080
	Rp5	.019
	Sinoptik	.031
Cloudness (L: .196)		
Percent Feelslike (L: .008)		0.9
	Gismeteo	.002
	Meta	.001
	Meteo	.001
	Rp5	.004
	Sinoptik	.001
Feelslike (L: .008)		
Percent Wind (L: .240)		23.9
	Gismeteo	.023
	Meta	.012
	Meteo	.020
	Rp5	.067
	Sinoptik	.111
Wind (L: .240)		

Рисунок 4.15 – Розгорнутий результат роботи програми Expert Choice

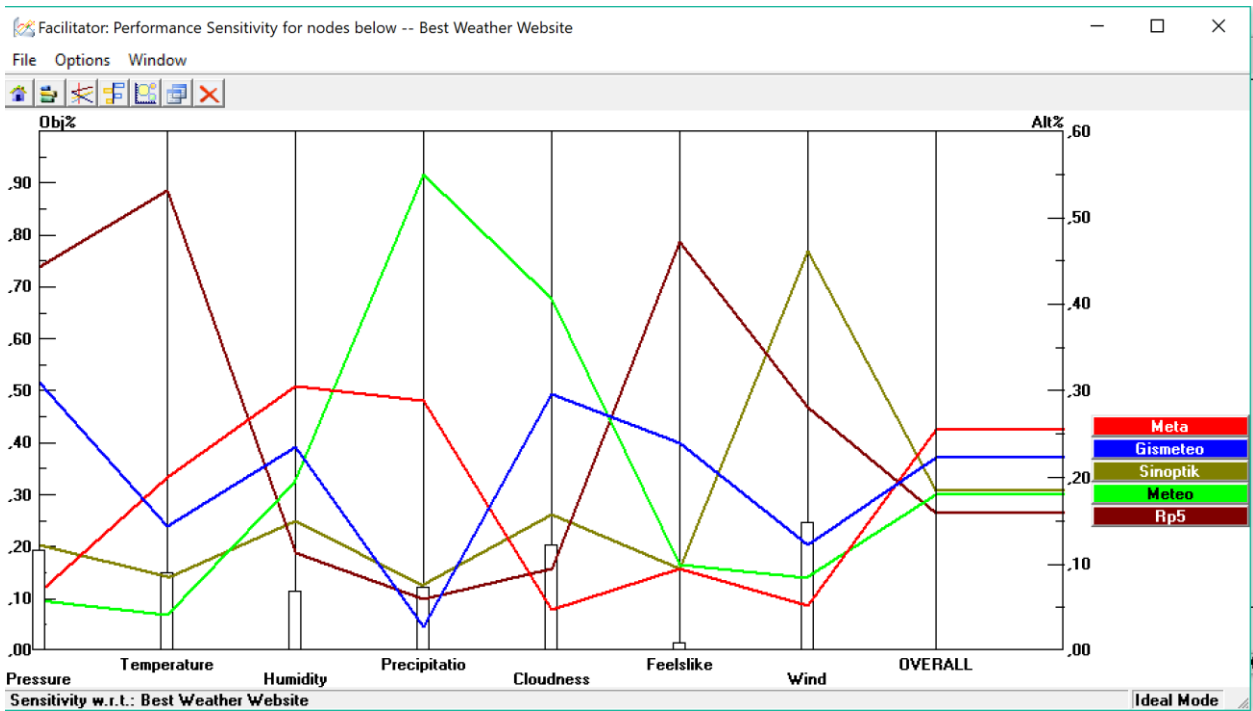


Рисунок 4.16 Графік чутливості альтернатив до критеріїв у програмі Expert Choice

У системі, що була розроблена у магістерській роботі, висновки що до найкращої з альтернатив формуються не на основі експертних оцінок, а на

основі того, наскільки достовірні дані надав той чи інший ресурс по кожному з пріоритетних критеріїв.

Так, згідно власного вектору матриці оцінки критеріїв, найважливішими виявилися показники вітру, хмарність та тиск. За результатами показників абсолютного відхилення показників вітру найкращою альтернативою виявився сайт Sinoptik, за показниками хмарності – Meta, за показниками атмосферного тиску – Meta. Це означає, у двох з трьох пріоритетних критеріїв ресурс Meta надає найкращі показники. Таким чином, Meta є найоптимальнішою з альтернатив з точки зору пріоритетності критеріїв та середніх абсолютних відхилень.

Варто зазначити, що у системі, що була розроблена у магістерській роботі також є можливість врахування чутливості альтернатив до критеріїв, але це не означає, що найкращою автоматично стане альтернатива, показники якою будуть найкращі для найважливішого критерію. Оскільки система дозволяє врахувати декілька пріоритетних критеріїв, це робить її більш універсальною та дозволяє отримувати більш точні результати.

ВИСНОВКИ

В ході дослідження було проаналізовано найпопулярніші методи прийняття рішень, які використовуються у сучасному світі для побудови СПР. Було описано метрики, за якими можливо оцінити універсальність методів та можливості їх застосування.

Були описані недоліки кожного з існуючих методів, на основі яких було обрано МАІ як найбільш наближений до вимог універсального методу прийняття рішень та було описано можливості його покращення з метою отримання більш точних результатів при виборі найкращої альтернативи.

Попередня обробка та аналіз даних, який є вдосконаленням МАІ, дозволяє вирішити проблеми достовірності даних адже під час аналізу обчислюються абсолютні відхилення кожного за показників погодим, що дозволяє зробити висновки про достовірність кожної з альтернатив.

Запропонована система покращує показники роботи МАІ, оскільки:

- дозволяє перевіряти дані на достовірність;
- дозволяє абстрагуватись від ОПР на етапі оцінки альтернатив;
- надає можливість враховувати більше одного пріоритетного критерію;
- дозволяє більш гнучко обирати шкали оцінки критеріїв;

Як видно з результатів, побудована система надає більш точні результати та надає можливість враховувати більше факторів, що впливають на оцінку результатів.

Система здатна суттєво підвищити рівень ефективності МАІ згідно отриманих даних, проте не дозволяє однозначно прогнозувати результати у майбутньому, оскільки на обраний часовий ряд діють багато зовнішніх факторів, які важко та іноді неможливо врахувати

ПЕРЕЛІК ПОСИЛАНЬ

1. Кини Р. Принятие решений при многих критериях: предпочтения и замещения / Р. Кини, Х. Райф. – М.: Радио и связь, 1981. – 560 с
2. Гавриленко О. В., Лекції з теорії прийняття рішень [Електронний ресурс] Режим доступу: <http://evgavrilenko.ucoz.ru/index/lekciji/0-134>
3. Ногин В. Д. Принятие решений в многокритериальной среде: количественный подход / Ногин В. Д. – М.: ФИЗМАТЛИТ, 2002. – 144 с
4. Лотов А. В. Многокритериальные задачи принятия решений / А. В.Лотов, И. И. Пospelова . – М.: МАКС Пресс, 2008. – 197 с
5. Ногин В. Д. Использование количественной информации об относительной важности критериев в принятии решений / В. Д.Ногин - СпБ.: НаучнойТехнсические ведомости Спбгту. 2000. – № 2. – С. 89 – 93.
6. Чернов В.Г. Основы теории нечетких множеств : учеб. пособие / В.Г. Чернову - Владим. гос. ун-т.- Владимир : Изд-во Владим. гос. ун-та, 2010. – 96 с.
7. Волошин О.В. Моделі та методи прийняття рішень / О. Ф. Волошин, С. О. Мащенко. – 2-ге вид., перероб. та допов. – К.: Видавничо-поліграфічний центр "Київський університет", 2010. – 336 с.
8. Саати Т. Аналитическое планирование / Т. Саати, К. Кернс. – М.: Радио и связь, 1991. – 320 с.
9. Саати Т. Принятие решений. Метод анализа иерархий / Т. Саати. – М.: Радио и связь, 1993. – 210 с.
10. Рыжов А.П. Элементы теории нечетких множеств и ее приложений / Рыжов А.П. -М.: Диалог-МГУ, 1998. - 81 с.
11. Бриллинджер Ф. Анализ временных рядов / Ф. Бриллинджер. -Г.: Мир, 1978. – 635с.
12. Афанасьев В.Н., Юзбашев М.М. Коэффициент корреляции рангов как показатель устойчивости динамики / Вестник статистики. – 1983. – №11

13. Грешилов А. А. Математические методы построения прогнозов / А. А. Грешилов, В. А. Стакун. – М.: Радио и связь, 1997. – 112 с.
14. Казинец Л.С. Темпы роста и абсолютные приросты / Казинец Л.С. - М: Статистика, 1975
15. Кендэл М. Временные ряды: Перю с англ. -М: Финансы и статистика, 1981
16. Крастинь О.П. Изучение статистических зависимостей по многолетним данным, -М.: Финансы и статистика, 1981