УДК 551.582.2, 551.583.1

ГОМОГЕНІЗОВАНА БАЗА ДАНИХ ДОВГИХ РЯДІВ СЕРЕДНЬОЇ МІСЯЧНОЇ ТЕМПЕРАТУРИ ПОВІТРЯ

В.І. Осадчий[1], О.А. Скриник[2], В.П. Сіденко[1, 3], Д.О. Бойчук[1, 3], Д.О. Ошурок[1], О.Я. Скриник[1]

[1]Український гідрометеорологічний інститут ДСНС та НАН України, просп. Науки 37, м. Київ 03028, Україна, e-mail: osad@uhmi.org.ua, dimaoshurok@gmail.com, skrynyk@uhmi.org.ua

[2]Національний університет біоресурсів та природокористування України, вул. Героїв Оборони 15, м. Київ 03041, Україна, e-mail: skrynyk@nubip.edu.ua,

[3]Київський національний університет імені Тараса Шевченка, вул. Володимирська, 64/13, м. Київ 01601, Україна, e-mail: vladyslavsidenko@gmail.com

Створено цифрову базу даних середніх місячних значень температури повітря, отриманих на 514 метеорологічних станціях, які працювали на території України в той чи інший період часу з 1812 по 2014 р. Для всіх станцій зібрано метадані (історичну інформацію про дати перенесень станцій чи інших змін у проведенні вимірювань). Показано, що надійні результати щодо гомогенізації часових рядів можна отримати тільки для періоду 1946–2014 рр. унаслідок великої кількісті пропусків протягом Першої та Другої світових воєн. Для зазначеного періоду гомогенізовано ряди даних 178 станцій за допомогою спеціалізованого/гомогенізаційного програмного забезпечення HOMER. Виділено точки розриву та визначено величини зсувів "станційних сигналів". Розраховано статистичні характеристики виявлених неоднорідностей. За порівняно великою кількістю метаданих та ідентифікованих розривів і їх розрахованими статистиками встановлено, що неоднорідність первинної інформації слід обов'язково враховувати під час проведення кліматологічних досліджень, зокрема і під час оцінювання регіональних кліматичних змін.

Ключові слова: середня місячна температура повітря, довгі часові ряди, кліматологічна неоднорідність, гомогенізація, HOMER, точки розриву, "станційний сигнал".

Вступ. Довгі ряди кліматологічних показників зазвичай містять різкі зсуви східчастого характеру та/або плавні тренди, які не є наслідком змін клімату чи його мінливості [12]. Причина їх появи — зміни у способі отримання емпіричної інформації на метеорологічних станціях (перенесення станцій, зміна їх оточення, заміна приладів чи зміна методики вимірювання/розрахунку) [35]. Такі зсуви/тренди інколи називають "станційним сигналом" або кліматологічною неоднорідністю [7]. Їх виявлення та вилучення, тобто приведення рядів до однорідного стану чи, іншими словами, гомогенізація, є важким завданням, оскільки в рядах кліматологічних даних завжди наявна інтенсивна шумова компонента, яка добре маскує будь-які регулярні зміни. За наявності детальної інформації про історії станцій та виконання паралельних вимірювань у "старих" і "нових" умовах можна було б точно оцінювати неоднорідності. Проте зазвичай історична інформація є доволі обмеженою та неповною, а паралельні вимірювання взагалі відсутні. Тому для гомогенізації здебільшого застосовують статистичні методи.

Очевидно, що першим етапом проведення кліматологічних досліджень має бути перевірка на однорідність/неоднорідність емпіричних даних (або обґрунтування однорідності даних). Така перевірка є особливо важливою, в разі досліджування чи оцінювання змін клімату, оскільки амплітуда "станційних сигналів" часто рівновелика величині самих змін [24]. Тому використання первинних (не гомогенізованих) даних може призвести до неправильного висновку щодо динаміки клімату.

За останні десятиріччя кліматологи згаданій проблемі приділяли значну увагу з метою її розв'язання [29]. Було створено велику кількість сучасних та ефективних гомогенізаційних алгоритмів, процедур і програмних пакетів [13, 15, 16, 19, 34, 37]. У результаті у багатьох країнах гомогенізовано національні бази даних (БД) довгих рядів основних кліматологічних показників [21, 32, 33]. Гомогенізаційні методи широко застосовують і в міжнародних кліматичних проектах [14, 31] для створення транскордонних БД і досліджень динаміки клімату.

Мета роботи — створення БД (максимальної за часовим періодом та кількістю станцій) довгих гомогенізованих рядів середніх місячних значень температури повітря, отриманих на території України. У подальшому БД можна використовувати для регіональних кліматологічних досліджень, зокрема, і для коректного оцінювання кліматичних змін температурного режиму.

Зауважимо, що незважаючи на довгу та багату на результати історію кліматологічних досліджень температурного режиму в Україні [1—4, 8, 10, 22], проблемі було приділено, на диво, мало уваги. Ймовірно, про перші спроби гомогенізувати ряди середніх місячних температур зазначено у публікаціях [27, 30]. Проте ці роботи не можна вважати остаточними через малу кількість станцій (33) чи малий часовий проміжок (1961—2010). У цій статті (як і у статті [28]) зібрано всі доступні дані та використано характеристики максимальної кількості станцій.

Емпіричні дані (середня місячна температура повітря). Середню місячну температуру повітря отримано із трьох джерел. Першу частину даних (від початку роботи станцій до 1950 р.) оцифровано з кліматологічного довідника [9]. Слід зазначити, що в ньому для деяких станцій наведено ряди спостережень. Проте ми оцифрували всі дані. Вважаємо, що інформація зі станцій з короткими періодами спостережень може бути використана для заповнення пропусків, що особливо актуально для кінця XIX і початку XX ст. Тоді працюючих станцій було дуже мало, тому кожен запис температури повітря є важливим.

Дані для періоду 1951–1965 рр. оцифровано з довідника [11]. Для періоду 1966–2014 рр. місячні значення температури розраховано за середніми добовими значеннями, які отримано в Центральній геофізичній обсерваторії. Для розрахунків використано рекомендації Всесвітньої метеорологічної організації [38].

Загальна кількість станцій, дані яких було опрацьовано, — 514. Імовірно, це всі (або майже всі) станції, які працювали в той чи інший час на території України починаючи з 1812 по 2014 р. На рис. 1 показано кількість працюючих станцій для кожного місяця періоду. Як видно з рисунка, регулярні спостереження на досить великій кількості станцій (~50) було розпочато на початку 80-х років XIX ст. Проте є два періоди часу, а саме Першої та Другої світових воєн, коли кількість працюючих станцій суттєво знижувалась. Для прикладу, на рис. 2,*а* зображено ситуацію для листопада 1941 р. Тільки 7 станцій (4 з яких знаходились у Луганській обл.) працювали того місяця. Це мінімальна кількість станцій для періоду 1881–2014 рр. Для порівняння, на рис. 2,*б* показано також максимальну кількість працюючих метеорологічних станцій для цього самого періоду.

Зазначені вище два проміжки часу з великою кількістю пропущених даних суттєво знижують надійність результатів гомогенізації: зазвичай гомогенізаційні алгоритми та програмні засоби є дуже чутливими до пропусків. Тому ми обмежили часовий проміжок для аналізу періодом 1946–2014 рр., де пропущених даних значно менше. Вважаємо, що відновлення температурних даних для попереднього періоду, а також їх гомогенізація мають бути предметом окремого дослідження, причому мають бути використані спеціальні методи просторово-часової інтерполяції (просторово-часового моделювання) та всі доступні дані (із залученням даних прикордонних районів суміжних країн).

Крім того, із набору станцій для періоду 1946–2014 рр. було виключено ті, які нині є закритими, та ті, що почали працювати після 1963 р. Остання умова узгоджується з умовою наявності 70 % даних у ряді як мінімального порогу для можливості включення ряду в подальший аналіз (гомогенізацію) [23]. У підсумку для гомогенізації (створення гомогенізованої БД) було використано 178 станцій, які рівномірно розташовані на території України. Їх просторовий розподіл показано на рис. 3.

Метадані. Історичну інформацію про дати змін у способі отримання емпіричних даних (метадані) отримано із двох кліматологічних видань [6, 5]. Довідник [6] містить історії станцій від початку

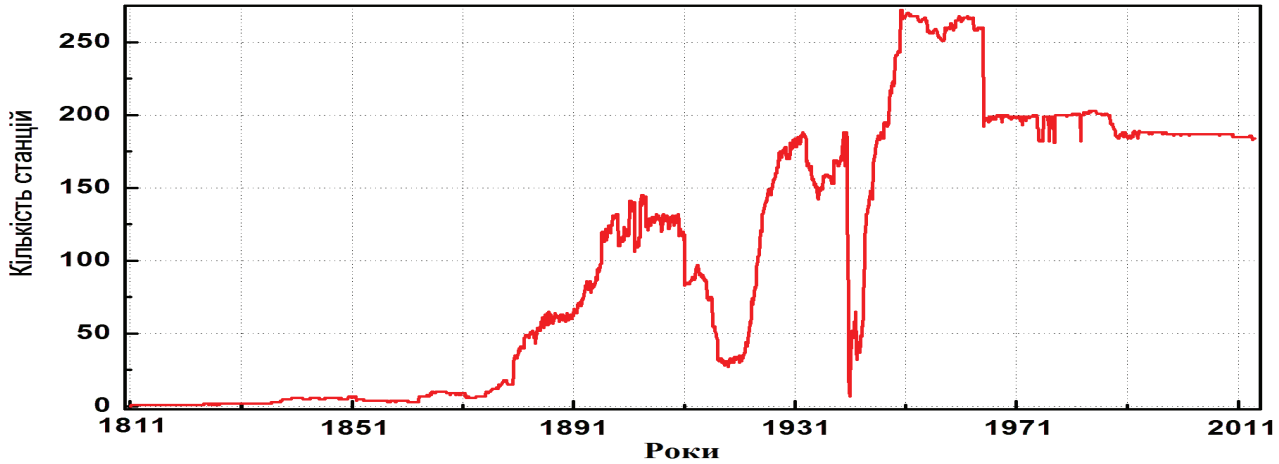
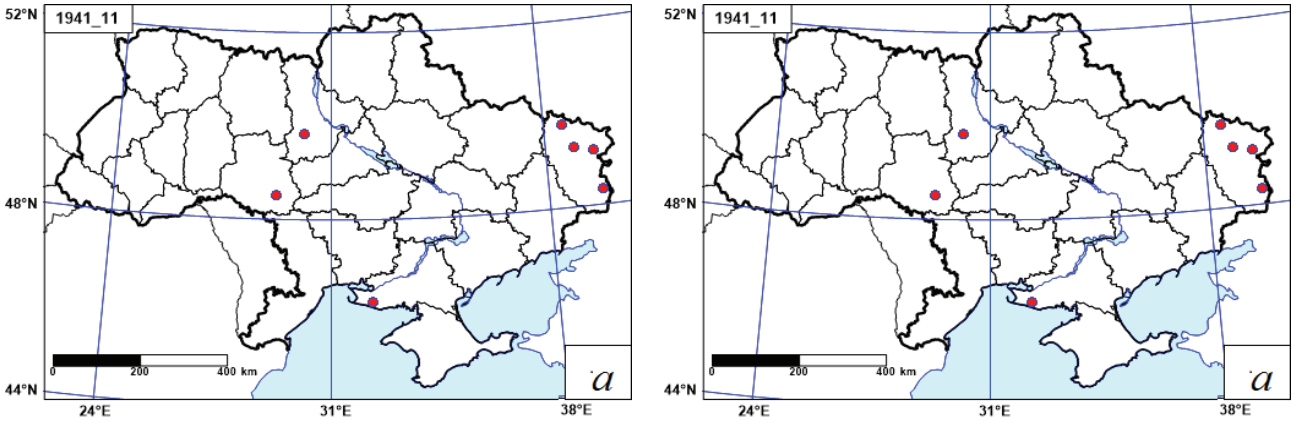


Рис. 1. Еволюція загальної кількості метеорологічних станцій, які працювали на території України протягом періоду 1812–2014 рр.

Fig. 1. Evolution of the total number of meteorological stations operating in Ukraine during the period of 1812–2014



а *б*

Рис. 2. Кількість метеорологічних станцій, що працювали на території України (за період 1881–2014 рр.): *а* – мінімальна (7, листопад 1941 р.); *б* – максимальна (272, квітень 1951 р.)

Fig. 2. The number of meteorological stations working on the territory of Ukraine (since 1881): *а* – minimum number (7, November 1941); *б* – maximum (272, April 1951)

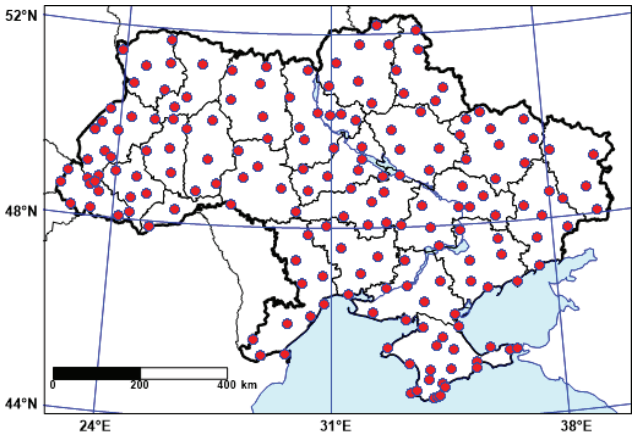


Рис. 3. Просторовий розподіл метеорологічних станцій, дані яких використано для гомогенізації (для створення гомогенізованої БД)

Fig. 3. Spatial distribution of meteorological stations, which data was used for the homogenization (to create a homogenized DB)

їх створення до 1967 р., сучасніша історія станцій представлена у довіднику [5]. Метадані було зібрано для всіх 514 станцій, для всього періоду їх роботи. У табл. наведено фрагмент зібраних та оцифрованих метаданих (для Волинської обл.). Зроблено спробу зібрати тільки надійні метадані, коли дата та причина можливої точки розриву чітко вказані. Тому отримана колекція метаданих складається в основному із дат переміщень станцій.

На рис. 4 представлено статистичну інформацію про метадані для 178 станцій, які було вибрано для проведення гомогенізації. Як видно з рис. 4,*а*, для 178 станцій існує велика кількість метаданих. Від початку роботи станцій в їх історичних описах було знайдено 420 чітко зафіксованих подій (в основному переміщень місць спостережень), які потенційно могли спричинити появу неоднорідностей (“станційних сигналів”). У середньому — це

Метадані для метеорологічних станцій України (фрагмент, Волинська обл.)
Metadata for meteorological stations in Ukraine (a fragment, Volyn Region/Oblast)

№ пп.	Станція	BMO ID	Рік початку роботи	Можливі точки розриву	Переміщення станції
1	Любешів*	33075	1950	–	–
2	Світязь	33067	1948	–	–
3	Маневичі	11143	1946	30.08.1963	210 м на південний захід
				01.1982	300 м на північний захід
				04.12.1999	600 м на південний захід
				11.11.2008	50 м на захід
4	Ковель	33173	1940	31.03.1945	1500 м на південний схід
				14.09.1970	1300 м на південний схід
5	Володимир-Волинський	33177	1929	28.11.1963	2000 м на південний схід, на околицю міста
				20.10.1989	100 м на південь
6	Луцьк (Луцьк, АМСЦ)	33187	1892	03.06.1945	500 м на північ, в аеропорт
				28.09.1951	1800 м на схід
				08.1961	500 м на північний захід
				11.04.1984	14000 м на південний схід, в аеропорт
				25.11.1998	8000 м на північний схід, у приміську зону

* Станція продовжує працювати.

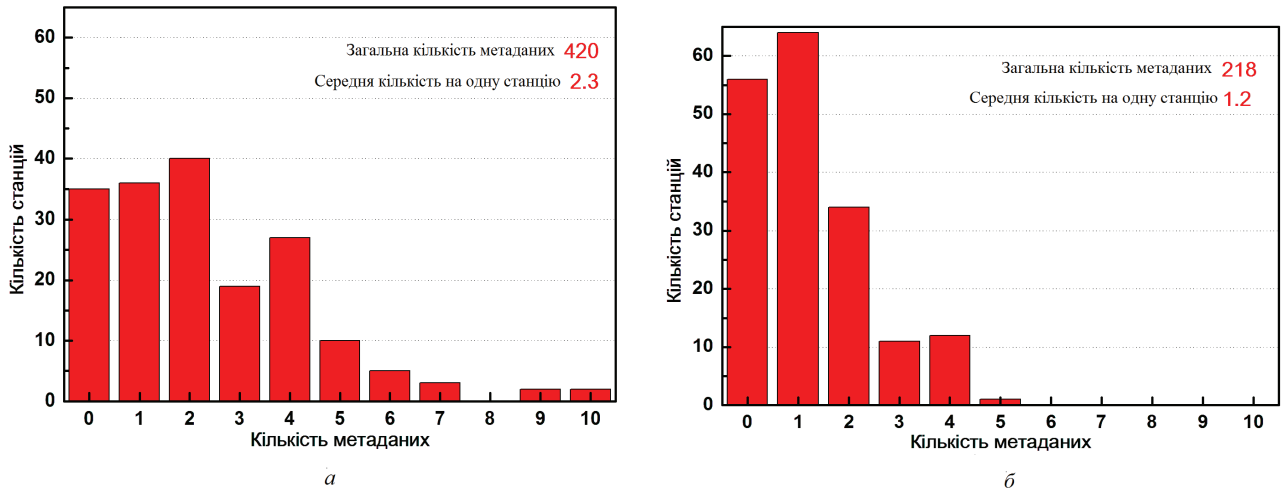


Рис. 4. Розподіл кількості станцій за кількістю метаданих (в основному — переміщень станцій: а — від початку роботи станцій; б — для періоду 1946—2014 рр.

Fig. 4. Distribution of the number of stations with the respect to the number reported changes in observing practice (mainly, relocation of the stations): a — since the beginning of the stations' history; b — for the period of 1946—2014

більш як 2 події на одну станцію. Тільки 35 станцій (~20 %) не зазнали змін у способі отримання емпіричної інформації (принаймні — вони не були зафіксовані у довідниках). "Рекордсменами" є станції Дніпро та Чернігів, кожну із яких переносили 10 разів.

Для періоду 1946—2010 рр. (рис. 4, *б*) загальна кількість метаданих зменшилася приблизно вдвічі (до 218, у середньому — понад 1 подію на станцію). Також збільшилася до 56 (~31 %) кількість станцій, які не зазнали змін у практиці проведення вимірювань. Проте кількість метаданих залишається значною.

Очевидно, що така велика кількість потенційних розривів/зсувів впливає на якість емпіричних даних/часових рядів, а також на надійність результатів їх аналізу. Тому їх обов'язково слід враховувати під час проведення кліматологічних досліджень.

За часовим розподілом метаданих (рис. 5) потенційно шкідливих впливи (~70%) на практику метеорологічних вимірювань/спостережень відбувалися в першій половині періоду 1946—2014 рр. (до 1980 р.).

Метод. Для проведення гомогенізації використано сучасне програмне забезпечення HOMER [26]. HOMER — спільний Європейський продукт, який створено в рамках реалізації проекту COST Action ES0601: (HOME) [20]. Програмне забезпечення об'єднує декілька добре відомих гомогенізаційних методів, було перевірено та верифіковано на незалежних реалістичних тестових даних [36]. Слід зазначити, що з часу створення HOMER активно використовують для гомогенізації національних БД кліматологічних показників [17, 18, 23].

Програмне забезпечення HOMER належить до відносних гомогенізаційних методів [36]. Тому всі

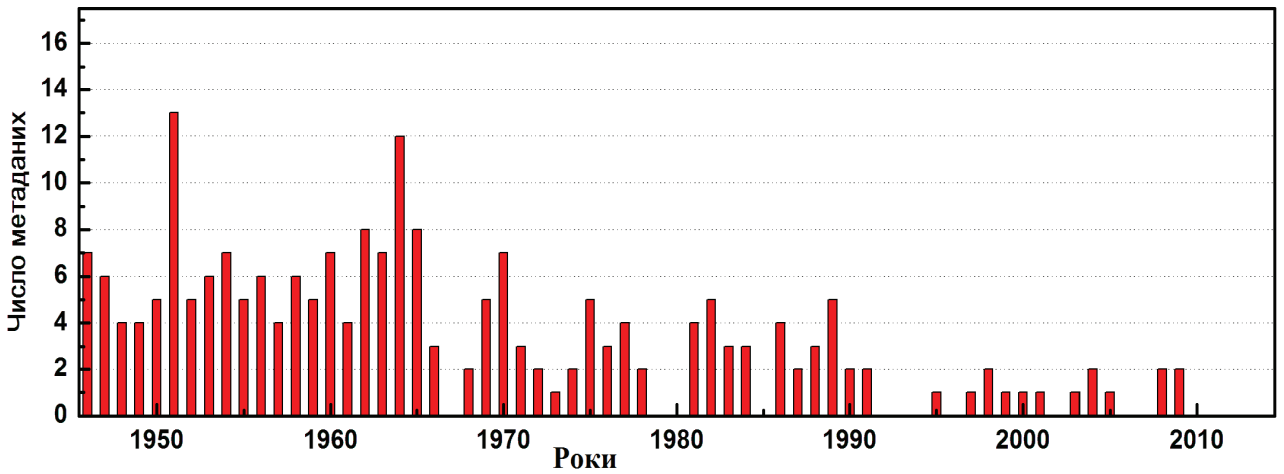


Рис. 5. Часовий розподіл метаданих

Fig. 5. Time distribution of metadata

статистичні розрахунки в цій програмі виконуються не для окремих рядів, а для певної кількості або різницевих рядів (для адитивних величин, таких як температура повітря), або рядів відношень (для кумулятивних величин, таких як сума опадів). Іншими словами, для довільного ряду, який необхідно гомогенізувати (candidate series), за певною ознакою (близькість географічного розміщення або висока кореляція даних) завжди вибирають деяку кількість інших довідкових/контрольних рядів (reference series), розраховують їх різниці/відношення, які в подальшому й аналізують. Детально математичну основу програми HOMER описано у документі [26].

У роботі використано інтерактивний гомогенізаційний алгоритм, який детально описано та обґрунтовано у документації програмного забезпечення [25].

Контроль якості даних. Першим етапом проведення гомогенізації є виявлення та вилучення викидів (outliers). В англомовній літературі цей процес називають контролем якості (quality control — QC). Викиди необхідно вилучити перед проведенням гомогенізації, оскільки вони суттєво впливають на її результати.

Для проведення контролю якості програма HOMER розраховує та подає у графічному вигляді аномалії досліджуваного ряду та вибраних довідкових рядів щодо середнього значення, визначеного за даними цих довідкових рядів (тобто щодо просторово-локального середнього). Ряди аномалій програма розраховує окремо для кожного місяця року. Значення досліджуваного ряду вважають викидом, якщо відповідна аномалія перевищує ±1,5 °C. Якщо ідентифікований викид не може бути виправлений на коректне/правильне значення, його замінюють на спеціальну мітку пропуску (–999,9).

Ефективність виявлення викидів програмним забезпеченням можна продемонструвати на прикладі станції Луцьк. На рис. 6 показано первинні (не гомогенізовані) ряди даних цієї станції для березня та грудня. Як видно із рисунка, жодне із значень, зокрема і значення для 1970 р., не викликає підозри і не подібна до викиду. Проте порівняння аномалій цих рядів із аналогічними аномаліями довідкових рядів (рис. 7) дає змогу чітко ідентифікувати значення для 1970 р. як помилки/викиди.

У результаті звернення до першоджерел, а також залучення додаткової інформації (щомісячників) виявлено, що в значеннях 5,0 та □4,0 °C було випадково/помилково змінено десятковий знак і насправді середня місячна температура в ці місяці дорівнювала 0,5 та □0,4 °C. На рис. 8 і 9 показано подібні до рис. 6 і 7 зображення після виправлення даних.

Парне порівняння (pairwise detection) Після вилучення викидів застосовують парне порівняння для виявлення можливих точок розриву. Програма HOMER розраховує різниці між значеннями досліджуваного ряду та довідкових рядів. Різниці розраховують для річних, сезонних та місячних значень і подають для аналізу в графічній формі. До кожного із різницевих рядів застосовують статистичні розрахунки з метою виявлення оптимальної кількості точок розриву та їхньої локалізації на часової осі.

Знайдені розриви зображають у вигляді вертикальних суцільних ліній чорного кольору. Розрив вважають достовірним/обґрунтованим,

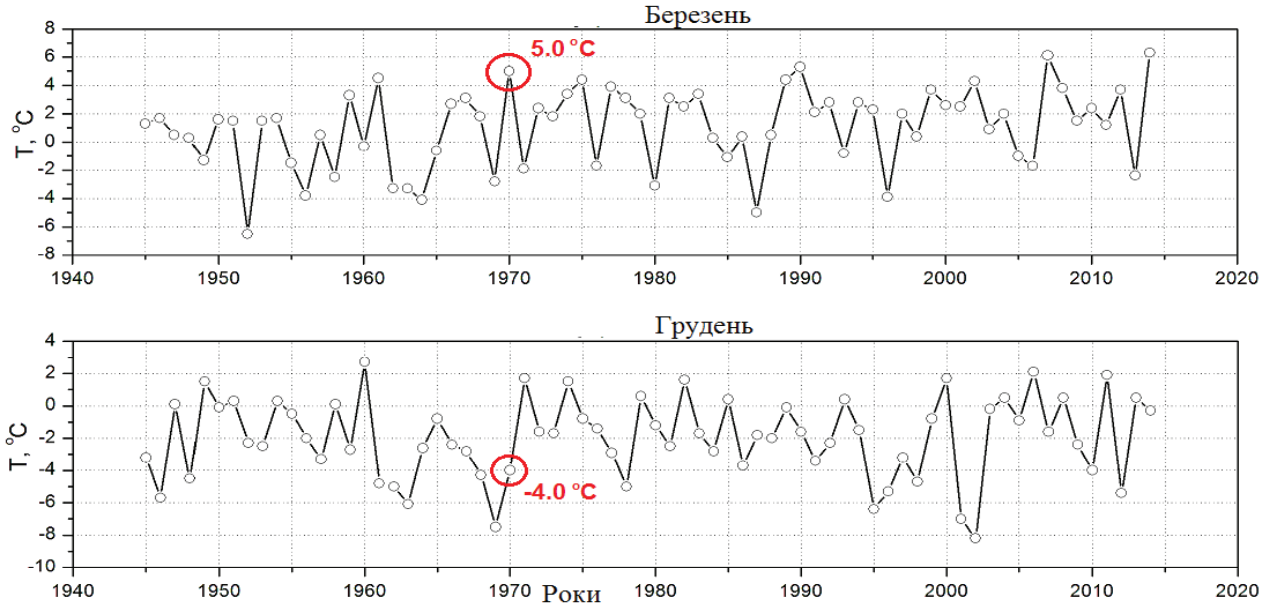


Рис. 6. Первинні (не гомогенізовані) ряди даних на станції Луцьк для березня та грудня

Fig. 6. Raw (non-homogenized) data series at Lutsk station for March and December

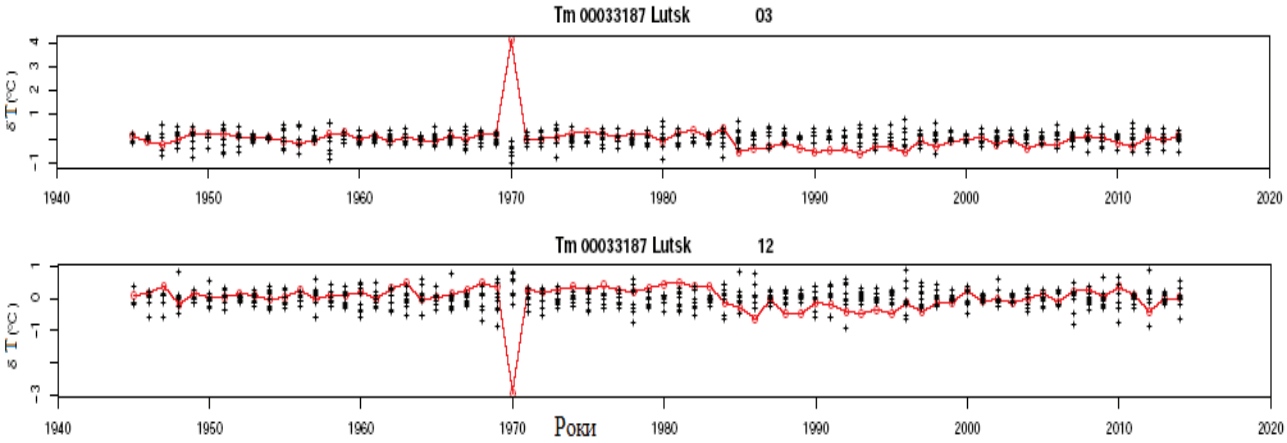


Рис. 7. Результати контролю якості рядів даних станції Луцьк для березня (03) та грудня (12) у програмі HOMER. По осі ординат відкладено аномалії рядів щодо локального середнього, розрахованого за всіма вибраними довідковими рядами. Суцільна лінія — аномалії для станції Луцьк, точки — аномалії довідкових рядів

Fig. 7. The QC results of Lutsk data series for March (03) and December (12) in the HOMER software. On the ordinate axis the anomalies of the series relative to the local average (calculated based on all selected reference series) are depicted. The continuous line shows anomalies for the station Lutsk, the dots mean anomalies of reference series

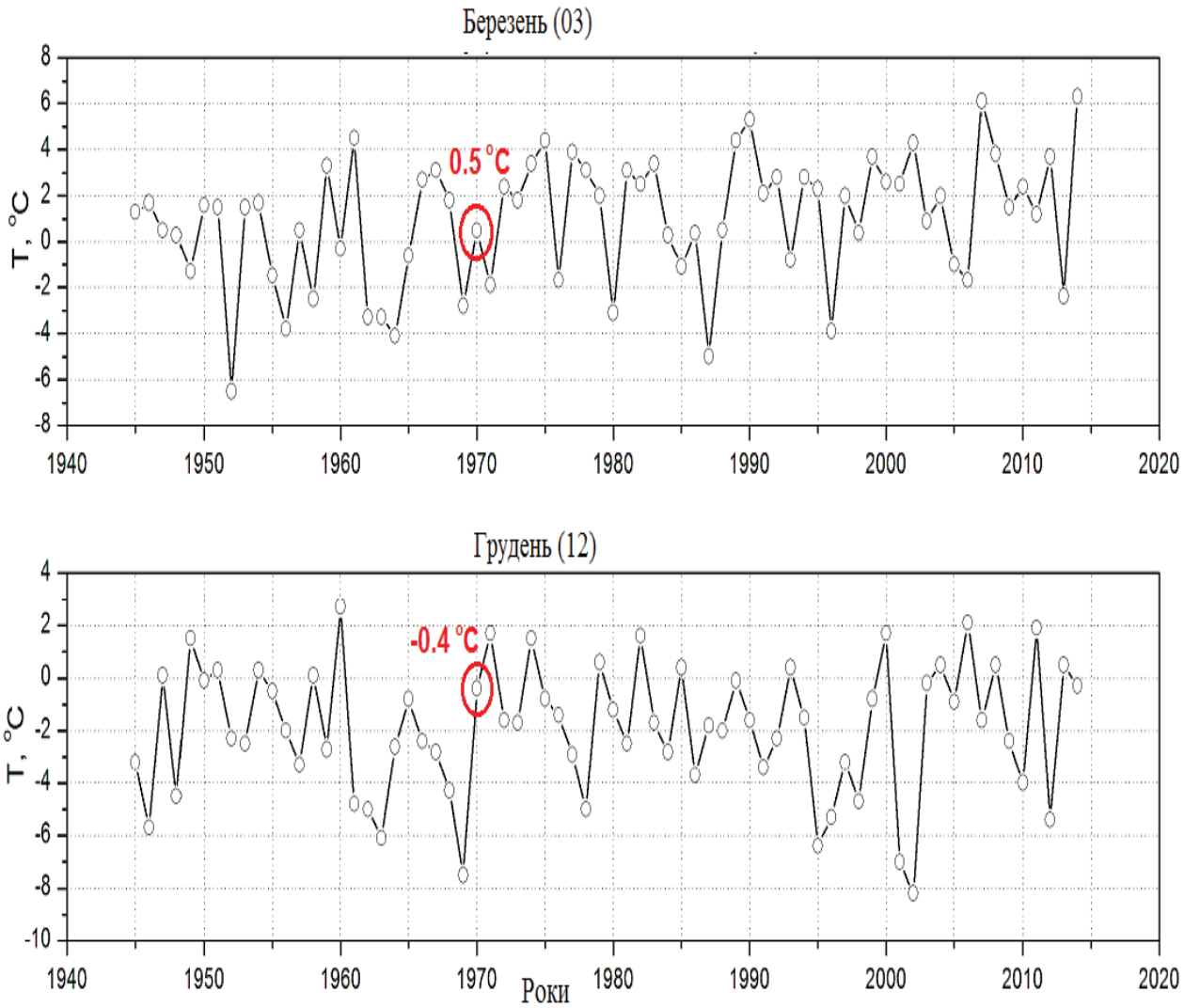


Рис. 8. Виправлені ряди даних на станції Луцьк для березня (03) та грудня (12)

Fig. 8. Corrected data series at Lutsk station for March and December

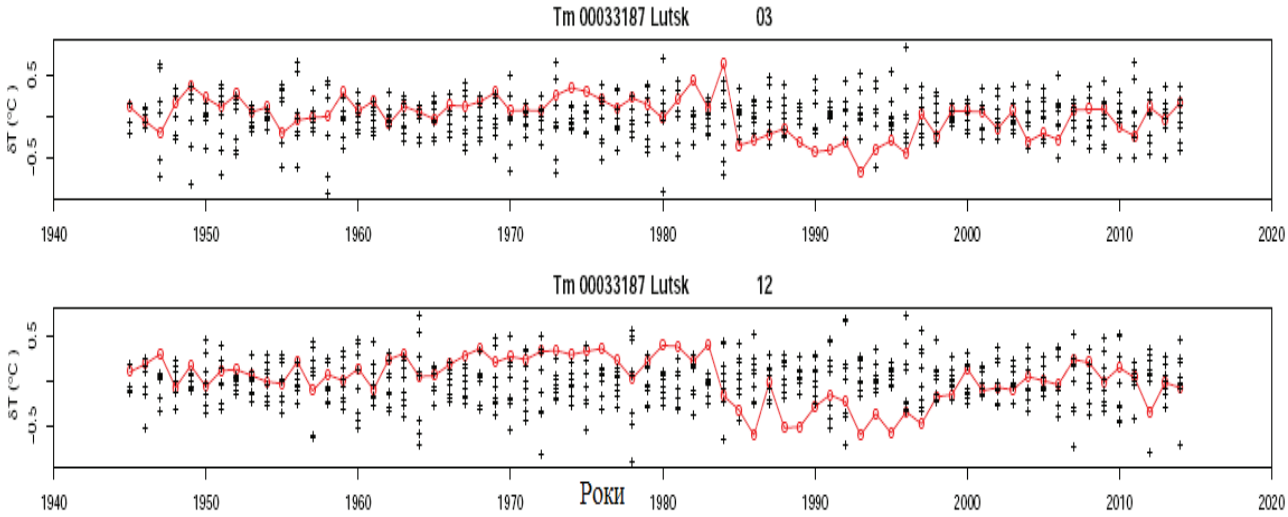


Рис. 9. Результати контролю якості рядів виправлених даних на станції Луцьк для березня (03) та грудня (12) у програмі HOMER

Fig. 9. The QC results of corrected data at Lutsk station for March (03) and December (12) in the HOMER software

якщо його позиція збігається у більшості різницевих рядів (як мінімум у трьох рядах). Для полегшення аналізу різницеві ряди розміщують у міру зростання їх середнього квадратичного відхилення (□).

На рис. 10 подано результати розрахунків для річних рядів станції Дніпро. Як видно з рисунка, для цієї станції дві точки розриву (у 1955 та 1970 роках) можуть бути чітко ідентифікованими (особливо, перша з них). Ще два розриви (близько 1960 та 1990 р.) є менш чіткими, тому їх необхідно додатково перевірити на наступних етапах гомогенізації. Зазначимо, що під час парного порівняння для підтвердження/спростування реальності існування деяких сумнівних розривів може бути залучено метадані.

Об'єднане порівняння (joint detection) Виявлені та занотовані можливі точки розриву мають бути підтвердженими під час наступного етапу: об'єднаного порівняння. Цей етап – інтерактивна процедура, протягом якої користувач має змогу коригувати результати розрахунків програми, спираючись на історичну інформацію/метадані. На рис. 11 показано приклад графічної інформації, яку отримує користувач під час об'єднаного порівняння (для станції Дніпро).

На рисунку зображено ідентифіковані на попередньому етапі розриви в усіх різницевих рядах для річних (∇), літніх (▲) і зимових (▲) даних. Крім того, представлено розриви (□), які розраховано принципово іншим гомогенізаційним алгоритмом (об'єднаною сегментацією). У документації HOMER рекомендують вважати розриви обґрунтованими/реальними, якщо навколо них (тобто – відповідних років) відбувається групування всіх типів ідентифікованих розривів, бажано із підтвердженням метаданими.

На рис. 11 таким є розрив у 1955 р. З історичного опису станції відомо, що в цьому році її було перенесено із середмістя в аеропорт. У межах 1987 р. теж групуються знайдені розриви, і тому цю точку розриву також можна вважати реальною. У 1970 р. групування не було таким значним, проте цю дату підтверджено метаданими. У той рік станцію знову було переміщено на достатньо велику відстань (1,3 км). Ідентифікований розрив у 1997 р. не підтверджено парним порівнянням і метаданими, тому його було анульовано користувачем (символ + біля □, який додає користувач в інтерактивному режимі, означає анулювання розриву, тоді як аналогічні символи в 1970 та 1987 рр., навпаки, означають додавання розривів).

Після опрацювання всіх станцій програма розраховує величини зсувів у ідентифікованих розривах, застосовуючи дисперсійний аналіз (ANOVA), та здійснює корекцію рядів (вилучає ідентифіковані "станційні сигнали" із вхідних даних).

Зазначені етапи за потреби можуть бути повторені. Рекомендують проводити не більш як дві ітерації. Інколи застосовують також третій статистичний алгоритм (ACMANT) для пошуку/ідентифікації розривів.

На завершальному етапі гомогенізації програма уточнює місяці, у яких відбулися ідентифіковані розриви, та остаточно корегує ряди даних.

Результати та їх обговорення. Результати контролю якості. Оскільки великий масив даних оцифровували напівавтоматичним способом (сканування з подальшим ручним коригуванням), було можливим отримання помилкових значень. Тому процедуру контролю якості проводили двічі. Першого разу виявлені викиди порівнювали із оригінальними записами в довідниках. Якщо інші першоджерела (щомісячники, добові значення та ін.) були доступними, то їх теж залучали для перевірки сумнівних значень. Якщо існувало підтвердження, що виявлені викиди є справді помилками, їх виправляли, у протилежному випадку – залишали в рядах. Таким чином, на виконання першому етапі контролю якості було виявлено та виправлено 118 помилко-

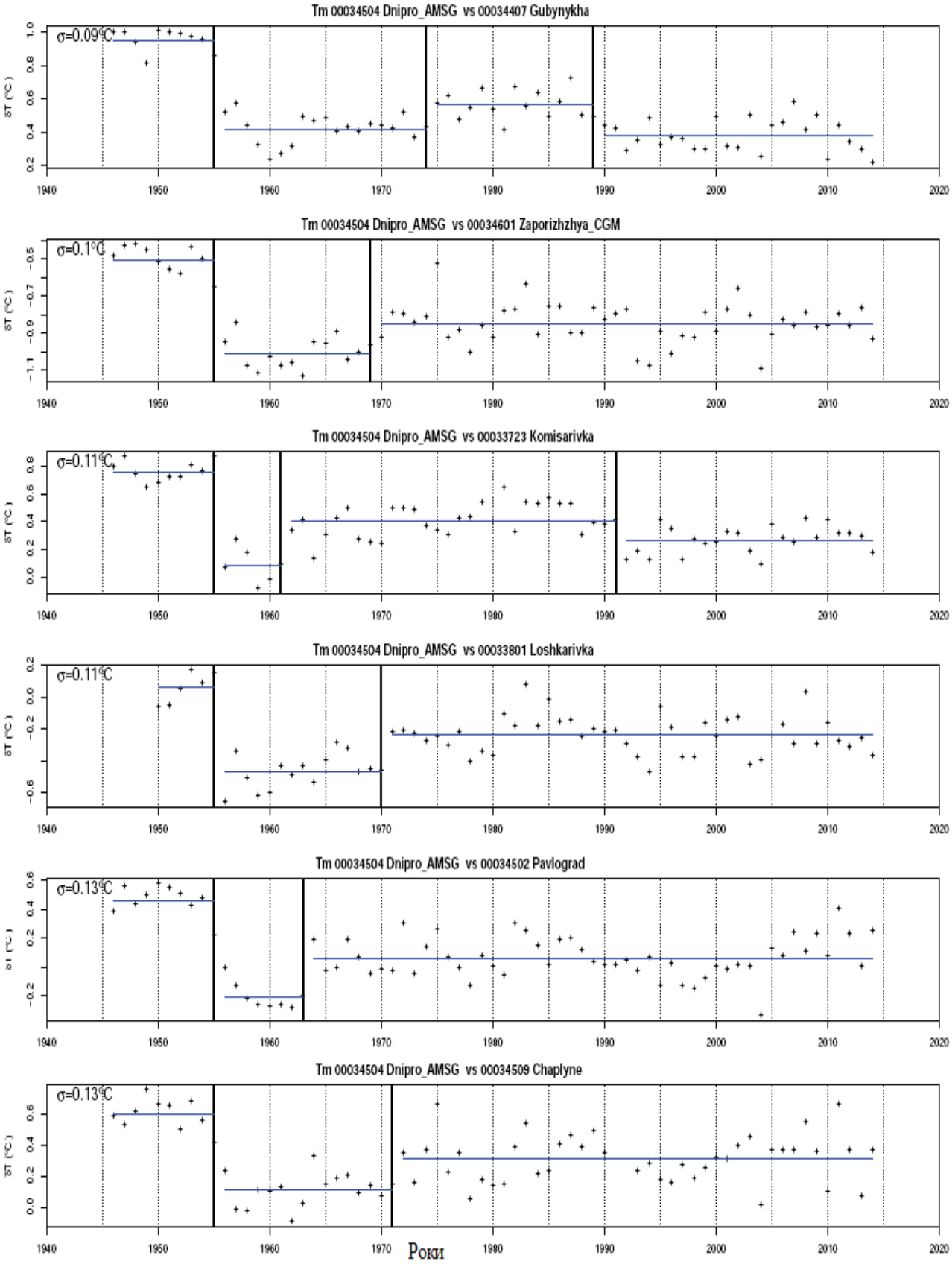


Рис. 10. Результати ідентифікації програмою HOMER точок розриву в річних різницевих рядах для станції Дніпро на основі парного порівняння (pairwise detection)

Fig. 10. HOMER's pairwise detection of the break points in annual difference series for Dnipro station

вих значень. З них 51 значеня (43 %) було ідентифі-

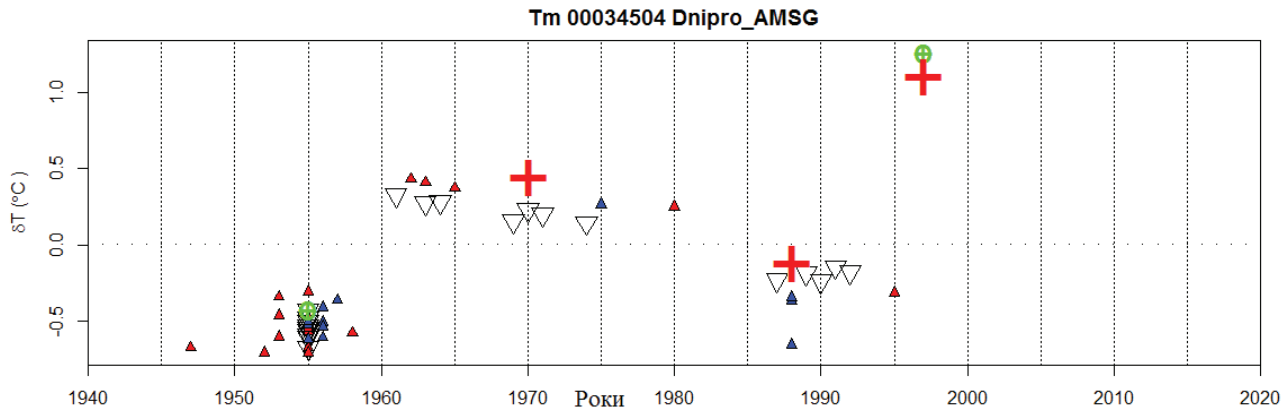


Рис. 11. Результати ідентифікації програмою HOMER точок розриву в рядах станції Дніпро на основі об'єднаного порівняння (joint detection)

Fig. 11. HOMER's joint detection of the break points for Dnipro station

ковано як помилки оцифрування, 49 (42 %) — віднесено до помилок у добових значеннях, 18 (15 %) — ідентифіковано як помилки в довідниках.

Після корекції контроль якості було проведено ще раз і виявлено 26 значень, аномалії яких перевищують порогові значення ±1,5 оС. Ці викиди було замінено на мітки пропусків.

Результати гомогенізації. Загальна кількість ідентифікованих розривів дорівнює 287. На рис. 12 gjrfpfyj розподіл кількості станцій за кількістю точок розриву. Як видно з рисунка, тільки 29 станцій (~16 %) статистичні/гомогенізаційні тести були успішними: не виявлено жодного розриву. Проте, відповідно до метаданих, тільки 11 із них не були переміщені чи зазнали впливу інших чинників. Кількість станцій, що мають 3 розриви, теж дорівнює 29. Більшість станцій (~63 %) мають один—два розриви, і тільки 7 станцій (~4 %) – 4—5. Середня кількість точок розриву на одну станцію дорівнює 1,6, що трохи перевищує середню кількість метаданих на станцію.

Рис. 12. Розподіл кількості станцій за кількістю ідентифікованих точок розриву:

Fig. 12. Distribution of the number of stations with the respect to the number of break points detected

Цікаво порівняти знайдені точки розриву із метаданими. 89 розривів (~31 %) чітко підтверджено з історичних описів станцій (рис. 13). Проте понад половини метаданих (потенційно шкідливих подій) не відображено в ідентифікованих розривах. Однією із можливих причин може бути особливість гомогенізаційних алгоритмів, згідно з якою простіше відшукати неоднорідності в середині часових рядів, ніж на початку чи в їх кінці. Такий висновок можна зробити із порівняння рис. 5 та 14. Інша можлива причина може бути пов'язана із самими потенційно шкідливими подіями. Неідентифіковані події/впливи можуть бути несуттєвими, тобто їхній ефект є незначний і вони не створюють розривів взагалі або розриви мають малу амплітуду і їх важко виявити. Деяким підтвердженням такого висновку може бути той факт, що всі "значні" переміщення станцій (наприклад, із середмість у аеропорти) чітко ідентифіковані програмним забезпеченням HOMER.

Величини розрахованих зсувів (амплітуди/фактори) в річних рядах не перевищують ±1 °С (зазначимо, що різницю між амплітудою та фактором зсуву розтлумачено на рис. 15). Мінімальне значення дорівнює □0,72/□0,68 °С, максимальне — 0,93/0,93 °С. Середнє значення величини зсувів дорівнює □0,03/□0,05 °С, середнє квадратичне відхилення — 0,21/0,22 °С. Частотний розподіл амплітуд/факторів показано на рис. 16. Слід зауважити, що в гістограмах є спад частоти близько 0 °С. Це можна пояснити тим, що дуже малі неоднорідності/зсуви важко виявити через наявність інтенсивної шумової компоненти.

Отримані результати свідчать, що виявлені неоднорідності часових рядів у середньому для території України дещо завищують температуру повітря в минулому, і як наслідок, – занижують інтенсивність її зростання. Зазначений висновок був детально перевірений і обґрунтований подальшими дослідженнями [28].

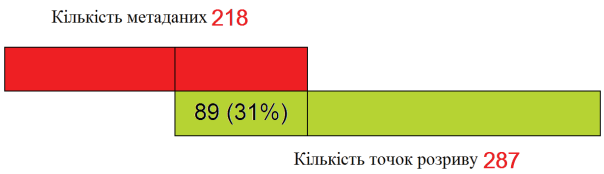


Рис. 13. Діаграма відповідності метаданих і точок розриву

Fig. 13. Correspondence between metadata and the break points detected

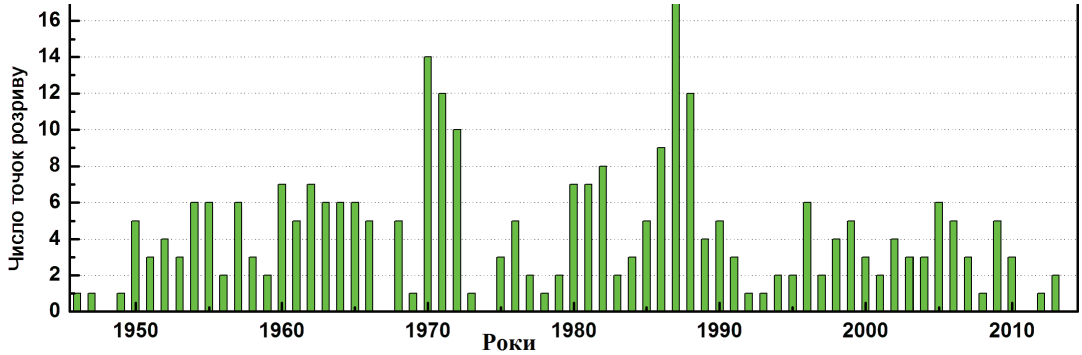


Рис. 14. Часовий розподіл ідентифікованих точок розриву

Fig. 14. Time distribution of the identified break points

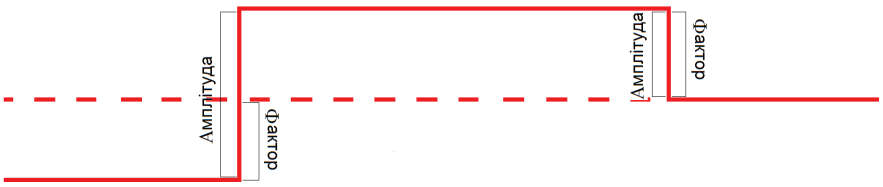
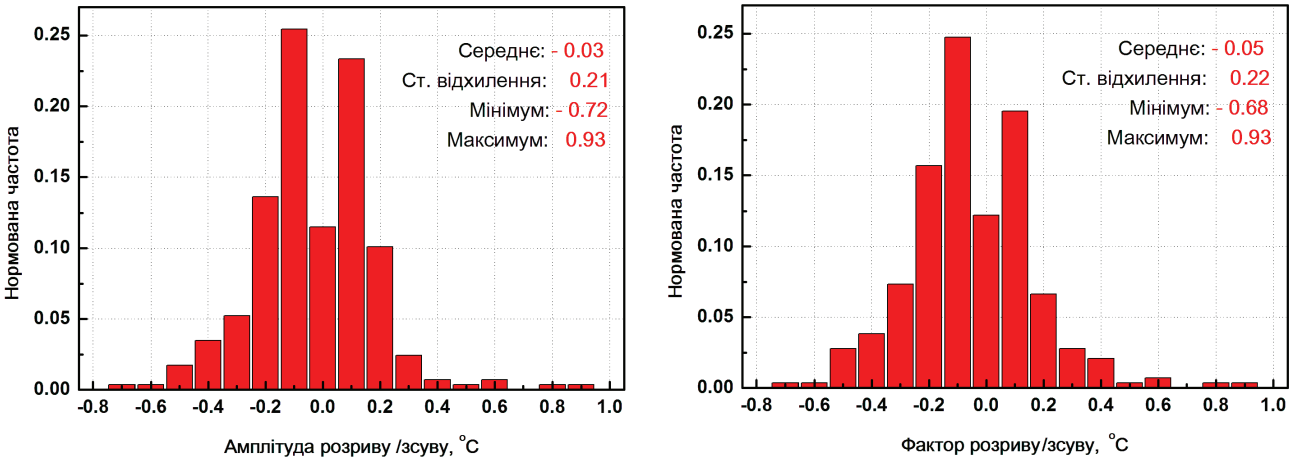


Рис. 15. Типовий "станційний сигнал" та його характеристики (амплітуди/фактори), які розраховує програма HOMER

Fig. 15. Typical "station signal" and its characteristics (amplitudes/factors) calculated by HOMER



а *б*

Рис. 16. Розподіл амплітуд (а) і факторів (б) виявлених зсувів/розривів у річних рядах

Fig. 16. Frequency distribution of amplitudes (a) and factors (б) of detected breaks in the annual series

На рис. 17 представлено виявлені "станційні сигнали" в річних рядах деяких станцій. Як видно із рисунка, багато переміщень мають "охолоджувальний" ефект, що також підтверджує зроблений вище висновок стосовно заниження інтенсивності підвищення температури через вплив неоднорідностей.

Висновки. У роботі наведено результа-ти збору, оцифрування та опрацювання даних середньої місячної температури повітря, які отримано на 514 метеорологічних станціях України за період 1812–2014 рр. З історичних описів станцій зібрано метадані – інформацію щодо потенційно шкідливих подій (перенесення станцій, замін приладів, змін методики та ін.), які можуть мати негативний ефект на часові ряди. Основна ідея такої роботи полягала в отриманні цифрової БД, яка б містила перевірені та гомогенізовані довгі ряди температурних даних

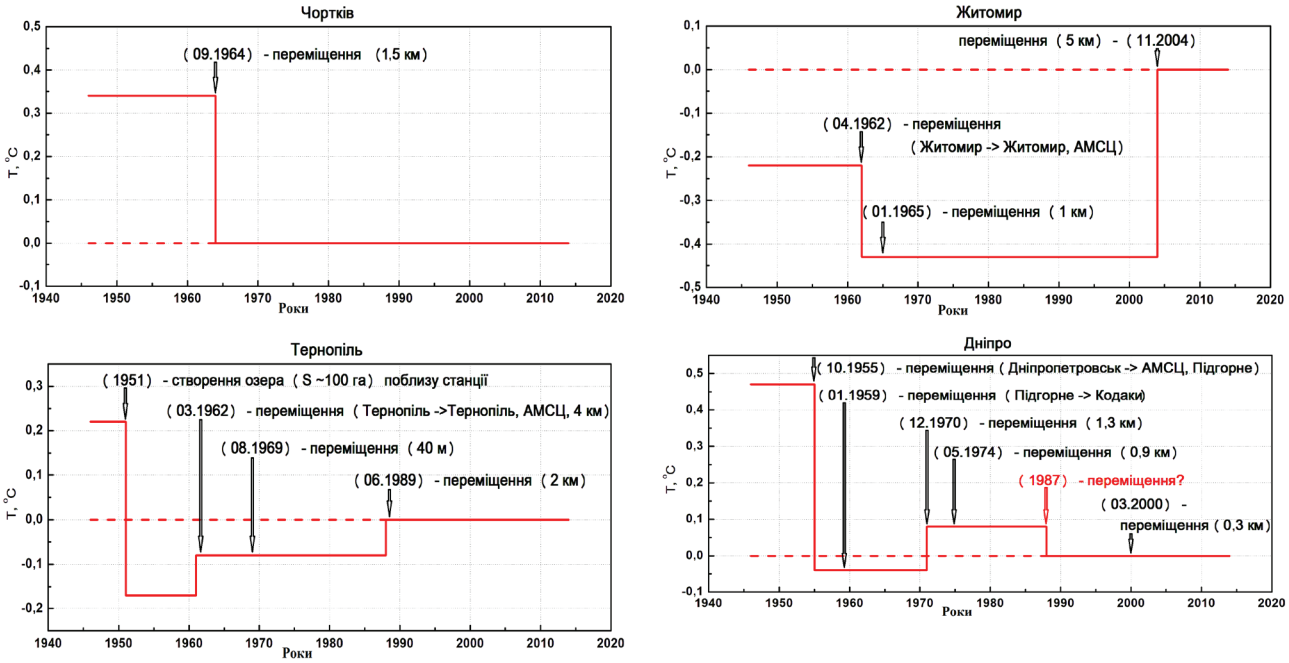


Рис. 17. “Станційні сигнали” у річних рядах деяких станцій (стрілками вказано метадані)
Fig. 17. “Station signals” in annual series of some stations (arrows indicate metadata)

і яку б можна було використати для кліматологічних досліджень температурного режиму України, зокрема для коректного та надійного оцінювання регіональних кліматичних змін.

Аналіз усієї сукупності даних показав, що, на жаль, через велику кількість пропусків протягом Першої та Другої світових воєн надійна гомогенізована БД може бути створена тільки для періоду 1946–2014 рр. Для зазначеного періоду було гомогенізовано часові ряди 178 станцій, що рівномірно розташовані на території України. Гомогенізацію проведено сучасним програмним забезпеченням HOMER, яке успішно верифіовано на реалістичному тестовому матеріалі і яке часто використовують у різних країнах для створення національних гомогенізованих БД кліматологічних показників.

Гомогенізація рядів виявила порівняно велику кількість точок розриву. Їх загальна кількість дорівнює 287 (~1,6 розривів на 1 станцію). Проте тільки 89 із них (31 %) мають чітке підтвердження метаданими. Величина зсувів (амплітуд та факторів “станційних сигналів”) не перевищує ±1 оС. Їх середнє значення дорівнює □0,03/□0,05 оС, середнє квадратичне відхилення — 0,21/0,22 оС. З урахуванням від’ємні середніх значень амплітуд і факторів, а також результатів аналізу ідентифікованих “станційних сигналів” можна припустити, що кліматологічна неоднорідність часових рядів дещо завищує температуру повітря в минулому і, як наслідок, — призводить до заниження оцінок інтенсивності зростання температури в середньому для території України. Втім надійніші висновки можна отримати після проведення детального порівняльного аналізу оригінальних рядів і відповідних рядів створеної гомогенізованої БД.

Список бібліографічних посилань

Бабіченко В.М., Ніколаєва Н.В., Гущина Л.М. Зміни температури повітря на території України наприкінці XX та на початку XXI століття. Укр. геогр. журн. 2007. № 4. С. 3-12.

Бабиченко В.Н., Рудишина С.Ф., Бондаренко З.С., Гущина Л.М Температура воздуха на Украине. Л.: Гидрометеоиздат, 1987. 399 с.

Волощук В.М., Бойченко С.Г., Степаненко С.М., Шищенко Г.П., Бортник С.Ю. Глобальне потепління та клімат України, регіональні екологічні та соціально-економічні аспекти. К.: ВПЦ “Київський університет»”, 2002. 117 с.

Клімат України / За ред. В.М. Ліпінського, В.А. Дячука та В.М. Бабіченко. К.: Вид-во Раєвського, 2003. 343 с.

Кліматологічний довідник. Історія та фізико-географічний опис метеорологічних станцій України. К., 2011. 462 с.

Климатологический справочник СССР. Выпуск 10. По Украинской ССР. История и физико-географическое описание метеорологических станций. Киев. 1968. 456 с.

Кобышева Н.В., Наровлянский Г.Я. Климатологическая обработка метеорологической информации. Л.: Гидрометеоиздат, 1978. 296 с.

Мартазинова В.Ф., Иванова Е.К., Щеглов А.А. Тенденция современного температурно-влажностного режима Украины к аномальности за счет атмосферных процессов в летний сезон. Наукові праці УкрНДГМІ. 2016. Т. 268. С. 15—24.

Метеорологические данные за отдельные годы. Выпуск 10а. По Украинской ССР и Молдавской ССР. Часть 1. Температура воздуха. Том 1. К.: Изд-во Академии наук УССР. 1953. 577 с.

Осадчий В.І., Бабіченко В.М. Температура повітря на території України в сучасних умовах клімату. Укр. геогр. журн. 2013. № 4. С. 32-39.

Справочник по климату СССР. Данные за отдельные годы. Выпуск 10. Украинская ССР. Часть 1. Температура воздуха. К., 1971. 960 с.

Aguilar E., Auer I., Brunet M., Peterson T.C., Wieringa J. WMO Guidelines on climate metadata and homogenization. WCDMP-No 53, WMO-TD No 1186, 2003. WMO, Geneva.

Alexandersson H. A homogeneity test applied to precipitation data. Int. J. Climatol. 1986. V. 6. P. 661-675.

Auer I., Bцhm R., Jurkovic A., Lipa W., Orlik A., Potzmann R. a.o. HISTALP – Historical instrumental climatological surface time series of the greater Alpine region 1760-2003. Int. J. Climatol. 2007. V. 27. P. 17-46.

Caussinus H., Mestre O. Detection and correction of artificial shifts in climate series. J. R. Stat. Soc. Ser. C. 2004. V. 53. P. 405-425.

Easterling D.R., Peterson T.C. A new method for detecting undocumented discontinuities in climatological time series. Int. J. Climatol. 1995. V. 15. P. 369-377.

Freitas L., Pereira M.G., Caramelo L., Mendes M., and Nunes L.F. Homogeneity of Monthly Air Temperature in Portugal with HOMER and MASH. Idojaras, Quart. J. Hungarian Meteorol. Ser. 2013. V. 117, iss. 1. P. 69-90.

Gibelin A.-L., Dubuisson B., Corre L., Madec T. Evolution of temperature and precipitation in France since the 1950s: a new homogenized dataset. Geophysical Research Abstracts. 2015. V. 17.

Guijarro J.A. User's guide to CLIMATOL. An R contributed package for homogenization of climatological series. State Meteorological Agency (AEMET). Balearic Islands Office, Spain, 2014. http://www.climatol.eu/.

HOME. Homepage of the COST Action ES0601 – Advance in Homogenization Methods of Climate Series: An Integrated Approach (HOME), http://www.homogenisation.org

Karl T.R., Diaz H.F., Kukla G. Urbanization: Its detection and effect in the United States climate record. J. Clim. 1988. V. 1. P. 1099-1123.

Krakovska S., Balabukh V., Greben V., Marushevska O. Regional features of global climate change, their assessment, projections and problems of adaptation in the Ukrainian Polesie protected areas. Proc. of the International Conference on Regional Climate CORDEX (Brussels, Belgium), 2013.

Mamara A., Argiriou A.A., Anadranistakis M. Detection and correction of inhomogeneities in Greek climate temperature series. Int. J. Climatol. 2014. V. 34. P. 3024-3043.

Menne M.J., Williams C.N. Jr., Vose R.S. The U.S. historical climatology network monthly temperature data, version 2. Bull. Amer Meteorol. Soc. 2009. V. 90, iss. 7. P. 993-1007.

Mestre O., Aguilar E. HOME_R. Fast documentation. HOMER Training School, 2011. 36 p., http://www.homogenisation.org/HOME_R.pdf!

Mestre O., Domonkos P., Picard F., Auer I., Robin S., Lebarbier E. a.o. HOMER: a homogenization software – methods and applications. Idojaras, Quart. J. Hungarian Meteorol. Ser. 2013. V. 117, iss. 1. P. 47-67.

Osadchyi V., Nabyvanets Y., Skrynyk O.A., Krakovska S., Skrynyk O.Y. Homogenization of long monthly air temperature time series collected in Ukraine. Proceeding of Seventh Seminar for Homogenization and Quality Control in Climatological Databases jointly organized with the Meeting of COST ES0601 (HOME) Action MC Meeting. Budapest: Hungary. 2011. WCDMP-No. 78. P. 92-98.

Osadchyi V., Skrynyk O.A., Radchenko R., Skrynyk O.Y. Homogenization of Ukrainian air temperature data. Int. J. Climatol. 2018. V. 38 (1). P. 497-505.

Peterson T.C., Easterling D.R., Karl T.R., Groisman P., Nicholls N., Plummer N. a.o. (and others) Homogeneity adjustments of in situ atmospheric climate data: a review. Int. J. Climatol. 1998. V. 18. P. 1493-1517.

Skrynyk O.A., Savchenko V., Radchenko R., Skrynyk O.Y. Homogenization of monthly air temperature and monthly precipitation sum data sets collected in Ukraine. Proceedings of Eighth Seminar for Homogenization and Quality Control in Climatological Databases and Third Conference on Spatial Interpolation Techniques in Climatology and Meteorology. Budapest: Hungary. 2014. WCDMP-No. 84. P. 128-133.

Spinoni J., Szalai S., Szentimrey T., Lakatos M., Bihari Z., Nagy A. a.o. Climate of the Carpathian Region in the period 1961–2010: climatologies and trends of 10 variables. Int. J. Climatol. 2015. V. 35, iss. 7. P. 1322-1341.

□tрбnek P. Homogenization of air temperature series in the Czech Republic during a period of instrumental measurements. Proceedings of the Fourth seminar for homogenization and quality control in climatological databases. WMO. Budapest, Hungary. 2004. WCDMP-No. 56. P. 117-133.

Syrakova M., Stefanova M. Homogenization of Bulgarian temperature series. Int. J. Climatol. 2009. V. 29. P. 1835-1849.

Szentimrey T. Multiple analysis of series for homogenization (MASH). Proceedings of the Second Seminar for Homogenization of Surface Climatological Data. Budapest, Hungary. WMO,1999. WCDMP-No. 41. P. 27-46.

Trewin B. Exposure, instrumentation, and observing practice effects on land temperature measurements. WIREs Clim. Change. 2010. V. 1, iss. 4. P. 490-506.

Venema V., Mestre O., Aguilar E., Auer I., Guijarro J.A., Domonkos P. a.o. Benchmarking monthly homogenization algorithms. Clim. Past. 2012. V. 8. P. 89–115.

Vincent L.A. A technique for the identification of inhomogeneities in Canadian temperature series. J. Clim. 1998. V. 11. P. 1094-1104.

WMO. Calculation of monthly and annual 30-years standard normals. World Meteorological Organization, World Climate Data and Monitoring Programme Series, 1989. WCDP 10, WMO-TD 341, Washington D.C., USA.

Надійшла до редакції 22.01.2018 р

ГОМОГЕНИЗИРОВАННАЯ БАЗА ДАННЫХ ДЛИННЫХ РЯДОВ СРЕДНЕМЕСЯЧНОЙ ТЕМПЕРАТУРЫ ВОЗДУХА

В.И. Осадчий[1], О.А. Скриник[2], В.П. Сиденко[1, 3], Д.А. Бойчук[1, 3], Д.А. Ошурок[1], О.Я. Скриник[1]

[1]Украинский гидрометеорологический институт ГСЧС и НАН Украины, просп. Науки, 37, г. Киев, 03028, Украина, e-mail: osad@uhmi.org.ua, dimaoshurok@gmail.com, skrynyk@uhmi.org.ua

[2]Национальный университет биоресурсов и природопользования Украины, ул. Героев Обороны, 15, г. Киев, 03041, Украина, e-mail: skrynyk@nubip.edu.ua

[3]Киевский национальний университет имени Тараса Шевченко, ул. Владимирская, 64/13, г. Ки□в 01601, Украина, e-mail: vladyslavsidenko@gmail.com

Создана цифровая база данных (БД) среднемесячных значений температуры воздуха, полученных на 514 метеорологических станциях, которые работали на территории Украины в тот или иной период времени с 1812 по 2014 г. Для всех станций собраны метаданные (историческая информация о датах переноса станций или других изменениях в проведении измерений).

Показано, что надежные результаты относительно гомогенизации временных рядов можно получить только для периода 1946–2014 гг. вследствие большого количества пропусков в течение Первой и Второй мировых войн. Для указанного периода проведена гомогенизация рядов данных 178 станций. Выявлены точки разрыва и определены величины смещений “станционных сигналов”. Рассчитаны статистические характеристики обнаруженных неоднородностей. Сравнительно большое количество метаданных и идентифицированных разрывов и их рассчитанные статистические характеристики показывают, что неоднородность первичной информации следует обязательно учитывать во время проведения климатологических исследований, в частности при оценке региональных климатических изменений.

Ключові слова: среднемесячная температура воздуха, длинные временные ряды, климатологическая неоднородность, гомогенизация, программа HOMER, точки разрыва, “станционный сигнал”.

HOMOGENIZED DATABASE OF LONG MONTHLY MEAN AIR TEMPERATURE TIME SERIES

V.I.Osadchyi[1], O.A. Skrynyk[2], V.P. Sidenko[1,3], D.O. Boichuk[1,3], D.O. Oshurok[1], O.Y. Skrynyk[1]2

[1]Ukrainian Hydrometeorological Institute under the State Emergency Service of Ukraine and the NAS of Ukraine, 37 Nauky Ave., Kyiv, 03028, Ukraine, e-mail: osad@uhmi.org.ua, dimaoshurok@gmail.com, skrynyk@uhmi.org.ua

[2]National University of Life and Environmental Sciences of Ukraine, 15 Heroyiv Oborony Str., Kyiv, 03041, Ukraine, e-mail: skrynyk@nubip.edu.ua

[3]Taras Shevchenko National University of Kyiv, 64/13 Volodymyrska Street, Kyiv, 01601, Ukraine, e-mail:

Purpose. The main purpose of the study is to create the database (with the maximum observation period and the maximum number of stations) of long homogenized time series of monthly mean air temperature recorded on the territory of Ukraine.

Methodology. The monthly mean air temperature data obtained at 514 meteorological stations located on the Ukrainian territory during the period of 1812–2014 have been processed from different sources and digitized where it was necessary. Besides, metadata (an information about potentially adverse events that can have a negative effect on the time series) have been also collected from historical descriptions of the stations. It was shown that the reliable homogenized database can be created only for the period of 1946–2014 because the great amount of missing data during World War I and II. For this period, we have homogenized time series of 178 stations by means of the HOMER software. Detected breaks/inhomogeneities have been analyzed and their statistics have been calculated.

Findings. The total number of breaks that we have identified is 287. However, only approximately 31% of them can be explained by the station relocations or other reasons. The shift magnitudes (amplitudes and factors) in annual series do not exceed ±1 □C with the small negative mean values. The negative mean values of amplitudes and factors, as well as the analysis of the identified “station signals” suggest that the climatological inhomogeneity of the time series increases the air temperature in the past and, as a result, leads to the underestimation of the temperature trend rise on average for the territory of Ukraine.

Practical value. The large number of metadata (station relocations and other events) along with identified breaks and their statistics show that the inhomogeneity of the raw climatological data must be taken into account when studying the climate dynamics. The created homogenized database can be used for regional climatological studies, in particular, for the correct estimation of the air temperature changes..

Keywords. monthly mean air temperature, homogenization of time series, HOMER, break points, “station signal”1. Atmospheric electricity. Proceedings of the Third All-Union Symposium. Tartu, 28-31 October 1986. Hydrometeoizdat, 1988. 288 p. [in Russian].

References:

Babichenko V.M. Nikolaeva N.V., Gushchina L.M. Changes in air temperature in Ukraine in the late XX and early XXI century. *Ukrainian Geographical Journal*, 2007. № 4. pp. 3-12. [In Ukrainian].

Babichenko, V.N., Rudyshina, S.F., Bondarenko, Z.S. & Gushchyna, L.M. Air temperature in Ukraine. Leningrad: Gidrometeoizdat, 1987. 399 p. [in Russian].

Voloshchuk, V.M., Boychenko, S.G., Stepanenko, S.M., Shyshchenko, G.P. & Bortnyk, S.Y. Global warming and climate of Ukraine, regional, ecological and socio-ecological aspects. Kyiv: VPC "Kyiv University", 2002. 117 p. [in Ukrainian].

Climate of Ukraine. Ed. V.M. Lipinski, V.A. Diachuk, V.N. Babichenko. Kyiv: Raievski publ, 2003. 343 p. [in Ukrainian].

Handbook. 2011. Climatological Handbook "History and Physiographic Description of Ukrainian Meteorological Stations". Publishing of Academy of Science of USSR: Kiev, Ukraine, 462 p. [in Ukrainian].

Handbook. 1968. Climatological Handbook USSR, Issue 10, Ukrainian SSR "History and Physico-Geographical Description of Meteorological Stations". Publishing of Academy of Science of USSR: Kiev, Ukraine, 456 p. [in Russian].

Kobysheva, N.V. & Narovlyansky, G.Y. Climatological processing of meteorological information. Leningrad: Gidrometeoizdat, 1978. 296 p. [in Russian].

Martazinova V.F., İvanova E.K., Shchehlov O.O. The trend of the modern temperature and humidity regime of Ukraine to abnormality due to atmospheric processes in the summer season. *Naukovi pratsi UkrNDHMI*. 2016. vol. 268. – pp. 15-24. [In Ukrainian].

Handbook. 1953. Meteorological Data for Separate Years. Issue 10a. Ukrainian SSR and Moldavian SSR. Part 1. Air Temperature. Publishing of Academy of Science of USSR: Kiev, Ukraine, 577 pp. [in Russian].

Osadchyi V.I., Babichenko V.M. The air temperature on the territory of Ukraine in modern climate conditions. *Ukrainian Geographical Journal*, 2013. № 4. pp. 32-39. [In Ukrainian].

Handbook. 1971. Handbook on Climate of USSR. Data for Separate Years. Issue 10. Ukrainian SSR. Part 1. Air Temperature. Publishing of Academy of Science of USSR: Kiev, Ukraine, 960 p. (in Russian).

Aguilar E., Auer I., Brunet M., Peterson T.C., Wieringa J. WMO Guidelines on climate metadata and homogenization. WCDMP-No 53, WMO-TD No 1186, 2003. WMO, Geneva.

Alexandersson H. A homogeneity test applied to precipitation data. *Int. J. Climatol.* 1986. vol. 6. pp. 661-675.

Auer I., Bцhm R., Jurkovic A., Lipa W., Orlik A., Potzmann R. a.o. HISTALP – Historical instrumental climatological surface time series of the greater Alpine region 1760-2003. *Int. J. Climatol.* 2007. vol. 27. pp. 17-46.

Caussinus H., Mestre O. Detection and correction of artificial shifts in climate series. *J. R. Stat. Soc. Ser. C.* 2004. vol. 53. pp. 405-425.

Easterling D.R., Peterson T.C. A new method for detecting undocumented discontinuities in climatological time series. *Int. J. Climatol.* 1995. vol. 15. pp. 369-377.

Freitas L., Pereira M.G., Caramelo L., Mendes M., and Nunes L.F. Homogeneity of Monthly Air Temperature in Portugal with HOMER and MASH. *Idojaras, Quart. J. Hungarian Meteorol. Ser.* 2013. vol. 117, iss. 1. pp. 69-90.

Gibelin A.-L., Dubuisson B., Corre L., Madec T. Evolution of temperature and precipitation in France since the 1950s: a new homogenized dataset. *Geophysical Research Abstracts.* 2015. vol. 17.

Guijarro J.A. User's guide to climatol. An R contributed package for homogenization of climatological series. State Meteorological Agency (AEMET). Balearic Islands Office, Spain, 2014. http://www.climatol.eu/.

HOME. Homepage of the COST Action ES0601 – Advance in Homogenization Methods of Climate Series: An Integrated Approach (HOME), http://www.homogenisation.org

Karl T.R., Diaz H.F., Kukla G. Urbanization: Its detection and effect in the United States climate record. *J. Clim.* 1988. vol. 1. pp. 1099-1123.

Krakovska S., Balabukh V., Greben V., Marushevska O. Regional features of global climate change, their assessment, projections and problems of adaptation in the Ukrainian Polesie protected areas. *Proc. of the International Conference on Regional Climate CORDEX (Brussels, Belgium)*, 2013.

Mamara A., Argiriou A.A., Anadranistakis M. Detection and correction of inhomogeneities in Greek climate temperature series. *Int. J. Climatol.* 2014. vol. 34. pp. 3024-3043.

Menne M.J., Williams C.N. Jr., Vose R.S. The U.S. historical climatology network monthly temperature data, version 2. *Bull. Am. Meteorol. Soc.* 2009. vol. 90, iss. 7. pp. 993-1007.

Mestre O., Aguilar E. HOME_R. Fast documentation. HOMER Training School, 2011. 36 p., http://www.homogenisation.org/HOME_R.pdf!

Mestre O., Domonkos P., Picard F., Auer I., Robin S., Lebarbier E. a.o. HOMER: a homogenization software – methods and applications. *Idojaras, Quart. J. Hungarian Meteorol. Ser.* 2013. vol. 117, iss. 1. pp. 47-67.

Osadchyi V., Nabyvanets Y., Skrynyk O.A., Krakovska S., Skrynyk O.Y. Homogenization of long monthly air temperature time series collected in Ukraine. Proceeding of Seventh Seminar for Homogenization and Quality Control in Climatological Databases jointly organized with the Meeting of COST ES0601 (HOME) Action MC Meeting. Budapest: Hungary. 2011. WCDMP-No. 78. pp. 92-98.

Osadchyi V., Skrynyk O.A., Radchenko R., Skrynyk O.Y. Homogenization of Ukrainian air temperature data. *Int. J. Climatol.* 2018. V. 38 (1). P. 497-505.

Peterson T.C., Easterling D.R., Karl T.R., Groisman P., Nicholls N., Plummer N. a.o. (and others) Homogeneity adjustments of in situ atmospheric climate data: a review. *Int. J. Climatol.* 1998. vol. 18. pp. 1493-1517.

Skrynyk O.A., Savchenko V., Radchenko R., Skrynyk O.Y. Homogenization of monthly air temperature and monthly precipitation sum data sets collected in Ukraine. Proceedings of Eighth Seminar for Homogenization and Quality Control in Climatological Databases and Third Conference on Spatial Interpolation Techniques in Climatology and Meteorology. Budapest: Hungary. 2014. WCDMP-No. 84. P. 128-133.

Spinoni J., Szalai S., Szentimrey T., Lakatos M., Bihari Z., Nagy A. a.o. Climate of the Carpathian Region in the period 1961–2010: climatologies and trends of 10 variables. *Int. J. Climatol.* 2015. vol. 35, iss. 7. pp. 1322-1341.

□t□рбnek P. Homogenization of air temperature series in the Czech Republic during a period of instrumental measurements. Proceedings of the Fourth seminar for homogenization and quality control in climatological databases. WMO. Budapest, Hungary. 2004. WCDMP-No. 56. pp. 117-133.

Syrakova M., Stefanova M. Homogenization of Bulgarian temperature series. *Int. J. Climatol.* 2009. vol. 29. pp. 1835-1849.

Szentimrey T. Multiple analysis of series for homogenization (MASH). Proceedings of the Second Seminar for Homogenization of Surface Climatological Data. WMO, Budapest, Hungary. 1999. WCDMP-No. 41. pp. 27-46.

Trewin B. Exposure, instrumentation, and observing practice effects on land temperature measurements. *WIREs Clim. Change*. 2010. vol. 1, iss. 4. P. 490-506.

Venema V., Mestre O., Aguilar E., Auer I., Guijarro J.A., Domonkos P. a.o. Benchmarking monthly homogenization algorithms. *Clim. Past.* 2012. vol. 8. pp. 89–115.

Vincent L.A. A technique for the identification of inhomogeneities in Canadian temperature series. *J. Clim*. 1998. vol. 11. pp. 1094-1104.

WMO. Calculation of monthly and annual 30-years standard normals. World Meteorological Organization, World Climate Data and Monitoring Programme Series, 1989. WCDP 10, WMO-TD 341, Washington D.C., USA.

Received 22/01/2018