

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

бакалавр

(назва освітнього ступеня)

на тему: Процес підвищення якості кліматичних даних
в процесі їх збору та реєстрації

Виконав(ла): студент(ка) 4 курсу, групи СН-41
спеціальності 122 Комп'ютерні науки

(шифр і назва спеціальності)

(підпис)

Крушельницький В.О.

(прізвище та ініціали)

Керівник

(підпис)

Приймак М.В.

(прізвище та ініціали)

Нормоконтроль

(підпис)

Шимчук Г.В.

(прізвище та ініціали)

Завідувач кафедри

(підпис)

Боднарчук І.О.

(прізвище та ініціали)

Рецензент

(підпис)

Оробчук О.Р.

(прізвище та ініціали)

Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

Факультет комп'ютерно-інформаційних систем і програмної інженерії
(повна назва факультету)

Кафедра комп'ютерних наук
(повна назва кафедри)

ЗАТВЕРДЖУЮ

Завідувач кафедри

Боднарчук І.О.

(підпис)

(прізвище та ініціали)

«25» січня 2021 р.

ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ

на здобуття освітнього ступеня бакалавр

(назва освітнього ступеня)

за спеціальністю 122 Комп'ютерні науки

(шифр і назва спеціальності)

студенту Крушельницький Вадим Олегович

(прізвище, ім'я, по батькові)

1. Тема роботи Процес підвищення якості кліматичних даних в процесі їх збору та реєстрації

Керівник роботи д.т.н., проф. Приймак М.В.

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом ректора від «16» березня 2022 року № 4/7-161

2. Термін подання студентом завершеної роботи 17 червня 2022 р.

3. Вихідні дані до роботи Літературні джерела з тематики роботи

4. Зміст роботи (перелік питань, які потрібно розробити)

ВСТУП 1. ЗБІР ДАНИХ ПРО ПОГОДУ 1.1 Метеорологічні мережі 1.2 Синоптичні від кліматичні дані 1.3 Історія станцій та їх метадані 1.4 Доступність та вартість РОЗДІЛ 2. ОЧИЩЕННЯ ПОГОДНИХ ДАНИХ 2.1 Просторова інтерполяція 2.2 Часова інтерполяція 2.3 Складності при очищенні даних про погоду 2.4 Вплив зміни/переміщення приладів 2.5 Аналіз стану метеорологічних станцій у Сполучених Штатах Америки 2.6 Аналіз стану метеорологічних станцій у інших країнах 2.7 Методи опрацювання розривів в серіях погодних даних 2.8 Методи визначення результатів та коригування даних про погоду 2.9 Перевірка методологій для визначення та кількісної оцінки результатів РОЗДІЛ 3. БЕЗПЕКА ЖИТТЄДІЯЛЬНОСТІ, ОСНОВИ ХОРОНИ ПРАЦІ ВИСНОВКИ ПЕРЕЛІК ПОСИЛАНЬ

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень, слайдів)

1. Тема дослідження. 2. Актуальність дослідження. 3. Задачі дослідження. 4. Збір метеорологічних даних. 5. Доступність даних про погоду та метаданих. 6. Причини помилок в метеорологічних даних. 7. Методи покращення якості даних. 8. Розриви у наборах даних. 9. Виявлення розривів. 10. Кумулятивна різниця температур між двома станціями. 11. Висновки

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Безпека життєдіяльності, основи охорони праці	Гурик О.Я., к.т.н., доц.		

7. Дата видачі завдання 25 січня 2022 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1.	Ознайомлення з завданням до кваліфікаційної роботи	25.01.22-27.01.22	<i>Виконано</i>
2.	Підбір джерел по темі роботи	28.01.22 – 01.04.22	<i>Виконано</i>
3.	Оформлення першого розділу	15.04.2022	<i>Виконано</i>
4.	Оформлення другого розділу	30.04.2022	<i>Виконано</i>
5.	Виконання завдання до підрозділу «Безпека життєдіяльності, основи охорони праці»	10.05.2022	<i>Виконано</i>
6.	Оформлення кваліфікаційної роботи	17.02.2022	<i>Виконано</i>
7.	Перевірка на плагіат	02.06.2022	<i>Виконано</i>
8.	Нормоконтроль	08.06.2022	<i>Виконано</i>
9.	Попередній захист кваліфікаційної роботи	09.06.2022	<i>Виконано</i>
10.	Захист кваліфікаційної роботи	21.06.2022	

Студент

(підпис)

Крушельницький В.О.

(прізвище та ініціали)

Керівник роботи

(підпис)

Приймак М.В.

(прізвище та ініціали)

АНОТАЦІЯ

Процес підвищення якості кліматичних даних в процесі їх збору та реєстрації // Кваліфікаційна робота освітнього рівня "Бакалавр" // Крушельницький Владислав Олегович // Тернопільський національний технічний університет імені Івана Пулюя, факультет комп'ютерно-інформаційних систем і програмної інженерії, кафедра комп'ютерних наук, група СН-41 // Тернопіль, 2022 // с. – 52, рис. – 5, табл. – 2, кресл. – 11, бібліогр. – 35.

Ключові слова: якість даних, дані про погоду, покращення якості даних, опрацювання даних.

Високоякісні погодні дані використовуються для управління ризиками, пов'язаними з бізнесом, залежним від погодних умов та для прийняття рішень на основі цих даних. Крім того, дані про погоду реальному часі надають учасникам ринку дискретні значення, щоб планувати щоденні угоди та дії. Наприклад, це може бути сільське господарство, транспорт, управління у сфері охорони здоров'я та інші предметні області.

Ринок погоди вимагає даних, які відповідають ключовим критеріям для використання їх при оцінці. В загальному йдеться про необхідність оперувати якісними даними. У цій роботі визначаються та обговорюються ці критерії, основні проблеми з даними погоди та засоби їх вирішення.

ABSTRACT

The process of the quality improvement the of climate data in the process of their collection and registration // Qualification work of the educational level "Bachelor" // Vladyslav Krushelnytskyi // Ternopil Ivan Puluj National Technical University, Faculty of Computer Information Systems and Software Engineering, Department of Computer Science, Group CH-41 // Ternopil, 2022 // p. – 52, fig. – 5, references – 35, posters – 11.

Key words: data quality, weather data, data quality improvement, data processing.

Weather data of high quality is used to manage the risks associated with weather-dependent businesses and for decisions making based on these data. In addition, real-time weather data gives market participants discrete values to plan daily transactions and actions. For example, this could be agriculture, transport, health care management and other subject areas.

The weather market requires data that meet key criteria for use them in evaluation process. In general, it is about the need to operate with quality data. This paper recognizes these criteria and discusses them. The core complications with data about weather and the means to solve them are discussed too.

ЗМІСТ

ВСТУП.....	7
РОЗДІЛ 1. ЗБІР ДАНИХ ПРО ПОГОДУ	9
1.1 Метеорологічні мережі.....	10
1.1.1 Сполучені Штати Америки	11
1.1.2 Європа.....	12
1.1.3 Японія та Австралія.....	12
1.1.4 Спеціальні мережі	13
1.2 Синоптичні та кліматичні дані	14
1.3 Історія станцій та їх метадані	16
1.3.1 Метадані станцій США.....	17
1.3.2 Міжнародні метадані	18
1.3.3 Доступність та вартість.....	18
РОЗДІЛ 2. ОЧИЩЕННЯ ПОГОДНИХ ДАНИХ.....	21
2.1 Просторова інтерполяція	22
2.2 Часова інтерполяція.....	24
2.3 Складності при очищенні даних про погоду	24
2.4 Вплив зміни/переміщення приладів	25
2.5 Аналіз стану метеорологічних станцій у Сполучених Штатах Америки	27
2.6 Аналіз стану метеорологічних станцій у інших країнах	29
2.7 Методи опрацювання розривів в серіях погодних даних.....	30
2.7.1 Визначення потенційних дат розривів	31
2.7.2 Виявлення розривів	32
2.8 Методи визначення результатів та коригування даних про погоду.....	34
2.8.1 Суб'єктивні методи.....	34
2.8.2 Об'єктивні методи	36
2.8.3 Кількісна оцінка розривів.....	39
2.9 Перевірка методологій для визначення та кількісної оцінки результатів	39

РОЗДІЛ 3. БЕЗПЕКА ЖИТТЄДІЯЛЬНОСТІ, ОСНОВИ ХОРОНИ ПРАЦІ ...	41
3.1 Аналіз небезпеки і шкідливості при розробці програмного забезпечення.....	41
3.2 Інформаційно-психологічні небезпеки.....	43
ВИСНОВКИ.....	48
ПЕРЕЛІК ПОСИЛАНЬ	50

ВСТУП

Основним джерелом даних про погоду в кожній країні є національна метеорологічна служба (НМС). Кожна НМС працює незалежно, встановлюючи свої стандарти для збору, архівування та розповсюдження даних про погоду.

Дані про погоду зазвичай класифікуються як синоптичні або кліматичні дані. Синоптичні дані – це дані в реальному часі, надані для використання в моделюванні безпеки авіації та прогнозуванні. Кліматичні дані – це офіційний запис даних, який зазвичай надається після певного контролю якості даних. У багатьох країнах також існують спеціальні мережі, які в деяких випадках можуть використовуватися для надання додаткових кліматичних даних. Ці теми та пов'язані з ними питання будуть додатково обговорені нижче.

Дані про погоду, опубліковані НМС, часто повинні бути «очищені», тобто має бути виконана заміна відсутніх та помилкових даних, перш ніж ринок погоди зможе їх використовувати. Типи помилок, яких слід очікувати, та методології їх усунення обговорюються у відповідному розділі роботи.

Після очищення даних про погоду можуть все ще існувати проблеми, які потрібно виправити, перш ніж дані можна буде використовувати для моделювання погодних ризиків. Ці проблеми відомі як «неоднорідності» або «розриви», викликані переміщенням станції, зміною приладів або змінами навколишнього середовища поблизу станції. Опрацювання неоднорідностей у часових рядах погодних даних зазвичай називають «гомогенізацією» або «коригуванням однорідності». У цьому розділі наведено ретельне обговорення покращення даних, яке є коригуванням даних у результаті дискретних змін часового ряду даних про погоду, викликаних змінами на метеорологічній станції. Недискретні неоднорідності, такі як ефекти урбанізації, розглядаються як тенденції.

Існує багато застосувань для даних про погоду в моделюванні погодних ризиків, включаючи потенційне використання даних дистанційного зондування

для населення. Висновки щодо використання погодних даних в управлінні ризиками представлені також в цій роботі.

РОЗДІЛ 1. ЗБІР ДАНИХ ПРО ПОГОДУ

В ідеалі на ринку потрібні своєчасні та точні дані про погоду. Для досягнення цього необхідно постійно записувати дані зі станцій, які належним чином ідентифіковані, укомплектовані навченим персоналом або автоматизовані з регулярним обслуговуванням, у належному робочому стані та захищених від втручання. Станції також не повинні піддаватися переміщенню.

На жаль, основним призначенням НМС у більшості, якщо не в усіх країнах є «захист життя та власності». Основна увага цих відділень зосереджена на точній діагностиці, відстеженні та прогнозуванні суворих погодних явищ. Збір та архівування даних про погоду є важливими, оскільки це дає економічну вигоду, але місцеві/національні економічні потреби не так залежать від високої якості даних, як ринок погодних ризиків. Таким чином, більшість державних метеорологічних служб не акцентують увагу на питаннях якості даних.

Всесвітня метеорологічна організація (ВМО) є міжнародною асоціацією членів НМС, але не має повноважень щодо виконання [1]. Хоча різні організації мають стандартизовані продукти, очевидно, що кожна організація стоїть окремо у своїй практиці у встановленні своїх стандартів для збору, архівування та розповсюдження даних. Наприклад, хоча ВМО встановила стандартний графік подання даних у певний час кожного дня, вона не має можливості отримати дані, які повідомляються в цей час. Щоденні мінімальні та максимальні температури повідомляються з півночі до півночі того самого календарного дня в США та Німеччині. Однак інші європейські країни повідомляють ці значення по-різному (наприклад, метеорологічне управління Великобританії (УКМО) повідомляє про добову мінімальну температуру з 09:00 за всесвітнім координованим часом (UTC) попереднього дня до 09:00 UTC поточного дня, а максимальна температура поточного дня з 09.00 UTC цього дня до 09.00 UTC наступного дня) [2]. Ці окремі стандарти та методи звітності створюють труднощі в роботі з даними про погоду в глобальному

масштабі та вимагають процедур контролю якості, що залежать від країни. Знання цих питань стандартизації та інших питань, включаючи підтримку історичної інформації станцій (метаданих), обслуговування мереж спостережень, доступність даних, права на використання даних є надзвичайно важливими для ринку погоди; без розуміння того, як збираються та обробляються дані про погоду, а також без знання історії станції, дуже важко оцінити надійність цих даних і використовувати їх. Оскільки індустрія погодних деривативів продовжує розширюватися в усьому світі, ці проблеми вирішує приватна індустрія погоди. З НМС укладаються угоди щодо перерозподілу вихідних даних та доступу до інформації про місцеві методи збору даних про погоду та історію станцій.

1.1 Метеорологічні мережі

Дані про погоду записуються з кількох метеорологічних мереж. Основна мета метеорологічних мереж – надавати синоптичні та кліматичні дані. Як правило, синоптичні дані – це дані, отримані одночасно на великій площі атмосфери. Дані SYNOP (дані, що знаходяться на станціях, які звітують відповідно до стандартів синоптичної звітності ВМО) піддаються дуже незначному контролю якості та розповсюджуються між НМС для створення прогнозів.

Кліматичні дані, які іноді отримують із синоптичних даних, є даними, які зазвичай використовуються для кліматологічних досліджень. Ці дані підлягають процедурам контролю якості та найчастіше формують офіційний запис історичної погоди та клімату. Набори кліматичних і синоптичних погодних даних можуть бути однаковими або не однаковими. Для збору обох наборів даних можна використовувати ті самі станції. Однак більше перевірок контролю якості буде здійснюватися з кліматичними даними. В інших випадках для запису синоптичних і кліматичних даних використовуються фізично окремі прилади спостереження.

Існують різні конвенції для ідентифікації станцій, створені різними організаціями. ВМО створила п'ятизначну систему нумерації для ідентифікації станцій для міжнародного звітування синоптичних даних з цих станцій. Існує також умова найменування для погодинних звітів METAR (акронім приблизно перекладається з французької як Aviation Routine Weather Report), які складаються з чотирьох символів. У США, WBAN (Weather Bureau Army Navy) п'ятизначна конвенція часто використовується при написанні контрактів для ідентифікації станції. Метеорологічні мережі окремих країн визначаються за допомогою умов, вибраних окремою країною. Відсутність всеосяжного міжнародного стандарту ідентифікації іноді викликає плутанину на ринку.

1.1.1 Сполучені Штати Америки

Офіційним джерелом погодних даних у Сполучених Штатах є Національний центр кліматичних даних (NCDC) [3], агентство Національного управління океанічних і атмосферних досліджень (NOAA [4]). США не роблять суттєвої різниці між синоптичними та кліматичними мережами. Класифікація мереж США базується на системі рівнів із станціями найвищої якості, які називаються станціями «першого порядку» (основними станціями).

Відмінності в класі станцій спостережень полягають у вимірюваних змінних, частоті спостережень, які повідомляються, загалом тривалості історії даних та типі операторів, які обслуговують ці станції. Ринок погоди найчастіше вибирає станції першого порядку для управління погодними ризиками США. Ці станції працюють 24 години на добу та обслуговуються Національною службою погоди (NWS) [5] підготовленим і сертифікованим персоналом. Ці станції першого порядку відповідають за спостереження за офіційними кліматичними даними, а також за надання синоптичних даних для цілей прогнозування. Спостереження, зроблені на цих станціях, охоплюють велику кількість змінних, включаючи температуру, опади, приземний тиск, вологість, швидкість і напрям вітру, хмарний покрив, глибину снігу, видимість і сонячну

радіацію. Ці спостереження проводяться та розповсюджуються часто, щогодини або в певний час доби. Ці спостереження також очищаються (National Weather Service – NWS) і повідомляються щомісяця у звітах місцевих кліматологічних даних (Local Climatological Data – LCD).

Підготовлений персонал, який контролюється персоналом NWS, також підтримує станції «другого порядку». Ці станції реєструють температуру, опади та багато, але не обов'язково всі, інші змінні, виміряні на станціях першого порядку. Ці станції надають неофіційні кліматичні дані, а також деякі синоптичні дані. Однак для цих станцій не надаються офіційні щомісячні LCD-звіти.

1.1.2 Європа

Так само в Європі існує кілька метеорологічних мереж. Велика Британія також класифікувала свої станції за системою рівнів на основі якості спостережень за погодою. Систему Великобританії можна порівняти з класифікацією систем США першого та другого порядку, але офіційно вона не класифікується таким чином. Багато з цих станцій надають як кліматичні, так і синоптичні дані. Станції, які повідомляють кліматичні дані, розташовані в комбінації комерційних аеропортів, баз Королівських ВПС та обсерваторій.

У Німеччині існує кілька мереж, кожна з яких визначається типом спостережень за погодою (наприклад, кількість опадів, клімат, синоптичні тощо) [6]. Мережа опадів містить до 4000 станцій, тоді як кліматична мережа складається з приблизно 200 станцій.

1.1.3 Японія та Австралія

В Японії існують два типи мереж : AMEDAS (Автоматизована система збору метеорологічних даних) [7] і SYNOP. Хоча станцій AMEDAS більше, ніж станцій SYNOP (більше 1000 порівняно з менш ніж 150), ринок погоди віддає

перевагу використанню станцій SYNOP, оскільки більшість із цих станцій керовані (близько 120), а добові температури визначаються з безперервних показань. Додаткові дані доступні через провінційні офіси, але якість даних може бути під сумнівом. Звіти про дані іноді відсутні, а якість даних не контролюється суворо. Крім того, станції можуть не забезпечувати безперервну 24-годинну звітність.

В Австралії як кліматичні, так і синоптичні дані спостерігаються на багатьох із тих самих станцій у межах різних метеорологічних мереж [8]. Станції, якими користується метеорологічний ринок, здебільшого складаються зі станцій, розташованих в аеропортах і обслуговуються державними службовцями.

Мережа кліматичних довідок була створена в Австралії для надання даних для досліджень зміни клімату. Наразі ці станції мало корисні для метеорологічної галузі через їх віддаленість та обмежену історію. Крім того, кооперативні волонтери, які зазвичай проживають в сільськогосподарських районах країни, забезпечують велику додаткову мережу станцій моніторингу опадів, що складається з приблизно 7000 станцій. Ця опадова мережа може мати проблеми з якістю, оскільки не дуже добре контролюється.

1.1.4 Спеціальні мережі

Окрім кліматичних та синоптичних мереж, НМС опікуються низкою спеціалізованих мереж спостережень. Ці мережі служать для забезпечення більшого просторового охоплення та/або надання спеціалізованих спостережень. Ці мережі значно відрізняються за точністю, надійністю та доступністю. NWS США керує мережею спільних спостерігачів [9]. Ця мережа з приблизно 12000 активних станцій служить для забезпечення більшого просторового охоплення метеорологічних спостережень. Підмножина цих станцій з кількома станціями першого та другого порядку є частиною історичної кліматологічної мережі США (HCN), що складається з 1221 станції,

бази даних, складеної NCDC та використовується для аналізу клімату США [10]. Багато з цих станцій фіксують лише максимальні та мінімальні температури й опади, як правило, щодня. Дані для цих станцій слід використовувати з обережністю. Є кілька відомих проблем з кооперативними мережами. Спостерігачі – це волонтери, а не професіонали, обладнання базове, дані часто відсутні або помилкові, доступність даних у реальному часі погана, а час проведення спостережень може відрізнятись («час упередження спостереження»).

1.2 Синоптичні та кліматичні дані

Однією з проблем, з якою стикається ринок погоди, є визначення кліматичних та синоптичних даних. Щоб полегшити плутанину та встановити стандарт, Асоціація управління погодними ризиками (WRMA) [11] визначила синоптичні дані – як «дані, які збираються в режимі реального часу на різних станціях по всьому світу та надаються через Глобальну телекомунікаційну систему (ГТС) [12]. Дані, мінімальна та максимальна температура та опади – зазвичай надаються чотири рази на день о 00:00 за Грінвічем (GMT), 06:00 GMT, 12:00 GMT та 18:00 GMT. Зазвичай реєструється 12-годинна мінімальна і 12-годинна максимальна температура, але час, який представляють 12-годинні значення, залежить від місцевого часу в країні вимірювання». Кліматичні дані визначаються як «дані, якість яких контролюється відповідними НМС, де дані збираються. «Кліматичні дані» є «офіційними» даними станцій цієї країни».

Під час огляду цих двох типів даних WRMA рекомендувала, щоб кліматичні дані були найбільш підходящими даними для використання в промисловості похідних погодних інструментів.

Історичні синоптичні дані, як правило, мають коротший запис, ніж кліматичні дані; найдовша історія синоптичних даних сягає лише 1974 року. Синоптичні записи, як правило, менш точні, оскільки до 1982 року всі температури були округлені до цілих градусів Цельсія, а процедури контролю

якості, можливо, не були ретельними. Синоптичні дані також можуть містити великий відсоток відсутніх значень.

Кліматичні дані – це офіційні історичні дані для країни. Цей статус диктує набір даних із тривалим періодом запису (як правило, що поширюється принаймні до 1961 року), значним контролем якості, меншою кількістю відсутніх значень, більшою роздільною здатністю та чітким визначенням даних. Клімат певного дня, визначений відповідними NMS, зазвичай базується на календарному дні (тобто середній показник клімату на цю дату в минулому), тоді як синоптичні спостереження повідомляються відповідно до графіків WMO. Розклади спостережень WMO створюються для надання даних для запуску прогнозних моделей, тоді як кліматичні дані служать цілям надання кліматологічних даних для даного місяця.

Відмінності між поточними синоптичними даними та кліматичними даними найбільш очевидні в термінах доступності, точності та надійності. За своєю природою синоптичні дані доступні протягом декількох хвилин або годин після спостереження, щоб надати дані для прогнозів, тоді як кліматичні дані можуть бути відкладені від годин до місяців. Ці відмінності в часі доставки є результатом різних рівнів контролю якості (Quality Control – QC). Без значного контролю якості користувач може передати великий відсоток відсутніх, а також, можливо, неправильних значень даних. Синоптичні дані ставлять під загрозу точність.

Через різницю в часі спостережень, при аналізі фактичних часових рядів, відмінності між синоптичними даними та кліматичними даними можуть бути досить великими. Наприклад, таблиця 1.1 ілюструє ці відмінності між двома наборами даних для Берлінського Темплогфа з 1 січня 1984 року по 31 грудня 2001 року (синоптичні дані до 1984 року були недоступні). Німецькі синоптичні дані спостерігаються щодня у двох 12-годинних блоках; максимальна температура спостерігається з 06.00 GMT до 18.00 GMT, а мінімальна температура спостерігається з 18.00 GMT до 06.00 GMT.

Таблиця 1.1 – Відмінності між синоптичними та кліматичними даними, Берлін Темплхоф, 1 січня 1983 р. – 31 грудня 2000 р. [13]

Відмінності в діапазоні	Кількість відмінностей для T _{min} (%)	Кількість відмінностей для T _{max} (%)
>10 °C	0,02	0,97
5,0 – 10 °C	0,41	1,31
2,0 – 5,0 °C	3,45	3,11
1,0 – 2,0 °C	4,78	5,74
0,5 – 1,0 °C	4,50	6,39
0,0 – 0,5 °C	6,96	9,42
будь-яка різниця	20,12	26,95

Кліматичні дані Німеччини спостерігаються за період від півночі до півночі. Різниця між синоптичними даними та кліматичними даними перевищує 1°C, що відбувається приблизно в 5% випадків.

Оскільки ринок метеорологічних деривативів продовжує розвиватися, NMS усвідомлюють необхідність надання кліматичних даних більш своєчасно. Для того, щоб задовольнити це, багато НМС зараз виробляють попередні кліматичні дані, на основі яких отримують остаточні офіційні кліматичні значення. Попередні кліматичні дані складаються із спостережень, які проходять обмежену процедуру контролю якості, але не ретельно відредаговані.

1.3 Історія станцій та їх метадані

Явні переваги існують для організацій, які проводять ретельний пошук історії станцій перед очищенням та покращенням даних. Переваги отримання високоякісної історії станцій полягають у тому, що вона дає аналітикові точне знання про те, коли виникли потенційні розриви та які фізичні причини розривів. Історична інформація, зібрана на станціях, зазвичай називається «метаданими».

Наявність майбутніх (поточних) метаданих також дуже важлива для прийняття обґрунтованих рішень щодо вибору станцій для погодних контрактів. Історична ціна торгівлі не має значення, якщо станція переміщується протягом періоду дії контракту і відбувається раптовий зсув у щоденних спостереженнях.

Здебільшого НМС не мають систем для інформування користувачів про майбутні зміни станцій. Це завдання вимагає знання організації окремих метеорологічних служб, робочих відносин з особами, відповідальними за зміни станцій, і вміння долати труднощі перекладу. Приватні постачальники зібрали та надають користувачам постійно оновлювану інформацію про історію станції.

Багато розривів є результатом не змін в інструменті або переміщення самого інструмента, а внаслідок фізичних змін навколишнього середовища інструменту. Ці зміни можуть бути такими ж великими, як будівництво будівель поблизу приладу чи встановлення нової стоянки, або, здавалося б, незначними, як видалення трави з-під датчика. Процеси звітності, створені різними НМС, не визначають ці зміни навколишнього середовища, які мають бути задокументовані; це створює недоліки в записі метаданих. Приватний сектор також розробив продукти для подолання цього недоліку. Це досягається шляхом застосування суворого підходу до дослідження метаданих, включаючи глибокі інтерв'ю з менеджерами аеропорту та персоналом місцевих метеорологічних служб.

1.3.1 Метадані станцій США

США зараз мають найповніші в світі метадані, доступні від NCDC через Інтернет; доступна інформація про рух станції (широта, довгота та висота), спостерігачів (NOAA, NWS, підрядники тощо) та адміністративні питання. Однак метадані NCDC не є повними, і важливі метадані можуть бути відсутні в архівах NCDC через те, що місцеві відділення не повідомляють про всі зміни та помилки під час введення даних в архів.

1.3.2 Міжнародні метадані

Офіційні метадані менш доступні в інших країнах, ніж у США. Метадані часто доступні в обмеженій кількості, у друкованому вигляді та важкодоступні. Наприклад, метадані доступні для Великобританії, але доступ до них займає багато часу, оскільки потрібні попередні дозволи уряду. Незважаючи на те, що Метеорологічний офіс Великобританії проводить програму модернізації, наразі для доступу до повних метаданих у Великобританії потрібно відвідати та шукати друковані матеріали в одному з трьох місць (залежно від станції): Брекнелл, Единбург або Белфаст.

Подібний процес необхідний для отримання голландських метаданих. Обмежена кількість німецьких метаданих доступна в електронному форматі, але для отримання всіх німецьких метаданих потрібен пошук у паперових копіях.

1.3.3 Доступність та вартість

Недоступність і вартість метеорологічних даних були одним із найбільших обмежуючих факторів у глобальному розширенні промисловості похідних погодних даних. Хоча дані про погоду в США коштують недорого та дуже легко доступні, багато національних метеорологічних служб погано обладнані для задоволення потреб галузі, а ціни часто встановлюються на рівнях, які роблять використання даних неекономічним для оцінки погодних ризиків.

Оскільки важливість даних про погоду за межами прогнозування стає все більш очевидною, НМС реагують по-різному. Деякі країни, як-от Нідерланди [14], перейшли до оновлення систем і впровадили політику безкоштовної обробки даних, тоді як такі країни, як Фінляндія [15], намагалися посилити контроль над своїми даними, обмеживши доступ.

Таблиця 1.2 – Доступність даних про погоду для окремих країн

Країна	Метеорологічні організації	Історичні дані (наявність)	Поточні дані (наявність)	Метадані (наявність)	Легкість доступу
Австралія	Австралійський, Бюро Метеорології	Так	Так	обмежений	Легко
Бельгія	Королівський метеорологічний Інститут Бельгії	Так	Так	обмежений	Дуже важко
Бразилія	Національний інститут Meteorologica (INMET)	Так	Так	обмежений	Важко
Канада	Довкілля Канади	Так	Так	обмежений	Легко
Фінляндія	Фінський метеорологічний інститут	Так	Так	Ні	Дуже важко
Франція	Метео-Франція	Так	Так	обмежений	Важко
Німеччина	Deutscher Wetterdienst	Так	Так	обмежений	Помірний
Італія	Італійські ВПС	Так	обмежений	Ні	Дуже важко
Японія	Японське метеорологічне агентство	Так	Так	Так	Помірний
Мексика	Метеорологічна служба Nacional	Так	Так	обмежений	Важко
Нідерланди	Королівські Нідерланди Метеорологічний інститут	Так	Так	обмежений	Помірний
Норвегія	Норвезький метеорологічний інститут	Так	Так	обмежений	Помірний
Іспанія	Національний інститут Meteorologica (INM)	Так	обмежений	обмежений	Помірний
Сполучене Королівство	Метеорологічний офіс	Так	Так	обмежений	Помірний
Сполучені Штати	Національна служба погоди	Так	Так	Так	Дуже легко

Права на використання даних варіюються від обмежувальних угод з кінцевим користувачем до безкоштовного розповсюдження всіх продуктів. Вартість даних може варіюватися від незначних сум до тисяч євро на станцію для історичних даних.

Права на використання даних і вартість даних стають все більш гострими питаннями серед різних НМС. Для вирішення цих питань у Європі в 1995 році було створено ЕСОМЕТ (група економічної погоди, що складається з представників 20 європейських країн-членів) [16]. Заявленою метою ЕСОМЕТ є «збереження вільного та необмеженого обміну метеорологічною інформацією між НМС для виконання їхніх операційних функцій у рамках правил ВМО та забезпечення найширшої доступності основних метеорологічних даних та продуктів для комерційного застосування». Теоретично, будь-який член ЕСОМЕТ повинен діяти як «єдине вікно» для європейських метеорологічних даних.

На практиці це просто не так. У глобальному масштабі, щоб отримати історичні та поточні кліматичні дані, необхідно зв'язатися з кожною країною окремо або співпрацювати з організацією приватного сектора для захисту даних. Кожна країна встановлює різні обмеження, що стосуються використання даних, перерозподілу та вартості даних. Таблиця 1.2 (див. вище) ілюструє поточну доступність даних для вибраних країн.

Частота помилок у США та європейських країнах низька, зазвичай менше 1%, з широким діапазоном від майже відсутності помилок до понад 2%, але деякі великі помилки трапляються.

РОЗДІЛ 2. ОЧИЩЕННЯ ПОГОДНИХ ДАНИХ

На жаль, «необроблені» дані про погоду, отримані від НМС, можуть бути відсутніми або повідомлятися неправильно. Джерелами відсутніх погодних даних може бути те, що прилад був несправний, а дані про це ніколи не записувалися, що була перерва в передачі погодних даних або що дані про погоду були записані та архівовані, але згодом втрачені.

Помилки в даних про погоду часто виникають через погано відкалібровані прилади, але також можуть бути викликані помилками під час запису даних або під час оцифрування старих паперових записів. У будь-якому випадку ці дані необхідно «очистити», перш ніж можна буде провести точний аналіз оцінки.

Дані про погоду подаються в різних форматах, METAR і SYNOPTIC, які є двома поширеними форматами. Формат METAR – це стандартний міжнародний формат звітності погодинних значень у метричних одиницях. Додаткові добові значення максимальної та мінімальної температури повідомляються опівночі за місцевим стандартним часом у США, надаючи найранішу інформацію ринку похідних погодних інструментів про значення попереднього дня.

Формат SYNOPTIC є міжнародним стандартним форматом для звітності значень у метричних одиницях у стандартні синоптичні години, як обговорювалося в попередньому розділі. Типовою помилкою звітування було б пропуск символу під час запису даних. Згодом усі значення будуть неправильно інтерпретовані.

Процедури контролю якості, застосовані до самих даних, виявлять інші типи помилок. Помилки включають добову мінімальну температуру, що перевищує максимальну добову температуру, та інші невинуваті значення в певний день, суперечливі значення з різних джерел і невідповідність значень послідовних днів.

Частота, тип і величина помилок залежать від країни, історичного періоду часу, мережі та типу погодних даних. Наприклад, у Китаї та Японії існує сувора система контролю якості для своєї системи синоптичної звітності в реальному часі, і виявлено мало помилок, тоді як дані з інших країн можуть містити відсутні та помилкові значення. Частота помилок у США та європейських країнах низька, зазвичай менше 1%, з широким діапазоном від майже відсутності помилок до понад 2%, але деякі великі помилки трапляються.

Очищення даних про погоду складається з двох процесів: заміни відсутніх значень і заміни помилкових значень. Ці процеси слід виконувати одночасно, щоб отримати найкращий результат.

Усунути чи замінити одне відсутнє денне значення досить легко. Однак проблема стає набагато складнішою, якщо є блоки щоденних відсутніх значень. Такі випадки не рідкість, особливо кілька десятиліть тому. Проблема очищення даних тоді стає проблемою заміни значень інтерполяцією між спостереженнями на кількох станціях (просторова інтерполяція) та інтерполяцією між спостереженнями в часі (часова інтерполяція).

2.1 Просторова інтерполяція

Існує багато потенційних підходів до просторової інтерполяції випадкових метеорологічних даних. Однією з перших робіт з інтерполяції в метеорологічній області була [17]. Метод інтерполяції тут коригував значення точки сітки, лінійно поєднуючи залишкові значення між прогнозованим і спостережуваним значенням. Автор [18] розширив цю роботу, використовуючи лінійну комбінацію самих спостережень для створення поля «першої здогадки». Потім це поле використовувалося як вхід для другого проходу даних. Цей підхід все ще використовується сьогодні при створенні об'єктивних полів аналізу для вхідних даних у числові метеорологічні моделі.

Згадані методи розроблені для інтерполяції набору значень сітки з випадкового набору вхідних даних, але можуть бути адаптовані для

інтерполяції будь-якого нестандартного значення точки сітки та для включення додаткових змінних, таких як кореляції точка-точка.

З появою технології географічної інформаційної системи (ГІС) розширилося дослідження методів інтерполяції для створення нелінійних поверхонь, з яких можна отримати будь-яку точку даних із географічними прив'язками. Ця технологія застосовна до проблеми заміни відсутніх погодних даних, оскільки просторова поверхня багатьох погодних змінних наближається до нелінійних поверхонь [19].

Однією з найпростіших форм інтерполяції, коли враховується лише відстань, є зважене усереднення в квадраті оберненої відстані або оберненої відстані. Цей метод добре працює для досить однорідних даних, таких як добова температура, коли є досить рівномірний розподіл даних навколо цільового місця. Цей метод погано працює для більш складних рішень, наприклад щоденних опадів.

Кілька підходів лінійної регресії також можуть бути застосовані, коли є більше вхідних даних, що визначають форму поля даних. Наприклад, такий підхід є кращим у порівнянні з методами, зваженими з оберненою відстанню, при інтерполяції кількості опадів як функції відстані, висоти та місяця року, оскільки цей підхід часто забезпечує кращу «підгонку» даних і його легше обчислити.

Нові методи просторової інтерполяції максимальних і мінімальних значень температури використовують методику штучної нейронної мережі (Artificial Neuron Network – ANN) [20]. Хоча ці підходи показують кращі статистичні результати в порівнянні з простими методами «зменшення» значень температури прогнозу в сітці до певних місць, вони не були ретельно перевірені на застосування нерівномірно розташованих вхідних даних і не були ретельно перевірені на інші погодні змінні, окрім температури.

2.2 Часова інтерполяція

Використано методи часової інтерполяції з використанням статистичних функцій розподілу ймовірностей. Автори робіт [21] та [22] використовують такі методи для відновлення щоденних часових рядів канадських опадів. Методи ієрархічної поліноміальної регресії також використовувалися для заповнення відсутніх даних про температуру та вітер для сільськогосподарських імітаційних моделей [23].

Такі методології можуть бути задовільними для історичного очищення невеликих прогалин відсутніх даних, але надають невелику цінність очищення даних в реальному часі, необхідного для найближчого розрахунку похідних погодних інструментів. Таким чином, для очищення в режимі реального часу переважні методи просторової інтерполяції.

2.3 Складності при очищенні даних про погоду

Фактичне застосування методів очищення даних ускладнюється тим фактом, що країни звітують із різними щоденними умовами часу. Наприклад, у США календарний день північ – північ за місцевим стандартним часом, канадський день – 06.00–06.00 GMT, а Велика Британія повідомляє про 09.00–09.00 GMT, при цьому мінімальна температура, яку повідомлялося, фактично припадала на попередній день.

Часи звітності також іноді відрізняються для різних параметрів. Наприклад, США повідомляють про щоденні максимальні/мінімальні температури опівночі за місцевим часом і повідомляють про цілодобові дощі о 12.00 GMT. Ці невідповідності у звітності вимагають розроблення спеціальних методів очищення даних для кожної змінної погоди та кожної країни.

Іншим ускладнюючим фактором очищення даних про погоду є те, що нерозумно вважати, що найближчі точки корелюють краще, ніж віддалені. Проблеми мікроклімату часто дають неадекватні результати просторової

кореляції при використанні стандартних методів географічного просторового очищення. Оптимальні методи, такі як модифікована техніка [18], оцінюють історичну кореляцію та відстані між станціями при розробці підходу до просторового очищення.

Ще одним ускладнюючим фактором є те, що кожна змінна погоди має свої унікальні характеристики просторової кореляції. Наприклад, кореляція кількості опадів з відстанню однозначно відрізняється від кореляції температури або вологості з відстанню.

Враховуючи всі ці вимоги до інтерполяції даних про погоду, модель інтерполяції даних повинна вирішувати проблему оцінки значень у нестандартних віднесених точках, використовуючи для отримання вагових коефіцієнтів як відстань, так і кореляцію погодних змінних.

2.4 Вплив зміни/переміщення приладів

Неупереджені дані про погоду мають вирішальне значення при розробці моделей для використання в бізнесі з погодними ризиками. Приватний сектор здійснив проекти для усунення розривів в історичних даних про температуру в США та інших країнах. Кліматологи вже давно визнали, що більшість наборів кліматичних даних містять неоднорідності або розриви, викликані некліматичними факторами, такими як зміна приладів, зміни фізичного розташування станції, зміни навколишнього середовища, зміни процедури експлуатації, людські помилки та зміни операторів.

Оскільки ці некліматичні «сигнали» є «шумами» для кліматичних досліджень, кліматологи довгий час намагалися виправити ці неоднорідності. Завдання ускладнюється випадковими несправностями інструментів (наприклад, поломка вентилятора, що спричиняє швидке надмірне нагрівання) або «вихід» з калібрування (наприклад, бруд і болото можуть викликати повільну тенденцію нагрівання/охолодження або «зміщення калібрування» інструмент).

У багатьох випадках ці інциденти тривають кілька тижнів або кілька місяців до наступного технічного обслуговування. Ці несправності зазвичай не є критичними, але вони створюють додатковий «шум» при виявленні розривів.

Як ілюстрація, на рисунку 2.1 показані річні індекси включення опалення (Heating Degree Day – HDD) для гіпотетичної ситуації, коли передбачається, що перерва вимірювання низької температури відбудеться в 1995 році, що спричинить зсув вгору в спостережуваних HDD до 1995 року. Нижній ряд розраховується на основі очищених даних, а верхній ряд оцінюється після врахування розривів; були накладені як очищені, так і покращені середні значення. Мета покращення даних полягає в тому, щоб усунути розриви та підлаштувати ряди погоди до сучасних спостережень. Оцінки середнього індексу та стандартного відхилення будуть упередженими, якщо не усунути розриви. Як наслідок, аналіз тенденцій також буде упередженим.

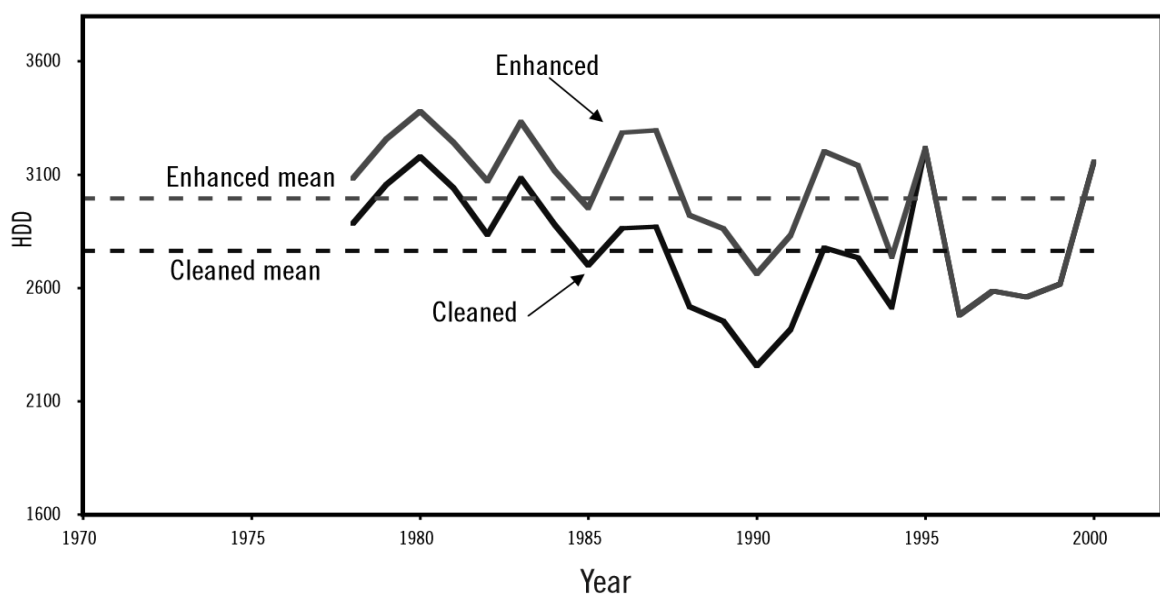


Рисунок 2.1 – Очищені та покращені річних рядів HDD

Вплив розривів на оцінку погодних контрактів може бути великим. Наприклад, 10-річний коефіцієнт для погодного контракту на опалення відповідно до HDD з листопада по березень може змінитися на цілих 60 градусів-днів, якщо в даних у 1995 році присутній розрив $0,8^{\circ}\text{F}$ ($0,8^{\circ}\text{F}$ – це

середня величина розриву, виявлена [24] під час аналізу різниці температур між спостереженнями за допомогою нещодавно встановленої автоматизованої системи спостереження за поверхнею (ASOS) з попередніми вимірюваннями). Це упередження може призвести до великого розкиду в оцінці, якщо припустити, що типова номінальна вартість становить 5000 доларів США за градус-день.

2.5 Аналіз стану метеорологічних станцій у Сполучених Штатах Америки

У роботі виявили під час дослідження 200 американських станцій першого порядку [25], які:

1. Майже всі станції мають хоча б один розрив неперервної роботи за останні 50 років.
2. Понад 50% цих станцій мають принаймні один розрив за останні 10 років.
3. Деякі станції мають розриви, що перевищують $2,5^{\circ}\text{F}$ в часовому ряді мінімальної або максимальної температури.

Ці висновки не є дивними при огляді історії цих станцій. З 1950 року на американських станціях першого порядку повідомлялося в середньому 13 подій, які були потенційними причинами розривів. Не менш важливим є те, що за останні 10 років темпи подібних подій знизилися незначно. З 1992 року було встановлено, що в середньому три зареєстровані події на станцію є причинами потенційних розривів.

За останні 40 років у США відбулося кілька національних змін приладів. Показання температури гігротермометра HO63 поступово замінили максимальні/мінімальні термометри як офіційні реєстратори максимуму/мінімуму між серединою 1960-х і початком 1980-х років, хоча спочатку HO63 часто використовувалися лише для запису погодинної температури. Нове покоління інструментів, HO83, встановлене між 1982 та

1986 роками, як правило, має теплий ухил. Автори [26] повідомили, що HO83 мали в середньому тепле зміщення $+0,57^{\circ}\text{F}$ порівняно з каліброваним польовим стандартом.

У 1990-х роках відбулася велика програма встановлення приладів типу ASOS для приблизно 850 станцій (як NWS першого порядку, так і другого порядку Airway). У багатьох аеропортах, одночасно з встановленням ASOS, прилади, як правило, були перенесені з районів поблизу Офісу метеорологічної служби до кінця злітно-посадкової смуги, як правило, у трав'яних районах. Наприклад, пристрій у Шарлотті, штат Північна Кароліна, був переміщений у 1998 році з ділянки поблизу бетону між офісом метеорологічної служби та авіаційним терміналом Батлера на низьку трав'яну ділянку приблизно на 500 футів перпендикулярно до кінця однієї з основних злітно-посадкових смуг. Через це переміщення було спричинено екстремальний розрив, що перевищує 2°F , що добре видно на рисунку 2.2, де показано місячний числовий ряд різниці температур між станцією в Шарлотті та двома сусідніми станціями.

Дослідження показали, що встановлення ASOS та переміщення приладів призвели до охолодження записів у багатьох місцях тобто для вибіркового дослідженні 76 станцій (не всіх станцій першого порядку) середня різниця температур між приладами ASOS і HO83 є охолодженням від $-0,79^{\circ}\text{F}$ з діапазоном від $-2,56^{\circ}\text{F}$ до $+0,61^{\circ}\text{F}$.

Нещодавнє опитування менеджерів ASOS на 200 станціях першого порядку показує, що в період з лютого 2001 року по лютий 2002 року на 10 станціях відбулися події, які можуть призвести до розривів (здебільшого переміщення станцій у зв'язку з будівництвом і розширенням аеропорту), а також швидкість роботи станцій, які не працюють.

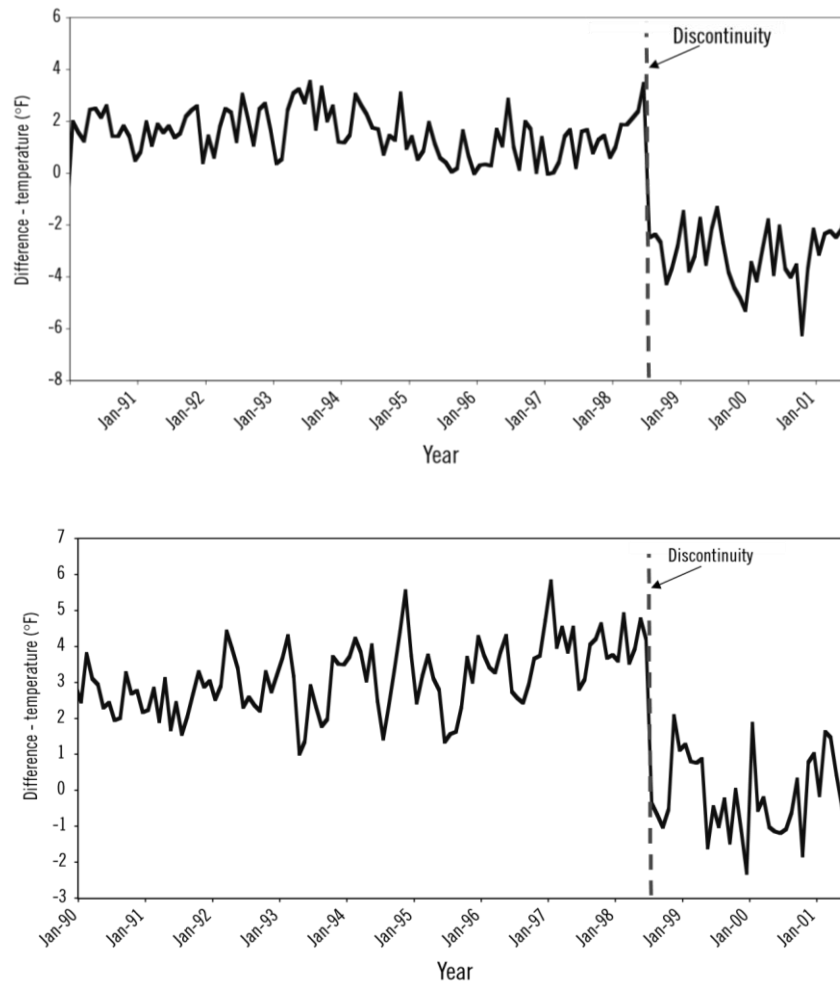


Рисунок 2.2 – Річна відмінність в показах температури між станцією в м. Шарлотта (США) та сусідніми станціями

Зміни протягом наступних 12 місяців будуть лише трохи меншими, ніж рівень подій, які відбулися за останні 12 місяців.

2.6 Аналіз стану метеорологічних станцій у інших країнах

Дослідження авторів [27], [28] також повідомляють про існування розривів у міжнародних метеорологічних даних. Розриви також викликані такими подіями, як зміна приладів і зміна корпусу станції (наприклад, нещодавнє встановлення напівавтоматичних станцій метеорологічного спостереження SAMOS у Великобританії).

В Японії огляд метаданих показує, що до початку 1990-х років відбулося дуже мало переміщень, після чого кількість переміщень станцій зросла.

В Австралії загальна політика Бюро метеорології полягає у будівництві нового майданчика, як правило, в аеропорту. Потім старий сайт, як правило, міський, ліквідується. Новим сайтам присвоюються нові ідентифікаційні номери. Якщо є необхідність забезпечити безперервний запис для двох ділянок, виконується порівняння історії опадів і температури, щоб усунути будь-яку однорідність, викликану переміщенням. Крім того, в аеропортах трапляються періодичні переміщення приладів, головним чином для будівельних робіт.

2.7 Методи опрацювання розривів в серіях погодних даних

Існує ймовірність того, що будь-які зміни приладів, оточення та переміщення приладів можуть створити серйозні розриви в ряді даних. Ідеальною обробкою рядів даних про погоду, що містять розриви, було б реконструкція історичних спостережень так, ніби поточні умови приладів і навколишніх станцій були однаковими протягом історичного періоду спостережень. Таке завдання дуже складне, оскільки вплив розривів на виміряні погодні змінні «розбавляється» впливом кліматичних змін. Хорошим рішенням є використання найкращих характеристик кількох методів для ідентифікації та кількісної оцінки розривів у погодних рядах.

Методи слід застосовувати до кожної змінної погоди окремо. Наприклад, кліматичні та некліматичні фактори по-різному впливають на часові ряди мінімальної та максимальної температури. Готуючись видалити розриви з ряду, необхідно відповісти на три питання:

1. Які можливі дати розриву?
2. Чи справді відбувся розрив?
3. Яка величина розриву?

На перше питання можна відповісти за допомогою інформації, доступної на станції (метаданих), або шляхом виконання додаткових статистичних тестів

для серії, якщо метадані є неповними, а це є нормою, як правило. На друге питання можна відповісти поєднанням суб'єктивних і об'єктивних методологій. На третє питання найкраще відповісти за допомогою математичних і статистичних моделей.

2.7.1 Визначення потенційних дат розривів

Метадані є неоціненними для визначення дати розривів. У метаданих повідомляється про два загальні класи подій:

- події, пов'язані з переміщенням станції або зміною приладів (так звані «правдиві» події)
- події, описані як адміністративні зміни.

Прикладом адміністративної зміни може бути перейменування аеропорту. Хоча адміністративні зміни не повинні впливати на погодні дані, існує обмежена кількість випадків, коли справжні події відбувалися випадково в ці дати. Отже, методи ідентифікації та кількісної оцінки розривів спочатку аналізують поведінку даних до і після справжніх подій. Подальший аналіз поведінки даних до та після адміністративних подій потім проводиться з суворішими критеріями прийняття, ніж для справжніх подій.

В ідеальній ситуації всі розриви пояснюють правдиві та адміністративні події, зазначені в метаданих. На жаль, в реальних умовах це не так. Не всі події повідомляються, і тести потрібно проводити на дати, відмінні від тих, які вказані в метаданих. Одним з підходів є виконання тестів щомісяця для виявлення потенційних розривів з використанням більш суворих критеріїв прийняття, ніж ті, що використовуються для справжніх або адміністративних подій.

2.7.2 Виявлення розривів

У кількох роботах узагальнено методи опрацювання розривів (наприклад, [29], [30]). Ці методи намагаються виявити та/або оцінити величину розривів, використовуючи лише дані з самої станції (одна станція), або шляхом розробки еталонних рядів із сусідніх станцій (кілька станцій), таким чином усуваючи вплив регіональних кліматичних сигналів.

Сам ряд може бути використаний для виявлення великих розривів після обробки даних. Такі прийоми, де фільтри застосовуються до ряду, або ті, де квантилі розподілу ряду нанесені на графік за різні періоди часу, можна використовувати для виявлення великих розривів більше ніж 2° . Наприклад, на рисунку 2.3 показано помісячний ряд температур у Шарлотті, Північна Кароліна, приблизно в 1998 році, коли було виявлено значний розрив. Хоча розрив нелегко виявити шляхом візуального огляду помісячних рядів, його легше побачити на накладеному графіку 12-місячного ковзного середнього.

Переважним рішенням для виявлення некліматичних сигналів у кліматологічній серії («цілі») є розробка другої серії («еталонної»). Посилання, як правило, є зваженим агрегатом часових рядів із сусідніх станцій відносно цільової станції. Різниця (температура) або співвідношення (опадів) між цільовим і контрольним рядами намагається зменшити, якщо не усунути, більшість кліматологічних сигналів.

Використання еталону при виявленні розривів у цілі передбачає, що контрольний ряд включає регіональні кліматичні тенденції та коливання, присутні в даних цілі, і що контрольний ряд не містить самих розривів протягом періоду аналізу розрив у цілі.

На практиці вкрай рідко зустрічаються сусідні станції, які мають точно такі ж регіональні та місцеві кліматологічні сигнали, що й цільова станція, і в той же час не містять жодних розривів.

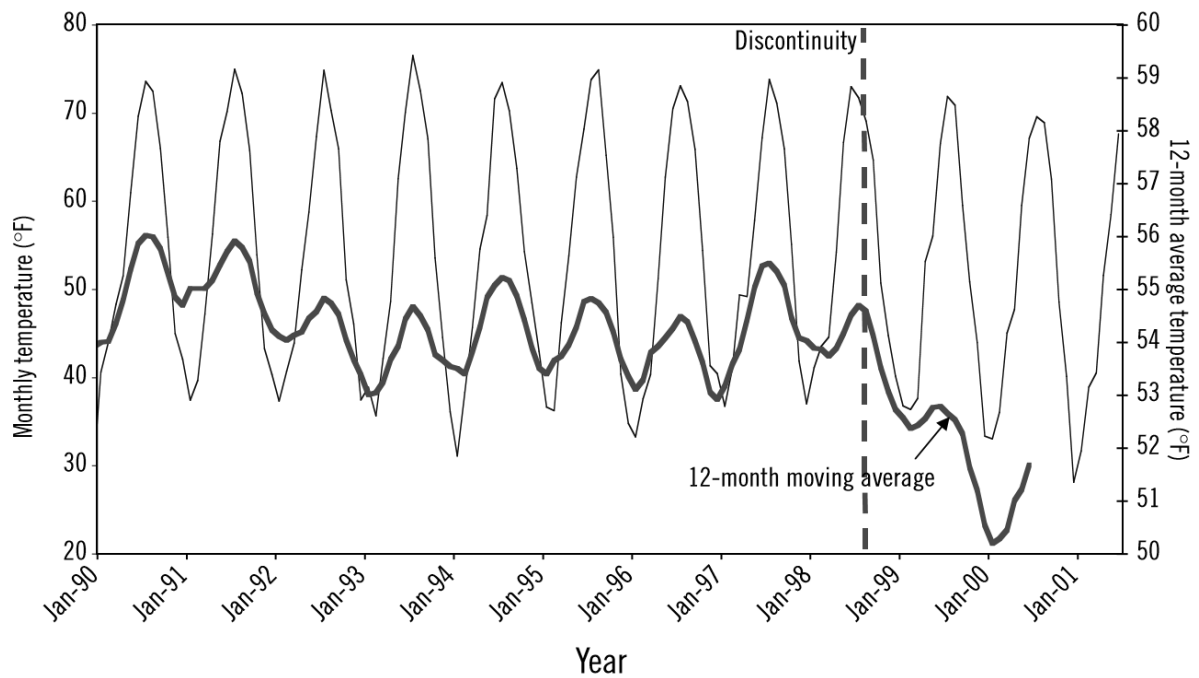


Рисунок 2.3 – Місячні значення температури для міста Шарлотта,
Північна Кароліна. США

Потрібна достатня кількість станцій, щоб значний розрив у часовому ряді однієї з сусідніх станцій не був значущим в агрегованих еталонних рядах і не перешкоджав ідентифікації розривів на цільовій станції.. Однак, якщо використовується занадто багато віддалених (або менш корельованих) сусідніх станцій, отримане посилення може не відобразити належним чином справжні кліматичні коливання цільової станції.

Простого тесту на відстань недостатньо, тому що найближчі станції можуть погано корелювати, тоді як інші станції можуть зазнавати подібних кліматичних умов, що й цільова станція. Тому першим кроком є вибір станцій, які сильно корелюють із цільовою станцією. Однак дві станції з сильною кореляцією можуть означати, що кореляція пов'язана з віддаленими телез'єднаннями. Отже, аналітик також повинен використовувати деяку функцію відстані або регіональну перевірку клімату при виборі опорних станцій.

З наведених вище причин довідки розробляються шляхом зважування значень з набору сусідніх станцій. Кожна вага є функцією відстані та коефіцієнта кореляції.

Після того, як довідник розроблено, основними проблемами у виявленні та кількісній оцінці розривів є:

- Як можна впевнено відрізнити сигнал від шуму?
- Які періоди часу слід перевірити до і після підозрюваного розриву?
- Що робити, якщо видимий сигнал змінюється з часом?
- Які тенденції в основних даних і як їх слід вирішувати?
- Які найменші значення можна виявити для розриву?
- Що потрібно зробити, щоб виявити розриви, близькі один до одного в часі?
- Що потрібно зробити, щоб усунути розриви на сусідніх станціях?
- Що потрібно зробити, щоб усунути мінливість змін у різні періоди часу?

2.8 Методи визначення результатів та коригування даних про погоду

Зазначені вище питання вирішуються як суб'єктивними, так і об'єктивними методами. Суб'єктивні методи мають кількісний компонент, але характеризуються остаточним судженням досвідченого метеоролога, тоді як об'єктивні методи – це повністю автоматизовані процедури без втручання людини.

2.8.1 Суб'єктивні методи

Суб'єктивне судження досвідчених метеорологів важливе для зміни ваги, наданої різним сусіднім станціям на основі багатьох факторів, які не можуть бути систематично запрограмовані. Наприклад, візуальний огляд різницевого ряду між цільовою станцією та її сусідніми станціями може дати вказівку на

якість даних сусідніх станцій та якість метаданих. Ці знання можуть призвести до суб'єктивної зміни параметрів тестів для визначення та кількісної оцінки розривів, виявлених за допомогою об'єктивних методів.

Перевага суб'єктивних методів полягає в тому, що вони надають метеорологу прості візуальні інструменти для швидкого визначення періодів часу, протягом яких може існувати одна або кілька розривів. Прийняття або відхилення цих розривів приймається після перегляду результатів, отриманих в результаті об'єктивного аналізу. Одним із перших розроблених суб'єктивних методів був «аналіз подвійної маси». Він відображає кумулятивні значення різниці температур у часі між метою та її еталом. Наприклад, на рисунку 2.4 показаний часовий ряд різниць між Шарлоттою та однією з сусідніх станцій і накладені кумулятивні відмінності. Якщо спостережувані температури вели себе однаково на обох станціях, то кумулятивна різниця між двома станціями рівномірно збільшується з часом. Графічно кумулятивні різниці температур у часі були б майже лінійними, що спостерігається до дати розриву.

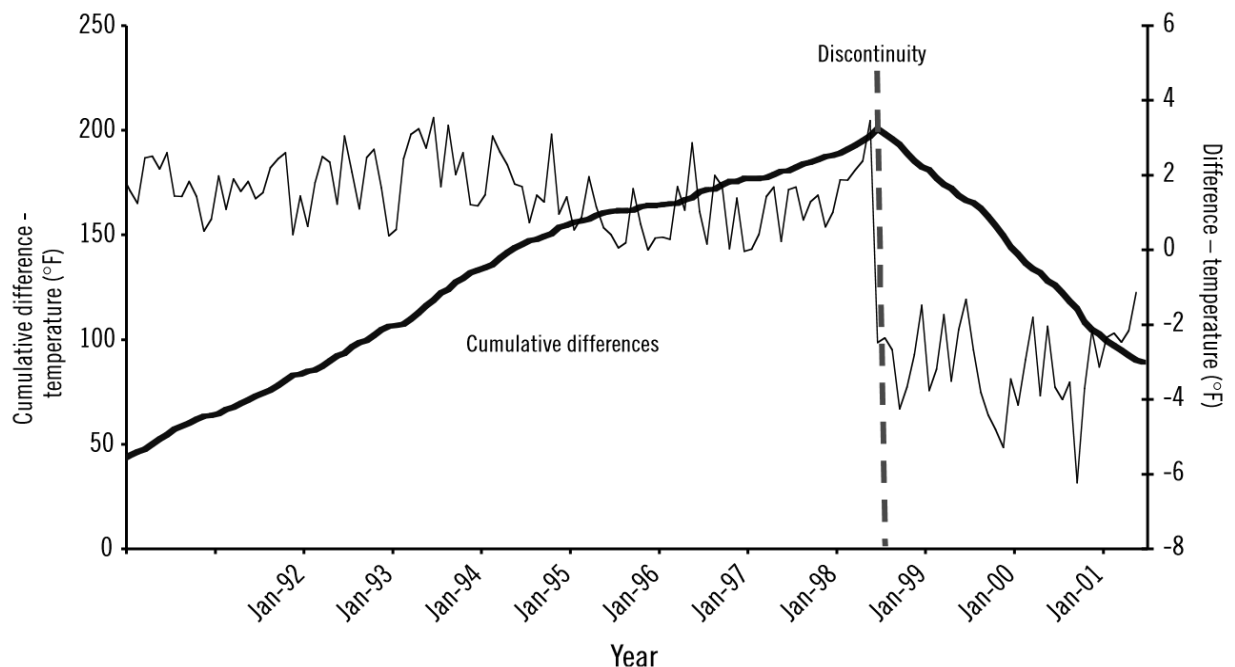


Рисунок 2.4 – Кумулятивні різниці температури між двома метеостанціями

Графічний аналіз залишків між спостережуваними значеннями цільового ряду та прогнозованими значеннями того самого цільового ряду, отриманий за допомогою аналізу лінійної регресії між цільовим показником та його посиленнями, корисний для виявлення розривів, оскільки графік залишків від часу буде показувати раптові зміни, якщо такі розриви існують. Також поширеним є припущення, що залишки є незалежними випадковими величинами з нормальним розподілом нульового середнього та постійного стандартного відхилення.

Розриви можна також візуально виявити шляхом побудови ліній регресії на ряді до та після певної дати. На рисунку 2.5 показано різницевий ряд між Шарлоттою та однією з сусідніх станцій і накладені дві лінії регресії до і після розриву 1998 року та лінію регресії за обидва періоди. Якби не було розривів, дві лінії регресії теоретично були б такими ж, як і лінія регресії для всього ряду.

Візуальний огляд цих ділянок зазвичай дозволяє успішно виявити великі розриви (більше $1,5^{\circ}\text{F}$) у межах серії. Візуальний огляд також дає «підказки» для значно менших розривів, які будуть далі аналізуватися за допомогою об'єктивних методів.

2.8.2 Об'єктивні методи

Більшість «окремих» об'єктивних методів можуть виявити окремі великі розриви ($1,5^{\circ}\text{F}$ або більше) з точністю понад 95% (99% для найкращих методів) у кліматологічних рядах. Рівень успіху зменшується приблизно до 50% для виявлення розривів порядку $0,5^{\circ}\text{F}$. Тому важливо не покладатися лише на об'єктивні методи при виявленні розривів.

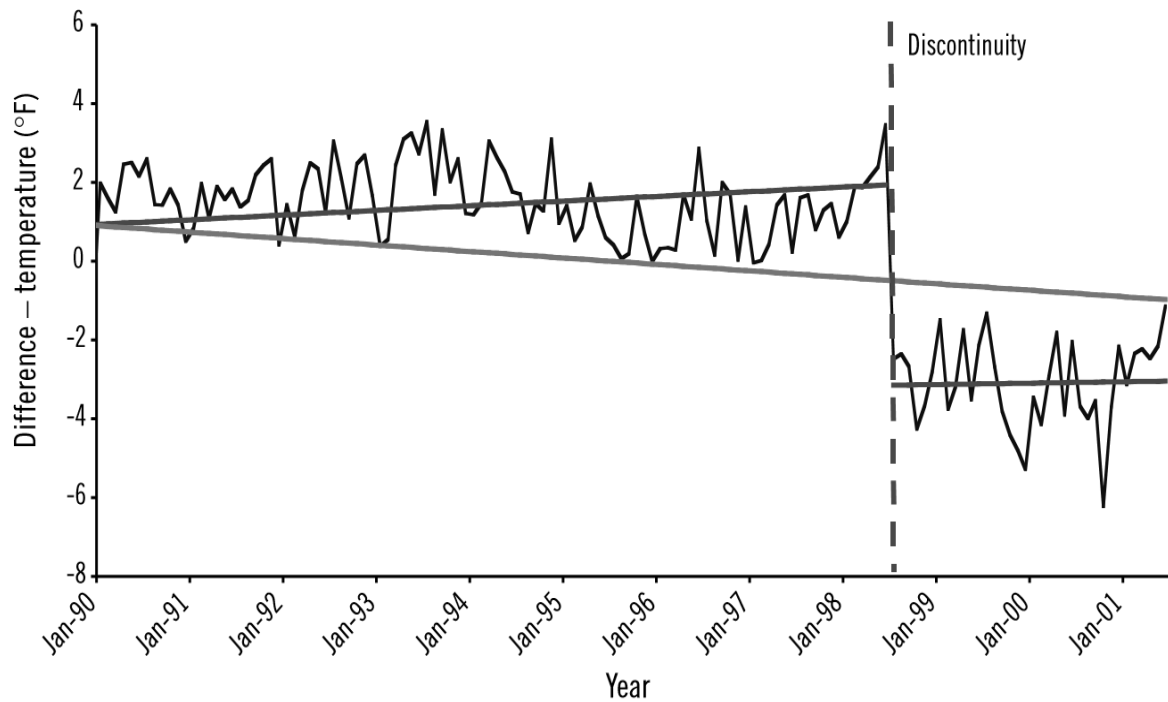


Рисунок 2.5 – Ряди різниці температур між сусідніми метеостанціями

Типовими об'єктивними методами є математичні інструменти, які виявляють час і величину розривів шляхом перевірки передбачуваних статистичних властивостей різниці (або співвідношення) передбачуваних справді однорідних цільових і контрольних рядів. Більшість статистичних тестів сформульовано для автоматичного виявлення розриву в різницевих рядах (або співвідношеннях) на певну дату, оскільки параметри розподілу ряду (або в часовій, або в частотній областях) значно змінилися статистично до і після цієї дати. Тести засновані на нульовій гіпотезі про те, що ряд є однорідним (наприклад, без змін середніх значень різницевого ряду до і після дати) і на альтернативній гіпотезі, що ряд стає неоднорідним на певну дату. Ці типи тестів використовуються в методах, розроблених у, наприклад, [31] та [32].

Більш новітні методик [33] та [34] одночасно генерують еталонні ряди та виявляють розриви. Ці методи спочатку визначають кілька дат, коли можливі розриви для набору рядів в одній кліматичній області. Кожен з цих рядів може потім використовуватися, як однорідне посилення на інші ряди для проміжків

часу, де розриви не виявлені. Коли розрив у серії узгоджено та знайдено з посиланнями, отриманими з інших серій, розрив приписується цій серії.

Деякі розриви можна легко виявити, оскільки вони спричиняють значну постійну зміну ряду різниці (або співвідношення), але інші очевидні розриви є короткочасними зрушеннями, які є результатом змін мікроклімату або зрушень внаслідок тимчасових несправностей приладів. З цих причин серії необхідно проаналізувати протягом достатнього періоду часу до та після потенційного розриву, щоб встановити реальний розрив. Як правило, для надійної кількісної оцінки розривів потрібно кілька років. Однак можна виявити та зробити початкові оцінки величини розривів протягом 15 місяців від дати розриву, а також навіть раніше, якщо є великі розриви та перевіряються підготовленими метеорологами.

В ідеалі не повинно бути розривів на сусідніх станціях приблизно в той час, коли буде виявлено розрив на цільових станціях. Якщо це станеться, розриви в цілі не будуть виявлені або виявлені помилково. Крім того, справжні розриви будуть «згладжені» на кількох станціях. Цих проблем можна запобігти шляхом аналізу різницевого рядів цільових станцій з окремими сусідніми станціями та шляхом оцінки величини розривів без використання цих сусідніх станцій.

Оскільки важливо з високим рівнем успіху виявити всі розриви, будь-яка методологія гомогенізації або покращення повинна спиратися на комбінацію як суб'єктивних, так і об'єктивних методів. Оптимальним методом є вибір знайдених розривів усіх величин, підтверджених метаданими та відгуками операторів станцій. Цю вибірку потім можна використовувати для розробки статистичних критеріїв для об'єктивних методів і критеріїв найкращого застосування суб'єктивних методів. Цей процес займає багато часу і відрізняється від звичайних методів гомогенізації, які застосовуються метеорологами, які вивчають кліматичні дані, але він потрібен для покращення

даних про погоду, які використовуються при оцінці операцій, залежних від погодних умов.

2.8.3 Кількісна оцінка розривів

Після визначення дати розривів було б ідеально, якби були доступні «лабораторні» експерименти для оцінки величини (зсуву) цих розривів. Наприклад, порівняння часових рядів, що перекриваються, між старим і новим приладами може виявити розриви, коли відбувається зміна або переміщення приладу. На жаль, ці порівняння проводяться лише на дуже обмеженій кількості станцій і зазвичай протягом дуже короткого періоду часу.

Величина розриву зазвичай розраховується як різниця середніх значень до і після дат, на які були виявлені розриви. Середні значення оцінюються протягом кількох років, оскільки ряди різниць насправді не є стаціонарними в часі. Вплив розривів на температурні ряди може залежати від сезону. Однак історичні детальні метеорологічні умови в будь-який час і фізичні причини існування розривів часто невідомі.

Після того, як розриви були застосовані, доцільно повторно запуснути методологію, щоб побачити, чи не були повністю усунені розриви під час першої ітерації методології.

2.9 Перевірка методологій для визначення та кількісної оцінки результатів

Відмінним способом здійснити валідацію методів є виконання попарних тестів, використовуючи ту саму методологію, що використовується при виявленні розривів, для систематичного порівняння цільової станції з декількома сусідніми станціями. Як суб'єктивні, так і об'єктивні тести необхідно виконувати окремо для цільової серії та найбільш корельованих сусідніх станцій, щоб переконатися, що розриви є правдивими та що їх значення

не пошкоджені. Якщо можливо, слід проводити інтерв'ю з операторами станцій для перевірки розривів, особливо протягом останніх 10 років. Проте, для розривів старше 10 років отримати такий зворотний зв'язок може бути проблемним.

Нарешті, гомогенізація будь-якої цільової станції не повинна кардинально змінити регіональну кліматологію. Ранговий тест зазвичай проводиться на щомісячних даних станції, щоб переконатися, що кліматологія станції не змінюється. Якщо кліматологія станції змінюється, наприклад, якщо в результаті гомогенізації створюються нові місячні екстремальні температури, тоді необхідні додаткові випробування для підтвердження.

РОЗДІЛ 3. БЕЗПЕКА ЖИТТЄДІЯЛЬНОСТІ, ОСНОВИ ХОРОНИ ПРАЦІ

3.1 Аналіз небезпеки і шкідливості при розробці програмного забезпечення

Організація робочого місця розробника ПЗ впливає на його працездатність.

У своїй діяльності розробник використовує комп'ютер, пристрої збереження інформації, а тому є необхідність забезпечення зручного доступу до всіх технічних засобів. Тому в даному розділі докладніше розглянемо відомості про систему ергономічних норм і принципів організації робочого місця, на котрому проводяться роботи зі створення модуля збору статистики.

Під робочим місцем розуміється зона, оснащена необхідними технічними засобами, у якій відбувається трудова діяльність виконавця або групи виконавців, які спільно виконують одну роботу або операцію.

Організація робочого місця полягає у виконанні заходів, які забезпечують безпечний і раціональний трудовий процес і ефективне використання знарядь та предметів праці, що підвищує продуктивність праці і знижує стомлюваність працівника.

Організація робочого місця залежить від характеру розв'язуваних задач і особливостей предметно-просторового оточення, що визначають робоче положення тіла і можливість пауз для відпочинку, типи і способи засобів відображення і керування, необхідність у засобах захисту, спецодягу, простору для налагодження і ремонту устаткування.

Одним з компонентів діяльності на робочому місці є робочі рухи. Їхня раціональна організація створює умови для зниження стомлення, резерви для підвищеної працездатності. Просторові характеристики руху оператора визначаються траєкторіями руху і розмірами моторного поля (зони досяжності).

При організації робочого місця необхідно забезпечити нормальні умови огляду. Зону огляду описує кут, вершина якого знаходиться в центрі ока, а сторони складають границі, в яких людина при фіксованому положенні голови й ока добре розрізняє їхнє місцезнаходження.

У горизонтальній площині цей кут складає 300 – 400. При організації робочого місця кут огляду можна взяти 500 – 600, включаючи зону менш ясного огляду. Допустимий кут огляду по горизонталі 900. У вертикальній площині оптимальний кут огляду 100 вгору і 300 вниз від лінії погляду, а допустимий 300 вгору і 400 вниз від лінії погляду.

Щоб зберегти нормальну гостроту зору, робочу поверхню розташовують від очей на відстані від 0,3 м до 0,75 м. Робочі меблі повинні бути зручними для виконання робочих операцій. В даному випадку робочий стіл є основним устаткуванням. Особливо важливе значення має висота столу, його конструкція, яка повинна передбачати шухляди для розміщення інструментів, документації.

Важливе значення має конструкція робочих крісел. Погано підібрані крісла можуть бути причиною надмірної стомлюваності.

Нахил і висота крісла повинні регулюватися відповідно до висоти робочої поверхні і росту працюючого. Рекомендована ширина крісла 370 – 400 мм, глибина 370 – 420 мм, висота спинки 370 – 1000 мм від рівня крісла. Для розміщення ніг необхідно передбачити вільний простір під робочою площиною [35].

Праця людини, що протікає в умовах надмірного нервово-емоційного напруження, довготривалих статичних навантажень, обмеженої рухової активності призводить до неврозів, відхилень у психіці, захворювань опорно-рухового апарату, серцево-судинної системи тощо. Комп'ютери, телебачення, системи зв'язку та інші засоби, що використовують досягнення радіоелектроніки, є генераторами цілої низки електромагнітних випромінювань, вплив яких на організм людини ще не зовсім вивчений.

З широким впровадженням автоматизації та комп'ютеризації виникла потреба врахування психологічних можливостей людини, таких як швидкість реакції, особливості пам'яті та уваги, емоційний стан та ін. Поява операторської діяльності призвела до суттєвих змін у фаховій структурі праці. Зменшились фізична важкість праці, ризик виробничого травматизму, однак разом з тим, на працюючу людину посилюється вплив нових, раніше не відомих чи мало вивчених несприятливих виробничих факторів фізичного, хімічного і особливо психофізіологічного характеру.

Проте, розвиток сучасної обчислювальної техніки відбувається не лише у бік покращення її технічних параметрів, але також звертається увага безпеку використання цієї техніки людиною шляхом зменшення потужності випромінювачів, зменшення рівня випромінювання з моніторів, зменшення напруг живлення, покращення ергономічних характеристик.

Таким чином, в розділі з охорони праці виконано огляд питань безпечної роботи при створенні сайту та встановлено, що умови такої роботи відповідають вимогам з охорони праці, які застосовуються в галузі інформаційних технологій.

3.2 Інформаційно-психологічні небезпеки

Сучасні реалії постіндустріального суспільства, зумовлені значним ростом інформації, відкривають ще одну сферу життєдіяльності людини – інформаційну. Сучасні засоби комунікації і обробки інформації створили принципово нові умови існування людини, що зумовило появу грандіозного проекту об'єднання національних інформаційних і телекомунікаційних структур в глобальну інформаційну інфраструктуру.

Життєдіяльність людини реалізується одночасно зі світом природи і у специфічному для людського суспільства інформаційному середовищі, що має свої закономірності розвитку і функціонування. Інформаційна сфера стає такою ж важливою складовою суспільного життя, як економічна, виробнича,

побутова, політична, військова та ін. Нові інформаційні технології, засоби масової комунікації багатократно підсилили можливості впливу на свідомість і підсвідомість як окремої людини, так і на великі групи людей та населення країни загалом.

Інформаційна сфера – сукупність таких елементів:

- об'єкти інформаційної взаємодії чи впливу;
- особисто інформація, призначена для використання суб'єктами інформаційної сфери;
- інформаційна інфраструктура, що забезпечує можливість здійснення обміну інформацією між суб'єктами;
- суспільні відносини, що складаються у зв'язку з формуванням, переданням, розповсюдженням і збереженням інформації.

Особистість, активний соціальний суб'єкт, його психіка піддаються безпосередньому впливу інформаційних чинників (передумов, що чинять опір чи утруднюють формування і функціонування адекватної інформаційно-орієнтуючої основи суспільної поведінки людини (життєдіяльності у суспільстві)), які трансформуються, через його поведінку, діяльність (бездіяльність), здійснюють деструктивний, дисфункційний вплив на його життєдіяльність.

До основних загроз інформаційно-психологічної безпеки відносять можливість настання негативних наслідків для суб'єктів, що піддаються інформаційно-психологічному впливу, які виражаються в таких формах:

- нанесення шкоди здоров'ю людини;
- блокування на неусвідомленому рівні волі, волевиявлення людини, штучне привиття їй синдрому залежності;
- втрата здатності до політичної, культурної, моральної самоідентифікації людини;
- маніпуляція суспільною свідомістю;

– руйнування єдиного інформаційного і духовного простору України, традиційних устроїв суспільства і суспільної моральності, а також порушення інших життєво важливих інтересів особистості, суспільства, держави.

Наприклад, культ жорстокості, насильства, порнографії, розбещеності тощо, які пропагують у засобах масової інформації, друкованих виданнях, комп'ютерних іграх, мережі Інтернет веде до неусвідомленого бажання у підлітків і молоді, а також дорослих з нестійкою психікою, копіювати запропоновані моделі поведінки. Цей вид пропаганди знижує рівень порогових обмежень і правових заборон, що поряд з іншими умовами відкриває шлях для багатьох правопорушень. Це своєю чергою наносить непоправну шкоду не тільки окремій особистості, але й суттєві збитки національним інтересам країни.

Отже, джерелом інформаційно-психологічної небезпеки є та частина інформаційного середовища, яка через визначені причини неадекватно відображає реалії, вводить в оману людину, засліплює її ілюзією.

Інформаційно-психологічні загрози зумовлені розробкою, виготовленням, розповсюдженням та використанням суб'єктами негативних інформаційно-психологічних впливів, спеціальних засобів і методів такого впливу.

Концепція інформаційно-психологічної безпеки.

Сучасне розуміння безпеки в контексті врахування відношення інтересів особистості, суспільства і держави висуває завдання розгляду нового аспекту цієї проблеми – безпеки в інформаційній сфері життєдіяльності людини, тобто інформаційно-психологічної безпеки.

В інформаційному середовищі, що є складовим системним утворенням, виділяється процесуальна складова як найбільш динамічна і змінна її частина – інформаційно-комунікативні процеси, які активно впливають на індивідуальну, групову і суспільну психологію (індивідуальну, групову, масову свідомість). Маніпулюючи станом інформаційного середовища, змінюється стан духовної

сфери суспільства, деформація і деструктивні зміни якої у формі психоемоційної і соціальної напруженості, спотворених норм і неадекватних соціальних стереотипів і установок, оманливих і неприродних орієнтацій та цінностей. Це своєю чергою впливає на стан і процеси у всіх основних сферах суспільного життя, в тому числі політичній і економічній.

Вперше у пострадянському просторі про проблему інформаційно-психологічної безпеки було зазначено в листопаді 1995 р. на науково-практичній конференції, організованій Інститутом психології Російської академії наук. На цій та подальших конференціях було розкрито роль знання технологій інформаційно-психологічного впливу, метою якого є маніпуляція, для вироблення напрямів реформування психологічного захисту особистості і особистої інформаційно-психологічної безпеки.

Інформаційно-психологічну безпеку особистості визначають такими основними причинами.

Зростання тиску інформаційного середовища визначає необхідність формування нових механізмів та засобів виживання людини як особистості й активного соціального суб'єкта у сучасному суспільстві.

Взаємодія психіки людини з інформаційним середовищем відрізняється якісною специфікою і не має аналогів у комунікації інших біологічних, технічних, соціальних і соціотехнічних структур.

Основною і центральною "мішенню" інформаційного впливу є людина, її психіка.

Отже, інформаційно-психологічну безпеку можливо розглядати як стан захищеності особистості, різних соціальних груп і об'єднань людей від дій, впливів, які здатні проти їхньої волі і бажання змінити психічні стани та психологічні характеристики людини, модифікувати її поведінку і обмежувати свободу вибору, зумовило потребу переосмислення інформаційної взаємодії, а також деяких інших соціально-психологічних процесів і явищ у сучасному суспільстві.

Інформаційно-психологічна безпека – стан захищеності окремих осіб чи груп осіб від негативних інформаційно-психологічних впливів і пов'язаних з цим інших життєво важливих інтересів особистості, суспільства, держави в інформаційному середовищі.

Негативний інформаційно-психологічний вплив – процес зміни психічних станів і характеристик людей під впливом інформаційно-комунікативних процесів як динамічного компонента інформаційного середовища. Цей вплив спрямований на людину чи групу осіб (у тому числі без їхньої згоди) з метою примусу до визначеної поведінки, оцінки ситуації, керування та корекції індивідуальної та колективної свідомості. Він здійснюється з використанням спеціальних засобів і методів впливу на психіку людини, унаслідок чого він приводить до негативних наслідків для особистості, суспільства і держави.

Спеціальні засоби впливу – технічні і програмні засоби, що використовують для використання з метою негативного інформаційно-психологічного впливу на людину чи групу людей.

Спеціальні методи впливу – послідовність прийомів впливу на психіку людини, використання яких приводить до негативних наслідків для особистості, суспільства та держави.

Головним об'єктом забезпечення інформаційно-психологічної безпеки в інформаційному середовищі у сфері індивідуальної безпеки є усвідомлення інформації, здатність людини адекватно сприймати навколишню дійсність, своє місце в зовнішньому світі, формувати відповідно до свого життєвого досвіду визначені переконання і приймати стосовно них рішення.

Інформаційно-психологічна безпека має спиратися на стандарти інформаційно-психологічної безпеки – затверджені у визначеному порядку інформаційно-психологічного впливу, який не викликає негативних наслідків для психіки людини.

ВИСНОВКИ

Надійний ринок послуг та продуктів на основі прогнозів погоди вимагає своєчасних, недорогих та високоякісних даних про погоду. Основним постачальником даних для кожної країни є НМС, яка несе остаточну відповідальність за надання даних про погоду. Дані про погоду зазвичай класифікуються як «кліматичні» або «синоптичні». Кілька країн, наприклад Австралія, не розрізняють кліматичні та синоптичні дані. Однак у більшості країн існує суттєва різниця між цими двома наборами даних. Кліматичні дані є офіційними даними, наданими НМС. Синоптичні дані – це дані, які надаються в регулярні звіти для введення в глобальні погодні моделі.

Дані про погоду, надані НМС, часто містять пропущені значення та помилки, і вони не коригуються на зміни розташування станцій протягом періоду запису. Приватна метеорологічна галузь відіграє важливу роль на ринку погодніх ризиків, надаючи методології очищення та покращення даних про погоду. Очищення даних про погоду виконується, щоб видалити відсутні значення та замінити помилкові значення. Покращення даних про погоду виконується, щоб налаштувати часові ряди даних про погоду для зміни розташування станції.

Існує багато методів очищення та покращення даних. Очищення даних досягається шляхом виконання ряду перевірок контролю якості даних для виявлення відсутніх значень і «позначення» підозрілих значень. Ці підозрілі або відсутні значення можна потім замінити за допомогою методів просторової або часової інтерполяції. Для очищення історичних даних можна використовувати будь-який метод, але для очищення даних про погоду в реальному часі перевага віддається методам просторової інтерполяції, оскільки вони не потребують даних до і після дати заміни для інтерполяції. Існує кілька методів інтерполяції просторових даних. Модель, яка має вирішити проблему оцінки нестандартного

інтервалу точок, буде використовувати як відстань, так і кореляцію від точки до точки для отримання ваг інтерпольованих значень.

Розширення даних виконується, щоб визначити, коли виникли розриви, і визначити величину розривів. Метадані дуже важливі для оцінки потенційних дат розривів. Для покращення часового ряду можна використовувати методології однієї станції або методології кількох станцій. Часто перевага віддається методології кількох станцій, що використовують еталонний часовий ряд, оскільки вони усувають вплив регіональних кліматичних тенденцій. Перевірку будь-якої методології слід проводити за допомогою парних тестів та інтерв'ю з оператором.

Нова технологічна галузь, яка ще не буде повністю вивчена індустрією погодних даних, — це технології дистанційного зондування. Сучасні радіолокаційні та супутникові технології можуть діагностувати температуру поверхні за допомогою інфрачервоних давачів і кількість опадів за допомогою мікрохвильових датчиків після належного калібрування. Це може бути рішенням для областей без приладів або там, де є проблеми з надійністю даних або проблеми з моральним ризиком.

ПЕРЕЛІК ПОСИЛАНЬ

1. World Meteorological Organization's website. URL: <https://www.wmo.int/index-en.html> (травень, 2022).
2. UK Met Office website. URL: <https://www.metoffice.gov.uk/> (травень, 2022).
3. The NCDC's website. URL: <https://www.ncei.noaa.gov/> (травень, 2022).
4. The NOAA's website. URL: <http://www.noaa.gov> (травень, 2022).
5. The NWS's website. URL: <https://www.weather.gov/> (травень, 2022).
6. Deutscher Wetterdienst website. URL: <http://www.dwd.de> (травень, 2022).
7. Japan Meteorological Agency's website. URL: <http://www.kishou.go.jp/english/index.html> (травень, 2022).
8. Bureau of Meteorology website. URL: <http://www.bom.gov.au> (травень, 2022).
9. Cooperative Observer Program website. URL: <https://www.weather.gov/coop/> (травень, 2022).
10. Easterling, D. R., & Peterson, T. C. (1995). A new method for detecting undocumented discontinuities in climatological time series. *International journal of climatology*, 15(4), 369-377.
11. The WRMA website. URL: <http://www.wrma.org> (травень, 2022).
12. Information on Global Telecommunication Systems website. URL: <https://community.wmo.int/activity-areas/global-telecommunication-system-gts> (травень, 2022).
13. Jewson, S., and D. Whitehead, 2001, "In Prase of Climate Data", *Environmental Finance*, 3:2, pp. 22–24.
14. KNMI website. URL: <http://www.knmi.nl> (травень, 2022).

15. Finnish Meteorological Institute website. URL: <https://en.ilmatieteentilaitos.fi/> (травень, 2022).
16. The ECOMET website. URL: <http://www.meteo.be/ECOMET> (травень, 2022)
17. Cressman, G. P. (1959). An operational objective analysis system. *Monthly Weather Review*, 87(10), 367-374.
18. Barnes, S. L. (1964). A technique for maximizing details in numerical weather map analysis. *Journal of Applied Meteorology and Climatology*, 3(4), 396-409.
19. Burrough, P., & McDonnell, R. (1998). Spatial information systems and geostatistics. P. Burrough, & R. McDonnell, *Principles of Geographical Information Systems*, 333.
20. Snell, S. E., Gopal, S., & Kaufmann, R. K. (2000). Spatial interpolation of surface air temperatures using artificial neural networks: Evaluating their use for downscaling GCMs. *Journal of Climate*, 13(5), 886-895.
21. Gullett, D., Skinner, W., & Vincent, L. (1992). Development of an historical Canadian climate database for temperature and other climate elements.
22. Mekis, E., & Hogg, W. D. (1999). Rehabilitation and analysis of Canadian daily precipitation time series. *Atmosphere-ocean*, 37(1), 53-85.
23. Acock, M. C., & Pachepsky, Y. A. (2000). Estimating missing weather data for agricultural simulations using group method of data handling. *Journal of Applied meteorology*, 39(7), 1176-1184.
24. Schrupf, A. D. (1997). Temperature Data Continuity with the Automated Surface Observing System. AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH.
25. Boissonnade, A. C., Heitkemper, L. J., & Whitehead, D. (2002). Weather data: cleaning and enhancement. *Climate Risk and the Weather Market*, 73-98.

26. McKee, T. B., Doesken, N. J., Davey, C. A., & Pielke Sr, R. A. (2000). Climate data continuity with ASOS: Report for period April 1996 through June 2000 (Doctoral dissertation, Colorado State University. Libraries).
27. Peterson, T. C., Easterling, D. R., Karl, T. R., Groisman, P., Nicholls, N., Plummer, N., ... & Parker, D. (1998). Homogeneity adjustments of in situ atmospheric climate data: a review. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 18(13), 1493-1517.
28. Peterson, T. C., Easterling, D. R., Karl, T. R., Groisman, P., Nicholls, N., Plummer, N., ... & Parker, D. (1998). Homogeneity adjustments of in situ atmospheric climate data: a review. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 18(13), 1493-1517.
29. Mekis, E., & Hogg, W. D. (1999). Rehabilitation and analysis of Canadian daily precipitation time series. *Atmosphere-ocean*, 37(1), 53-85.
30. Alexandersson, H., & Moberg, A. (1997). Homogenization of Swedish temperature data. Part I: Homogeneity test for linear trends. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 17(1), 25-34.
31. Karl, T. R., & Williams Jr, C. N. (1987). An approach to adjusting climatological time series for discontinuous inhomogeneities. *Journal of Applied Meteorology and Climatology*, 26(12), 1744-1763.
32. Allen, R. J., & DeGaetano, A. T. (2000). A method to adjust long-term temperature extreme series for nonclimatic inhomogeneities. *Journal of climate*, 13(20), 3680-3695.
33. Caussinus, H., & Mestre, O. (1996, October). New mathematical tools and methodologies for relative homogeneity testing. In *Proceedings of the Seminar for Homogenization of Surface Climatological Data* (pp. 63-82).
34. Venema, V. K., Mestre, O., ... & Brandsma, T. (2012). Benchmarking homogenization algorithms for monthly data. *Climate of the Past*, 8(1), 89-115.
35. Жидецький, В. Ц., Джигирей, В. С., & Мельников, О. В. (2000). Основи охорони праці. Львів: Афіша, 350, 132-136.