Bridging the gap: rescuing and digitizing historical meteorological records

Kofi Asare[1]*, Nana Ama Browne Klutse[2],
Stephen Aboagye-Ntow[1], Caroline Edinam Doe[1],
Antwi-Boasiako Amoah[3], Eric Asuman[4], Martin Addi[1] and
Benjamin Adjetey Wemegah[1]

[1]Remote Sensing and Climate Center, Ghana Space Science and Technology Institute, Accra, Ghana,
[2]Department of Physics, University of Ghana, Accra, Ghana, [3]Environmental Protection Agency, Accra,
Ghana, [4]Ghana Meteorological Agency, Accra, Ghana

This research article presents the findings of a data rescue and digitization initiative conducted in selected districts across Ghana, focusing on the Ghana Meteorological Agency (GMet) archive and field stations. The study involved 200 stations, spanning 130 districts, with the primary goal of rescuing and digitizing climatological data. The number of rainfall stations has reduced from 518 in 1976 to 87 in 2021, while the number of temperature stations have decreased from 138 in 1976 to 40 in 2021. Through extensive data searches at the GMet archive and fieldwork, data were successfully rescued, digitized, and quality-controlled, reducing the percentage of missing data and enhancing overall data availability. The study encountered challenges including a lack of proper record-keeping offices at observing stations, recently assigned observers unaware of historical data, insufficient station metadata, and faulty or broken equipment. Recommendations include regular audits by GMet to prevent further data loss, development of comprehensive station records and metadata, implementation of efficient data transmission methods, transition to electronic data transfer systems, and upgrading manual stations to automatic weather stations (AWS). These measures are essential for improving the robustness and reliability of meteorological data collection in Ghana, which is critical for accurate weather forecasting, climate monitoring, and informed decision-making across various sectors.

KEYWORDS

meteorological data, data rescue, digitization, climate research, historical records

1 Introduction

In an era of unprecedented climate change and technological advancements, the preservation and digitization of data are crucial pillars in safeguarding our collective understanding of the past, present, and future. The journey toward securing and breathing new life into historical records and valuable information is a pivotal task in the age of digital transformation. Meteorological data, which include information on temperature, precipitation, air quality, and other atmospheric variables, form the basis for climate research, enabling scientists to analyze trends, model future scenarios, and develop strategies to address climate change (Immerzeel et al., 2014). Meteorological data from historical records provide essential context for understanding climate variability and change. However, these records are often scattered, poorly preserved, or in obsolete formats, putting them at risk of loss and limiting their accessibility. Preserving historical meteorological data presents significant challenges. By analyzing historical meteorological data, researchers can reconstruct past climate conditions,

identify climate cycles and oscillations, and discern patterns of climate variability and climate extremes (Chevalier, 2022; Mateus et al., 2020; Brönnimann et al., 2018; Salonen et al., 2019; Klein Tank et al., 2009). Historical meteorological data provide crucial baseline information for climate research, aiding in the identification of long-term trends and natural climate variability. The loss or degradation of these records hampers our ability to assess historical climate conditions accurately and understand the drivers of climate change (Bloomfield et al., 2022).

Historical records are often dispersed across various formats, such as paper documents, handwritten logs, and outdated digital storage systems, making data rescue and digitization a complex and labor-intensive process (Capozzi et al., 2020). In addition, the deterioration of physical records due to aging, disasters, or neglect further complicates efforts to protect this valuable information (Brázdil et al., 2019). Additionally, technological obsolescence and traditional archives, repositories, and historical documents face the constant threat of deterioration, loss, or becoming obsolete (Rahman, 2020; Masenya and Ngulube, 2019; Rakemane and Mosweu, 2020; Evens and Hauttekeete, 2011; Rasaki et al., 2023). The loss or degradation of historical meteorological data poses significant challenges for climate research, hindering efforts to reconstruct past climate conditions, validate climate models, and assess long-term climate trends. Moreover, the inability to access or utilize these records limits our ability to understand the drivers of climate change, predict future climate scenarios, and develop effective adaptation and mitigation strategies (Mateus et al., 2020; Ryan et al., 2018; Christiansen and Ljungqvist, 2017).

Preserving historical meteorological records is essential as these data offer valuable insights into long-term climate patterns, aiding researchers in establishing baselines, identifying anomalies, and evaluating the impact of human activities on the environment (Seshaba, 2024). The urgency to preserve these climatological paper archives propels the data rescue and digitization initiatives forward. Data rescue initiatives play a vital role in preserving these records, preventing their loss for future generations (Erb et al., 2022; Ashcroft et al., 2016; Williamson et al., 2015). The need to rescue and digitize data has become paramount due to climate change and the digital revolution (Capozzi et al., 2020; Brönnimann et al., 2019; Ryan et al., 2018; Brunet and Jones, 2011; Wilkinson et al., 2011).

Beyond mere preservation, digitization opens the floodgates to accessibility, enabling a broader audience to engage with historical and scientific data once confined to the shelves of archives (Odunayo et al., 2021). Climate data rescue involves organizing and preserving recorded instrument observations and climate data at risk of loss. Data rescue is critically important to ensure that future generations of scientists and other data users have access to all the information necessary to assess climate variability and change and provide a range of climate services (Luterbacher, 2024).

Limited access to data is a major challenge hindering Ghana's efforts to strengthen climate resilience as part of its national adaptation strategies. Although livelihood activities are rainfall-dependent, scientific evidence does not inform decisions because of the lack of access and sparsity of observation stations. The limited and sparse network of weather stations creates gaps that make it difficult to use them for any useful climate analysis and adaptation purposes. Some of these datasets are archived in paper formats. Some archived datasets have been lost while the remaining continue to deteriorate. The need to rescue these archived datasets is much more compelling at present

than ever before. Our understanding of the past, present and future climate hinges on quality climate datasets (Zhou et al., 2022). We must rescue what is available now and put structures in place to continue. Data rescue and digitization efforts aim to address these challenges, ensuring the preservation and utilization of valuable meteorological information. This project is an initiative under the National Adaptation Programme to rescue and digitize rainfall and temperature data from selected districts in Ghana. The project was implemented in collaboration with the Ghana Meteorological Agency (GMet). The study focused on selected districts in Ghana. This paper presents a pilot data rescue and digitization of some climatological paper archives for selected districts in Ghana under the National Adaptation Programme (NAP). The aim of this study is to share how the data were searched, located, digitized, and quality-controlled.

2 Methodology

This study focused on rainfall and temperature, the most critical climate variables. The data rescue exercise at the archives involved 200 weather stations, spanning a total of 130 districts. However, due to limited funding, the fieldwork was only able to cover 40 stations across 13 districts. The search focused on locating relevant meteorological forms (such as 100, 101, 103/3, etc.) containing rainfall and temperature data. The data rescue and digitization efforts mainly focused on the 10 NAP districts, along with additional districts agreed upon with GMet. This study employed a well-thought-out method for data rescue and digitization, which helped in achieving the objectives and expected outcomes as indicated above. The adopted method is consistent with the WMO guidelines for data rescue and digitization.

2.1 Meeting GMet management and training junior scientist

The management of GMet consulted and briefed about the data rescue and digitization process to get their buy-in and ownership beyond the exercise for continuity and sustainability. The meeting highlighted the importance of data rescue and digitization to the Ghana Meteorological Agency. A training session was organized to understand and appreciate the processes involved in data rescue and digitization, enabling them to learn and contribute meaningfully. To promote inclusiveness, sustainability, and ownership by GMet, junior scientists from GMet, as well as from academic and research institutions with relevant expertise, were involved throughout the entire process. Gender balance was also prioritized in the selection of junior scientists to ensure diversity and equity (see Figure 1).

2.2 Data search and locate

Searching and locating the missing dataset is fundamental to the data rescue and digitization process. Permission was obtained from GMet leadership to begin the search and rescue process, both at the archives and in the field. The field visit involved liaising with heads of the regional and district offices and the officers at the observatory. These officers were given prior notice and were briefed about the purpose of the visit to their offices.

FIGURE 1
Head of PMU and DG of GMet interacting with a junior scientist at the workstation.

2.3 Data retrieval and formatting

The GMet database was queried to obtain all rainfall and temperature data with their available metadata. The stations were classified into station types (rainfall, agroclimatic, climatic, and synoptic). The data received were prepared and put in a suitable form to ensure smooth analysis. The rainfall and temperature data with the retrieved metadata were insufficient to provide spatial context. The metadata retrieved from the GMet database included the GPS location, elevation, station ID, station type, and station name. The stations initially lacked district or region information, but this data was subsequently added, with each station being assigned its corresponding district and region. Furthermore, GPS locations with errors were corrected in instances where the coordinates pointed beyond Ghana's borders. Years, months, and missing days were identified and inserted for all the stations for both rainfall and temperature. The missing years, months, and days were compiled for each station to aid in searching for these missing datasets.

2.4 GMet archives

The search and rescue mission was first conducted at the headquarters to look for all missing periods for the selected stations. The GMet archive is the repository for all the data received from the observation stations through the synoptic stations, district, and regional offices for preservation and archiving. Relevant details of all the selected stations, along with their missing periods (day, month, and year) were provided to the GMet archive office to assist in locating the stations and their corresponding rainfall and temperature data (see Figure 2).

Particular attention was given to the period of interest (missing periods). The archive team worked diligently and successfully retrieved some of the missing periods that were located. The junior scientist helped in searching, cleaning, and sorting the retrieved datasets, identifying the missing periods, and organizing and labeling them in envelopes. After completing the search and rescue exercise at the GMet Headquarters archives, an assessment was conducted to determine which data and periods had been rescued and identify what is still outstanding.

2.5 Fieldwork

The Ghana Meteorological Agency primarily relies on a manual data transmission system for the majority of its stations, with the exception of a few synoptic stations that have begun electronic transmission in recent years. Data are first obtained and recorded at the observing station on relevant forms (101, 657, 106, etc.), depending on the station type. These paper records are then transported to the district office or synoptic station, then to the regional office, and further transported to the headquarters. Visiting the regional, district, and observation stations became essential to search for missing data that could not be found in the GMet Headquarters archives. The paper records of rainfall and temperature rescued from the field were cleaned, sorted, enveloped, and labeled with their station IDs and transported to the GMet Headquarters in Accra. In some cases, the rescued records were photocopied, scanned, and printed. All the data rescued from the field and archives at the headquarters were submitted to the directorate for data processing and management for digitization (see Figures 3, 4).

2.6 Data preservation and storage

These paper records rescued from the archives were organized logically and enveloped with their station IDs. Similarly, historical datasets that were recovered from the field (regional, district, and observing stations) for the selected districts were organized logically and transported to the GMet Headquarters for digitization and subsequent archiving by the relevant GMet department following best practices. Although all datasets recorded at the observing stations were sent to the GMet Headquarters for digitization and archiving, there were many instances where some records were found at the regional, district, or observing stations but not in the

FIGURE 2
GMet archive.

FIGURE 3
Visit to Koforidua district Met station.

archives. This was due to the voluntary photocopying of records by the observers. In cases where photocopying was not possible, the found datasets were transported to the headquarters for proper archiving.

2.7 Digitizing data values

The objective of the digitization was to convert analog records into digital format. The GMet database was updated with all the

FIGURE 4
Handwritten data rescued.

rescued data from the archives and the field. The data rescued were mostly analog, with few in digital format. The analog dataset was transcribed into a digital format that could be stored as numerical data, directly readable by a computer and uploaded into the GMet database. All rescued datasets were stored at the workstation set up by the consultant, in consultation with GMet. Junior scientists did the data entry. Datasets were organized by station, year, month, and data type. In addition, measures were implemented to reduce the possible entry of wrongful data (human errors). This option of digitizing was adopted because it is the preferred choice of GMet. The paper records were digitized directly into the CLIDATA database following the recommendations, standards, and protocols of GMet (see Figure 5).

The keying of recovered data was performed by junior scientists and GMet staff due to the sensitivity and restricted access to the database, as determined by the GMet management. Other junior scientists provided support and assisted in the digitization process. The quality control process began during data entry and was finalized at the data quality control management unit at GMet.

2.8 Quality control, validation, and archiving

Quality control (QC) in the CLIDATA database is a critical component that ensures the integrity and reliability of the data collected for operational and research purposes. The methodology for quality control in the CLIDATA encompasses a comprehensive
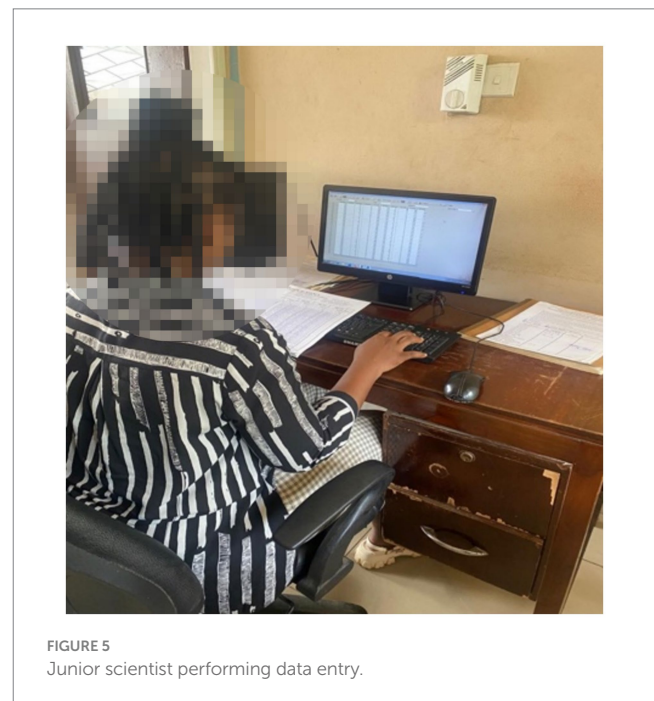

FIGURE 5
Junior scientist performing data entry.

approach that is implemented before, during, and after data collection. This multiphase strategy is essential for identifying and rectifying potential sources of error, thereby enhancing the overall quality of the database.

The automatic quality control (AQC) of data within the CLIDATA database is a critical process that ensures the integrity and reliability of meteorological data. The evolution of quality control methods from manual to fully automated systems has significantly enhanced the efficiency and effectiveness of data validation processes. Automatic quality control systems are designed to identify and flag erroneous data entries, thereby providing quantitative quality information such as quality flags or indices that are essential for data usability in research and operational contexts (Ośródka et al., 2022).

The digitization process followed the best management practices of the Ghana Meteorological Agency and that of the CLIDATA database. The quality control process was strictly followed during the data entry, involving all the missing gaps recovered from the archives and the field. The CLIDATA database used by the GMet has a robust quality control system. CLIDATA has a system in place to flag data (yellow, red, etc.) that exceeds certain thresholds based on the climatology of the station, allowing for the correction of erroneous entries and ensuring the accuracy of the records being entered into the various quality control management systems established by GMet (see Figure 6).

After entering the data and resolving any flagged issue, senior personnel at GMet validate the data by comparing what has been keyed using the paper records. The data keyed into the CLIDATA database are accepted and locked once the data have been validated and accepted as credible, following a rigorous verification process that may even extend to the station level, if necessary. The digitized data were archived after the quality control.

2.9 Analysis

R was used in formatting and visualizing the data. The number of stations per year was computed and plotted. The station network maps displaying the location of the station were generated using QGIS software.

3 Findings

3.1 Station network and baseline report

National adaptation efforts depend significantly on climate data for effective planning and for reporting to the international

FIGURE 6
GMet staff validating and performing quality control of data keyed in digitized.

community. These reports must reflect realities at the community level, which requires access to community-specific climate information. Therefore, assessing the spatial spread of Ghana's rainfall and temperature stations is crucial. Figures 7, 8 show the station network of GMet from its database, which served as the foundation for developing the baseline reference information. This includes both active and closed station networks. GMet had a total of 518 rainfall stations distributed across the country. In the past, the GMet database depicted a relatively dense station network for rainfall, particularly in the southern part of Ghana. The southern part of Ghana could generally be described as having a higher density than the Northern part of Ghana.

There are a total of 138 temperature stations distributed across the country. In comparison to rainfall stations, the distribution of temperature stations (both minimum and maximum) is less dense. Similar to rainfall, Southern Ghana has a relatively higher density of temperature stations than the North.

The baseline information included the first date the station reported data, the last date the station reported data, the number of years with reported data, the number of years with missing data, the percentage of data available, missing periods (day, month, and year), the percentage missing, the district and region where the station is located, station type, station ID, and various forms of duplicates. All the stations had missing gaps. Some stations had long years of observation, whereas others had a few years, depending on when they were established and closed down. The oldest station in Ghana was established at Aburi in 1891, initially as a rainfall station and later upgraded to an Agromet station. The station has the longest record of rainfall data, covering 80 years (see Figure 9).

The Aburi station has 28% of its rainfall data missing. The station is still active and reporting data. The last reporting date (end-date) of some stations is over 50 years ago. This means the station was closed down approximately 50 years ago.

3.2 Trends of the active station networks

The baseline report assessed the number of active rainfall and temperature stations across Ghana annually. This provided information on the interannual variation of active and reporting stations and evaluated the trends of the number of stations active and reporting to the present.

3.2.1 Rainfall

Figure 10 presents the trends of the active rainfall station networks of GMet from 1891 to date. The first rainfall station was set up in Aburi, Ghana in 1891. After this, the number of stations began to increase in the 1930s. The number of rainfall stations exceeded 100 in the early 50s. The gradual increase in rainfall stations continued steadily, reaching over half a million in the 70s. The year with the highest rainfall stations in Ghana was 1976, with approximately 518 rainfall station networks nationwide.

The number of rainfall station networks began to decline in the late 70s, reaching a significant decline level in 1987. The number of station networks began to rise again and stabilized in the 1990s and early 2000s. There has been a steady decline in the rainfall station networks from 2011 to date. The number of rainfall station networks has currently declined from 518 stations to less than 100 active

FIGURE 7
Rainfall stations of GMet.

stations. Approximately 17%, representing 87 of the rainfall station networks, are still active from the reference peak of 1976 as of 2021. This means that approximately 83% of the rainfall stations have been closed down.

3.2.2 Temperature

Figure 11 presents the trends of active temperature (minimum and maximum) stations owned by GMet across the country. The first temperature station was installed in 1946. The number of stations recording temperature data increased until it peaked in 1976, with 138 stations. This peak year for temperature stations is similar to that of the rainfall stations. However, the number of temperature stations began to decline in the late 1970s, reaching a new low in 1987. The temperature station networks began to rise again in the late 1980s and 1990s, reaching a new peak of 120 stations in 2000, though still below the 1976 peak of 138 stations. The number of station networks began to decline in late 2000, and this downward trend has continued, with fewer than 40 active stations recording temperature data as of 2021. Approximately 29% of the temperature station networks, or approximately 40 stations, are still active from the reference peak of 1976. This means that

approximately 71% of temperature stations have been closed down. Although this record is not impressive, it looks relatively better than the rainfall stations.

The decline in both rainfall and temperature stations may not be unique to only Ghana as Sfîcă et al. (2021) have also reported in their study. This is a worrying concern that requires urgent attention globally.

3.3 Impacts of rescue and digitization

The entire exercise of rescue and digitization has yielded results. Figure 12 presents the station distribution network where data were rescued and further digitized from the GMet Headquarters archives. The paper records rescued from the archives emanated from 130 districts covering 200 stations. Data rescued from the archives covered the entire 16 regions of Ghana. Some of the stations where data were rescued have been closed down.

The search involved looking for relevant meteorological forms (100, 101, 103/3, etc.) containing rainfall, temperature, and other
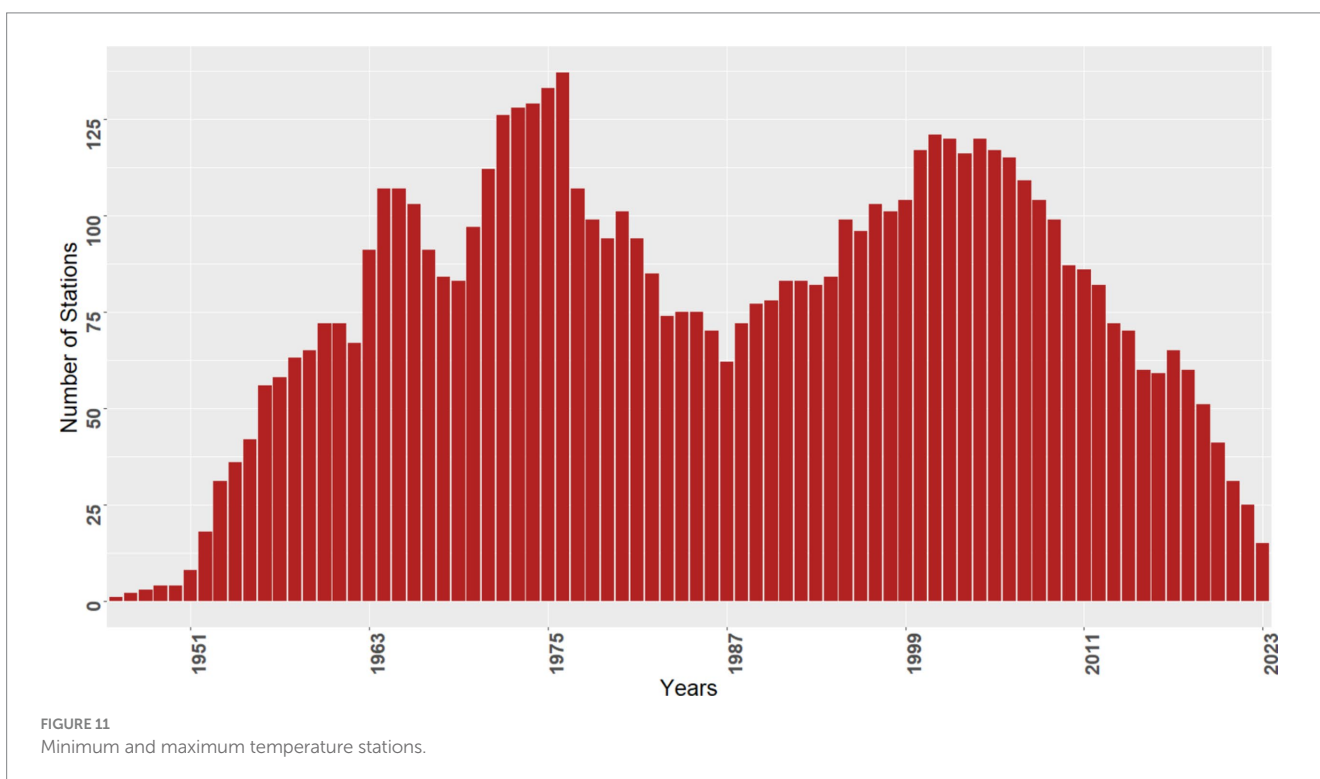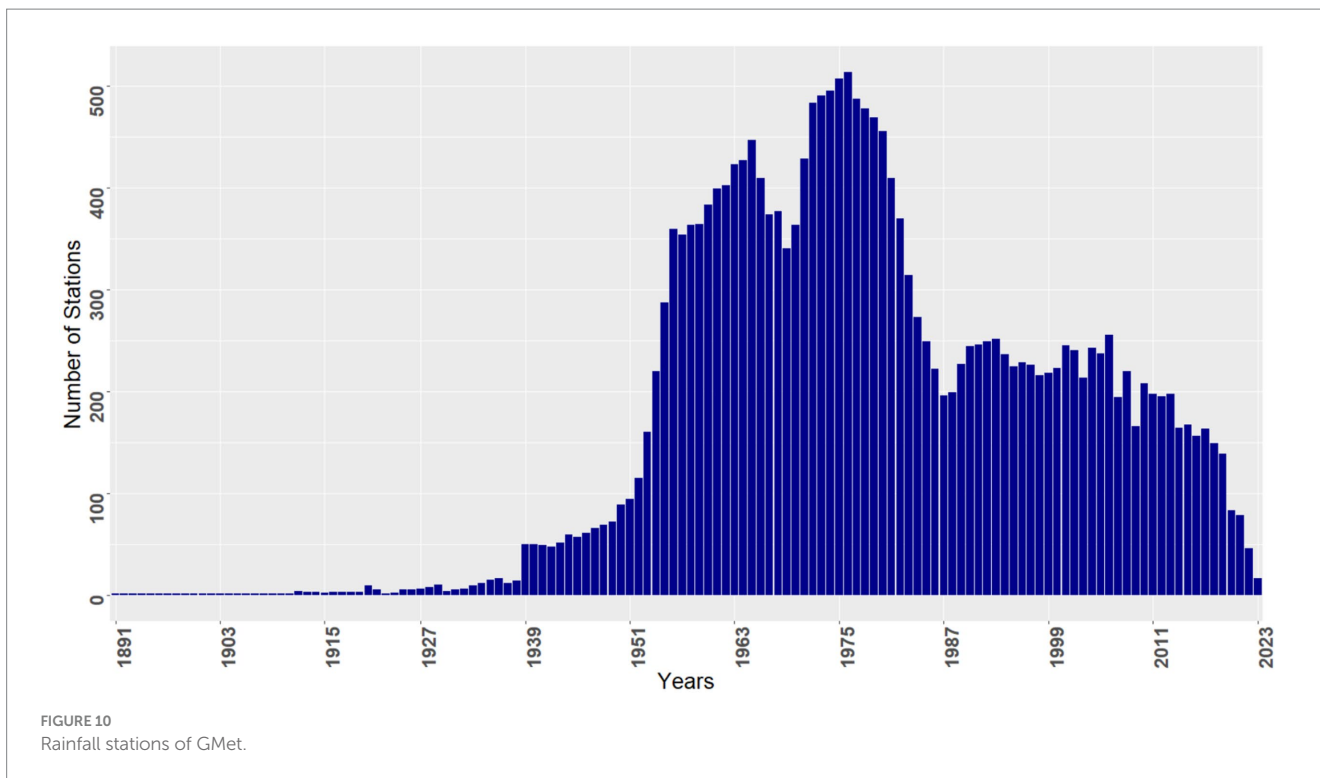
FIGURE 8
Minimum and maximum temperature stations.

StationName	ID	District	Lon	Lat	StartYear	EndYear	Type	Number_Y	Actual_Y	% Available_	% missing_u
National Achives-Accra	23097ACC	Ga Central	-0.21	5.56	2002	2017	Rainfall	13	16	52.4	47.6
ATTC- Nkrumah Circle	23095ACC	Ga Central	-0.21	5.57	2006	2017	Rainfall	2	12	9.8	90.2
Abakoase	21063ABA	Atiwa East	-0.63	6.42	1972	1975	Rainfall	4	4	5.6	94.4
Abamkrome	23075ABA	Prestea-Huni Valley	-0.65	5.82	1965	1970	Rainfall	6	6	6	94
Abenase	21000ABE	Akyemansa	-1.05	6	1939	1990	Rainfall	29	52	26.4	73.6
Abesewa Gyaman	18024ABE	Wassa Amenfi East	-2.02	5.82	1982	2008	Rainfall	24	27	47.1	52.9
Abetifi	14000ABE	Kwahu East	-0.75	6.68	1939	2017	Rainfall	54	79	56	44
Abetifi	14067ABE	Kwahu East	-0.75	6.68	1976	2023	Synoptic	44	48	86.6	13.4
Abetifi-Kwahu	14042ABE	Kwahu East	-0.75	6.68	1964	2018	Climate	28	55	39.8	60.2
Aboantem	14021ABW	Ejura-Sekyedumase	-1.4	7.18	1965	2019	Rainfall	50	55	74.1	25.9
Abofour	17000ABO	Offinso	-1.73	7.13	1956	2002	Rainfall	44	47	52.2	47.8
Abokobi	23000ABO	Ga East	-0.2	5.73	1956	2006	Rainfall	25	51	32.9	67.1
Abomosu	21035ABO	Shai Osudoku	0	6	1935	1961	Agromet	3	27	4.2	95.8
Abomosu	21035ABO	Shai Osudoku	0	6	1935	1961	Rainfall	3	27	4.2	95.8
Abor	23047ABO	Keta	0.87	6.05	1953	1981	Rainfall	29	29	38.3	61.7
Aboso	18039ABO	Prestea-Huni Valley	-1.97	5.37	1939	1982	Climate	39	44	59.1	40.9
Aboso	18039ABO	Prestea-Huni Valley	-1.95	5.35	1939	1982	Rainfall	39	44	59.1	40.9
Abra	18025ABR	Ahanta West	-2.05	4.9	1961	2000	Rainfall	21	40	21.9	78.1
Abrafo-Odumase	23043ABR	Twifo/Heman/Lower Denkyira	-1.38	5.33	1963	1979	Rainfall	17	17	20.7	79.3
Abrem Agona	17004ABR	KEEA Municipal	-1.43	5.18	1954	1979	Rainfall	25	26	30.7	69.3
Abrumase	07001ABR	East Gonja	-1.02	8.48	1961	1964	Rainfall	4	4	4.4	95.6
Abura-Dunkwa	23003ABU	Abura / Asebu / Kwamankese	-1.17	5.33	1958	1989	Rainfall	30	32	38.8	61.2
Aburi	23029ABU	Akuapem South	-0.17	5.85	1891	2022	Agromet	80	132	72	28
Aburi	23029ABU	Akuapem South	-0.17	5.85	1891	2022	Climate	80	132	72	28
Abutia Teti	07010ABU	Ho West	0.37	6.52	1954	1990	Rainfall	32	37	41.7	58.3
Accra	23088SFA	Awutu Senya East	-0.45	5.57	1987	2017	Rainfall	4	31	3.2	96.8
Accra Academy	23094ACC	Accra Metropolitan	-0.24	5.57	1963	2017	Rainfall	21	55	29.9	70.1
Accra-Korle Bu	23004KOR	Ablekuma South Metropolitan	-0.22	5.54	1954	2017	Rainfall	32	64	36.7	63.3

FIGURE 9
Baseline report of GMet stations.

FIGURE 10
Rainfall stations of GMet.

FIGURE 11
Minimum and maximum temperature stations.

climatological data. Some amount of rainfall and temperature data were rescued from the archives. Figure 13 presents a sample of missing periods rescued for selected stations.

The fieldwork was very useful as it led us to retrieve data from the private notebook of a retired staff at Akropong in the eastern region. Figure 14 presents the various stations where data were rescued from

the field visit. Data were rescued from 30 districts covering 40 stations across the country. Data were recovered from Ghana's Bono, Savannah, North East, Upper East, Upper West, Northern, Western, Greater Accra, Ashanti, and Eastern regions.

Although some datasheets were found, there were no records for the parameters of interest. Some observers indicated that due to
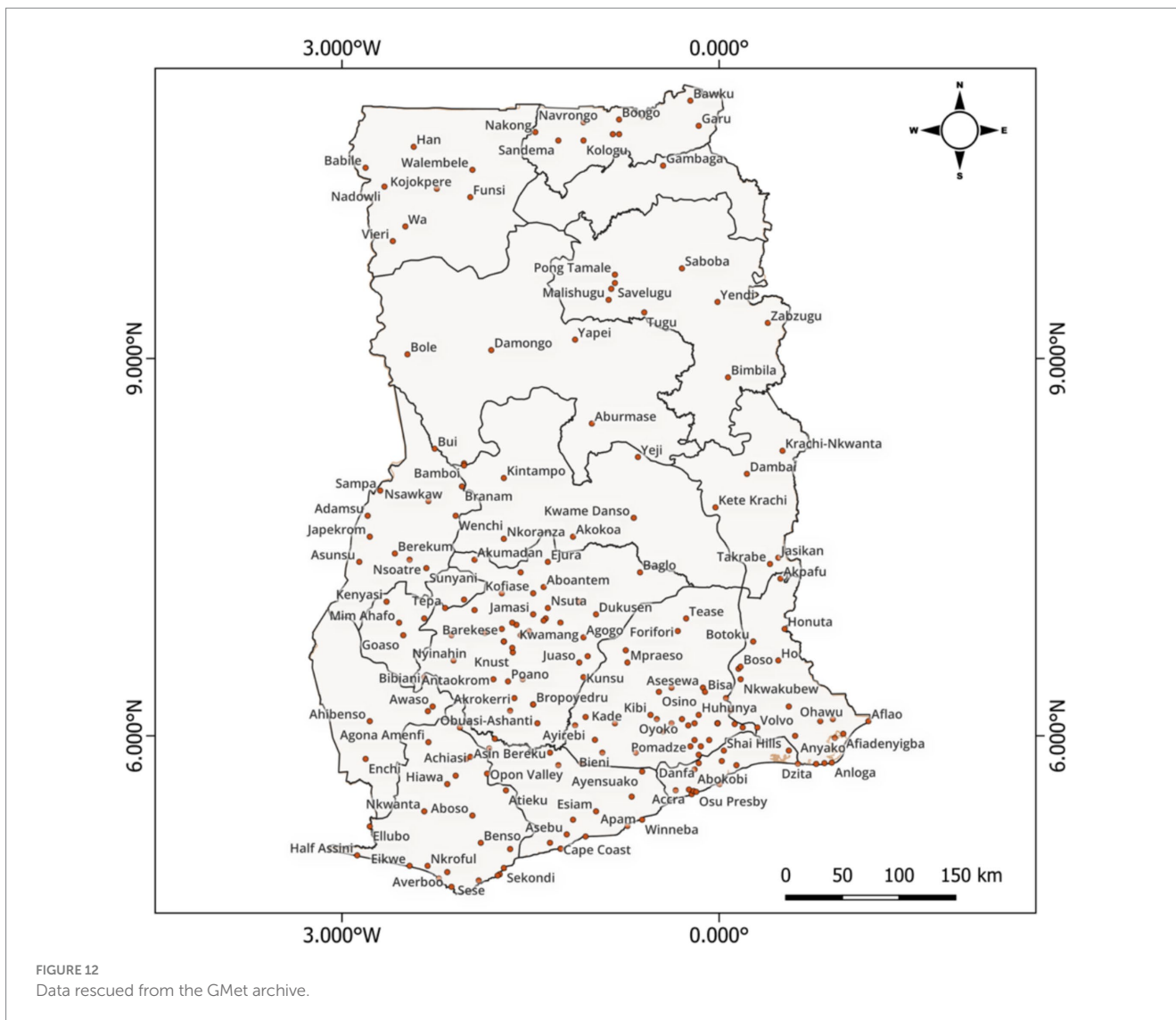
FIGURE 12
Data rescued from the GMet archive.

dysfunctional equipment, records were not taken. Some of the missing records could also not be located.

Table 1 presents the initial baseline information on the percentage of data available and the improvement in the percentage of data available after the rescue and digitization work for selected stations. This gives a quick visual impression of the study and the improvements seen as a result of this study.

The impact made will play a pivotal role in addressing critical gaps in climate modeling and policymaking by unlocking historical records and ensuring data quality for comprehensive analyses. Integrating rescued data into climate models enhances their accuracy by providing larger and more robust datasets, enabling better calibration and improved simulations of climate variability and extremes. These efforts also strengthen evidence-based policymaking, supporting risk assessments and the design of targeted interventions for disaster risk reduction, agriculture, water management, and health. Additionally, the rescued data will help close spatial and temporal gaps in climate records in Ghana, improving geographic coverage and reducing biases in assessments. This further facilitates the development of early warning systems by leveraging historical patterns to improve forecasting and disaster preparedness. In sectoral applications, such as

agriculture, water resource management, and renewable energy planning, enhanced data inform adaptive strategies and resource allocation. Moreover, the preservation of fragile historical records prevents data loss, while digital accessibility fosters global collaboration and innovation in climate science. By supporting localized insights through high-resolution regional models and scenario analyses, these efforts will enable tailored climate adaptation and mitigation strategies. Ultimately, climate data rescued and digitized will ensure the availability of reliable data to guide sustainable development, enhance resilience, and raise public awareness of climate challenges.

3.4 Challenges in rescuing data

The data rescue and digitization work were not without challenges. This section presents some challenges impeding the preservation and digitization of data that require attention.

The majority of the observing stations do not have offices to store records. One of the key challenges observed was the absence of proper office spaces at some stations where weather data are recorded, which

STATION NAME	RAINFALL	T MIN	T MAX
ABETIFI	2020–April to July 2022–Sept to Dec 2023–March-Nov	2012- Sept Nov 2013- Nov 2018-Nov 2020- April to July 2023-April to Nov	2014- Aug to Dec 2015-Jan to Dec 2016-Jan to Dec
SALTPOND	1956-Nov 1984- Mar 1985- July, Nov 1986- July 2020- Feb to July 2021- July to Sept 2022- Jan, Feb, Sept, Oct 2023- May, July to Dec		1986- Nov 2012- July, Sept, Oct, Nov 2015- Feb 2017- June, July, Aug 2020- Mar to July 2021- Mar, May, Jun, Aug to Dec 2022- Mar, Aril, May, June, July 2023- May, June
CAPECOAST	1972-Jan-Dec 1975-Jan-Dec 1984-Jan-Dec 2012-Feb, April to Dec 2015- Jan, April, June, Aug 2016- Jan-Dec 2017- July, Oct, Nov 2018- Jan, Aug, Oct to Dec 2019- Jan, April, Oct to Dec 2020- Jan to Oct 2021- Feb, Oct, Nov 2022- March 2023- Jan to Dec Note:1972 & is in inches		1996- Dec 2012- Feb to Dec 2013- Jan to Dec 2014- Mar to May, July, Dec 2015- Jan to Sept, no Aug 2017- Aug, Sept 2018- Aug 2020- Mar to Dec 2023- Jan to Dec

FIGURE 13
Year, month, and day rescued per station.

means there is no proper place to safely keep these records. This lack of proper offices for storing meteorological data can have significant implications for the preservation and management of these important records. This issue may lead to difficulties in storing and safeguarding weather data, potentially resulting in data loss or degradation. This finding has been corroborated by Coney et al. (2022) and Sfîcă et al. (2021).

Many of the observers we met at the stations were relatively new and had no idea of the older records that were of interest to us. During our visit to the weather stations, we found that most observers were new to their roles, and the lack of office facilities made it impossible to access or retrieve historical records.

Observers always claim that they have sent the records to the headquarters. The observers stated that, as per the policy of the Ghana Meteorological Agency, all paper records were submitted to the headquarters. Therefore, no records were kept at the station.

Some of the missing data gaps, particularly in recent times, were not due to missing records but rather the result of faulty or broken equipment. A number of instruments, especially thermometers, were found to be malfunctioning or broken. Although missing records were rescued at the station level, it was confirmed that no data had been recorded during those periods due to equipment failures, as noted by Brohan et al. (2006).

The metadata of stations were scanty without proper records as indicated by Hunziker et al. (2017) and Hunziker et al. (2018). The existence of a lack of comprehensive metadata of the weather stations was a challenge. Information about stations was scanty and without proper records. This makes it difficult to understand the history of the weather station. The lack of comprehensive metadata for weather stations poses challenges in utilizing and interpreting weather data effectively, especially when there are duplicate stations, or when stations undergo upgrades or relocations. Enhancing the collection, organization, and accessibility of station metadata is essential for improving the reliability, quality, and usability of weather station information for various applications in meteorology and climatology.

The closure of some weather stations has made it difficult to trace and rescue missing data periods in the field. When stations cease operations, rescuing data from those locations becomes problematic, resulting in gaps in historical records, as indicated by Yaseen and Abdulkareem (2022).

3.5 Recommendations

To enhance data rescue and digitization efforts at the Ghana Meteorological Agency (GMet), several recommendations can be drawn from reputable sources in the field.

First, it is suggested that the scope of the study should be broadened to encompass all remaining rainfall and temperature stations in Ghana, with a potential extension to include other meteorological datasets. This expansion would contribute to a more comprehensive and inclusive database for meteorological research and applications. Broadening the scope to include all rainfall and temperature stations, as well as other meteorological datasets, requires conducting an inventory assessment and securing the needed funding to cover the remaining stations and meteorological variables.

In addition, the implementation of a seamless and transparent system at the GMet Headquarters is advised for receiving, filing,

FIGURE 14
Data rescued from GMet fieldwork.
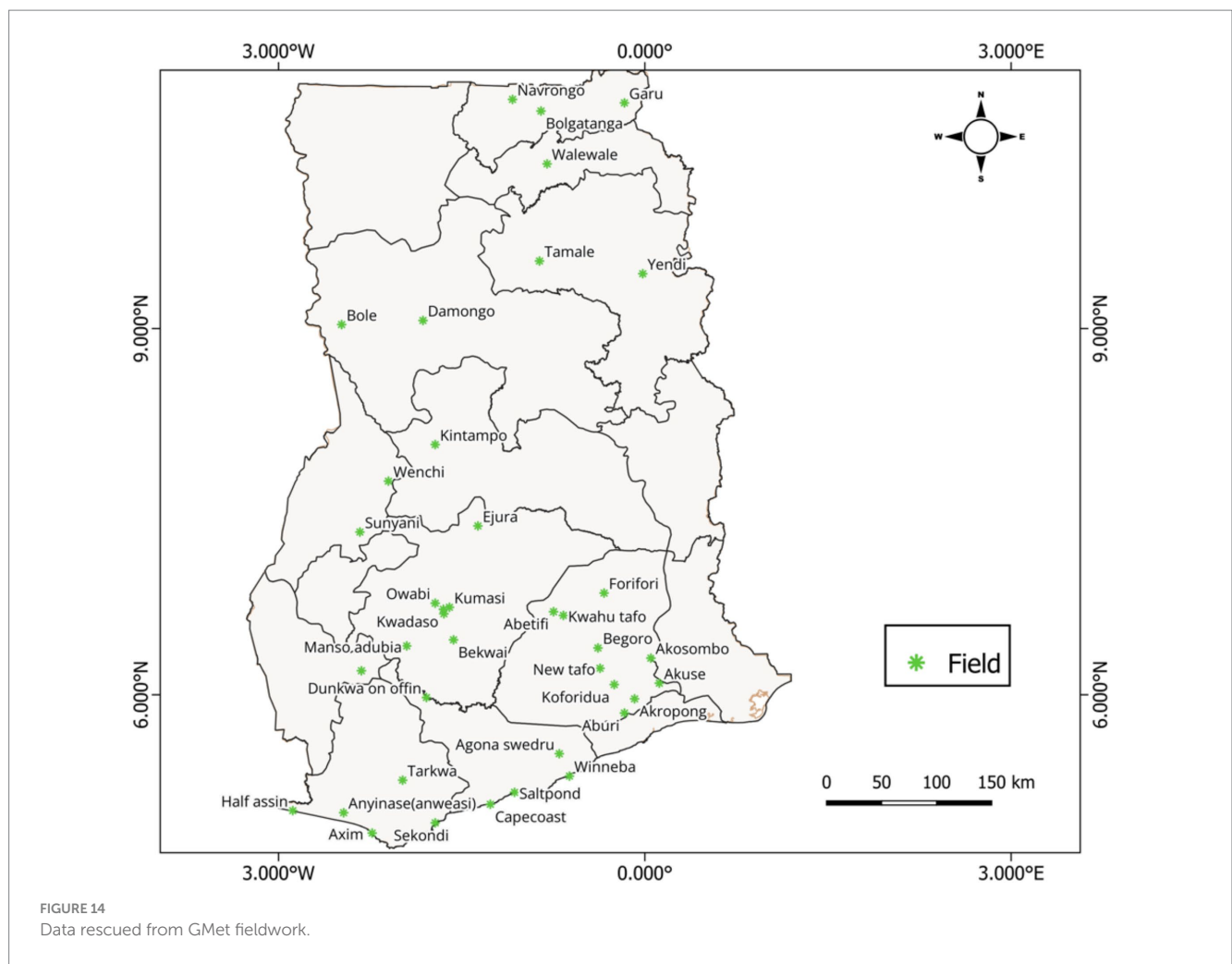
digitizing, conducting quality control, and archiving data with defined timelines and checks and balances. Such a system would ensure efficiency, accuracy, and accountability in the data management process. Establishing a seamless and transparent data management system involves designing workflows, based on the CLIDATA database tools, training and organizing refresher courses for staff, and introducing monitoring mechanisms to ensure efficiency and accountability.

Furthermore, it is recommended that all forms of duplicates within the database should be identified and removed to maintain data integrity and avoid redundancy. Addressing database duplicates can be achieved through data audits, the use of specialized tools for flagging redundant entries, and implementing policies to prevent future duplications. Regular data audits, facilitated by a dedicated team and a well-defined schedule, ensure the integrity of data transmission and archiving processes.

Regular audits by internal auditors at GMet are also proposed to oversee the data transmission process from the origin of the data to the point where it is archived. This will help prevent any loss of valuable data, thereby avoiding missing gaps. Internal audits are practical with a dedicated team and a well-defined audit schedule. The audit team should maintain detailed records of the audit, which should include issues and resolutions identified with follow-up measures ensuring the implementation of audit findings.

Developing comprehensive metadata records for all stations and ensuring effective and efficient data transmission methods, both analog and digital, are crucial steps to enhance data management practices at GMet. Creating comprehensive metadata for all stations following best practices and enhancing transmission methods necessitates defining metadata standards, upgrading infrastructure for reliable analog and digital transmissions, and training staff on their use.

Transitioning from paper-based data transfer to electronic transmission is advised to streamline processes and improve data accessibility and usability. Transitioning to electronic data transfer, while cost-effective, requires investment in improved Internet access and connectivity, capacity building of staff, and phased implementation starting with high-priority stations such as the synoptic weather stations.

Moreover, it is essential to address equipment issues by promptly replacing faulty devices and considering reopening closed-down stations if feasible. Addressing equipment issues and reopening closed stations involves auditing equipment, recruiting new staff, securing targeted funding, and developing maintenance plans to prevent recurring issues.

Additionally, efforts should be made to bridge any existing gaps in the network distribution, particularly focusing on closing the north–south divide to ensure uniform data coverage across regions.

TABLE 1 Impact of rescue and digitization.

Station name	ID	Available before (%)	Available after (%)	Impact (%)
Abetifi	14067ABE	86.6	92.8	6.2
Aburi	23029ABU	72.0	83.7	11.7
Akropong Akwapim	14062AKR	89.3	94.4	5.1
Asutsuare	07063ASU	71.6	83.5	11.9
Atebubu	11002ATE	97.2	97.2	0
Axim	23001AXM	98.1	99.1	1.0
Banko	14034BAN	55.1	38	17.1
Bibiani	16004BIB	63.9	78.0	14.1
Bogoso	18005BOG	90.5	94.8	4.3
Bolgatanga	04027BOL	90.9	95.2	4.3
Cape Coast	23048CAP	56.4	86.1	29.7
Damongo	01001DAM	95.3	98.4	3.0
Dunkwa	17015DUN	97.4	95	2.4
Ejura	14004EJU	99.3	99.7	0.4
Garu	04031GAR	94.5	89.6	4.9
Koforidua	22050KDA	98.9	99.6	0.7
Konongo	19007KON	82.6	90.5	7.9
Kumasi	17036KUM	95.9	83.3	12.6
Sunyani	01032SUN	93.2	98.1	4.9
Tamale	07006TLE	97.8	99.4	1.6
Winneba	23067WIN	67.8	86.2	18.4
Yendi	08017YEN	96.9	99.0	2.1

Bridging these distribution gaps requires spatial analysis to identify underserved areas, strategically deploying stations, partnering with local stakeholders, and continuously monitoring network performance.

The loss of climate data due to the challenges highlighted has far-reaching implications for climate science, policymaking, and societal resilience. Station closures disrupt the continuity of climate records, creating gaps in temporal and spatial datasets that are critical for understanding long-term climate variability and trends. There is an urgent need to rescue, digitize historical datasets, and further implement the recommendations stated above to avoid future reoccurrence. Efforts to rescue and digitize these data are urgent and invaluable. Restoring historical climate records helps fill critical gaps, enabling the reconstruction of long-term climate patterns and improving the granularity of regional and global climate assessments. Such efforts also preserve the scientific and cultural legacy embedded in these records, ensuring that decades of observations are not lost to time or neglect. By making rescued data accessible and integrating it into modern databases, climate scientists and decision-makers can unlock new insights into climate dynamics, strengthen resilience planning, and support sustainable development. The value of climate data rescue lies not only in addressing the deficiencies of the past but also in providing a robust foundation for future research, ensuring that society is better equipped to navigate the challenges of a changing climate. The integrity of climate science, the effectiveness of climate policy, and the resilience of societies to climate impacts hinge on our ability to rescue and maintain high-quality climate data. As such, concerted efforts to restore and enhance climate monitoring networks are not just beneficial but essential for addressing the pressing challenges posed by climate change.

4 Conclusion

This study followed the terms of reference and achieved the deliverables following international best practices. The World Meteorological Organization guidelines for data rescue and digitization, as detailed in the methodology, were followed. Relevant staff at GMet and junior scientists were involved in ensuring sustainability. The data rescue effort included 200 weather stations. Although 200 stations were involved, only 40 were involved in fieldwork after the rescue work at the GMet archives due to limited financial resources. The recommendations should be adhered to avoid the future reoccurrence of missing data where the paper records are available. The management and the majority of staff involved at GMet cooperated in many ways to ensure the success of this study. The GMet weather stations require augmentation to increase the density of the station network across the country and bridge the north–south gap. The data rescue and digitization exercise have been a vital endeavor that has, to some extent, bridged the past with the future, safeguarding invaluable information from the perils of obsolescence and decay. Through meticulous planning, careful execution, and ongoing commitment, these efforts ensure that

historical records, scientific data, and other valuable materials are digitally preserved and made accessible to researchers, educators, and the public for future generations.

Data availability statement

Author contributions

KA: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Supervision, Writing – original draft. NK: Conceptualization, Methodology, Supervision, Writing – review & editing. SA-N: Data curation, Formal analysis, Visualization, Writing – original draft. CD: Data curation, Formal analysis, Validation, Visualization, Writing – original draft. A-BA: Funding acquisition, Supervision, Writing – review & editing. EA: Project administration, Supervision, Writing – review & editing. MA: Data curation, Formal analysis, Visualization, Writing – original draft. BW: Data curation, Formal analysis, Writing – original draft.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

References

Ashcroft, L., Allan, R., Bridgman, H., Gergis, J., Pudmenzky, C., and Thornton, K. (2016). Current climate data rescue activities in Australia. *Adv. Atmos. Sci.* 33, 1323–1324. doi: 10.1007/s00376-016-6189-5

Bloomfield, H., Brayshaw, D., Deakin, M., and Greenwood, D. (2022). Hourly historical and near-future weather and climate variables for energy system modelling. *Earth Syst. Sci. Data* 14, 2749–2766. doi: 10.5194/essd-14-2749-2022

Brázdil, R., Valášek, H., Chromá, K., Dolák, L., Řezníčková, L., Bělínová, M., et al. (2019). The climate in south-east Moravia, Czech Republic, 1803–1830, based on daily weather records kept by the reverend Šimon Hausner. *Clim. Past* 15, 1205–1222. doi: 10.5194/cp-15-1205-2019

Brohan, P., Kennedy, J., Harris, I., Tett, S., and Jones, P. (2006). Uncertainty estimates in regional and global observed temperature changes: a new data set from 1850. *Journal of Geophysical Research Atmospheres*, 111. doi: 10.1029/2005jd006548

Brönnimann, S., Allan, R., Ashcroft, L., Baer, S., Barriendos, M., Brázdil, R., et al. (2019). Unlocking pre-1850 instrumental meteorological records: a global inventory. *Bull. Am. Meteorol. Soc.* 100, ES389–ES413. doi: 10.1175/bams-d-19-0040.1

Brönnimann, S., Brugnara, Y., Allan, R., Brunet, M., Compo, G., Crouthamel, R., et al. (2018). A roadmap to climate data rescue services. *Geosci. Data J.* 5, 28–39. doi: 10.1002/gdj3.56

Brunet, M., and Jones, P. (2011). Data rescue initiatives: bringing historical climate data into the 21st century. *Clim. Res.* 47, 29–40. doi: 10.3354/cr00960

Capozzi, V., Cotroneo, Y., Castagno, P., Vivo, C., and Budillon, G. (2020). Rescue and quality control of sub-daily meteorological data collected at Montevergine Observatory (Southern Apennines), 1884–1963. *Earth Syst. Sci. Data* 12, 1467–1487. doi: 10.5194/essd-12-1467-2020

Chevalier, M. (2022). *crestr*: an R package to perform probabilistic climate reconstructions from palaeoecological datasets. *Clim. Past* 18, 821–844. doi: 10.5194/cp-18-821-2022

Christiansen, B., and Ljungqvist, F. (2017). Challenges and perspectives for large-scale temperature reconstructions of the past two millennia. *Rev. Geophys.* 55, 40–96. doi: 10.1002/2016rg000521

Coney, J., Pickering, B., Dufton, D., Lukach, M., Brooks, B., and Neely, R. (2022). How useful are crowdsourced air temperature observations? An assessment of Netatmo stations and quality control schemes over the United Kingdom. *Meteorol. Appl.* 29:e2075. doi: 10.1002/met.2075

Erb, M., McKay, N., Steiger, N., Dee, S., Hancock, C., Ivanovic, R., et al. (2022). Reconstructing Holocene temperatures in time and space using paleoclimate data assimilation. *Clim. Past* 18, 2599–2629. doi: 10.5194/cp-18-2599-2022

Evens, T., and Hauttekeete, L. (2011). Challenges of digital preservation for cultural heritage institutions. *J. Librariansh. Inf. Sci.* 43, 157–165. doi: 10.1177/0961000611410585

Hunziker, S., Brönnimann, S., Calle, J., Moreno, I., Andrade, M., Ticona, L., et al. (2018). Effects of undetected data quality issues on climatological analyses. *Clim. Past* 14, 1–20. doi: 10.5194/cp-14-1-2018

Hunziker, S., Gubler, S., Calle, J., Moreno, I., Andrade, M., Velarde, F., et al. (2017). Identifying, attributing, and overcoming common data quality issues of manned station observations. *Int. J. Climatol.* 37, 4131–4145. doi: 10.1002/joc.5037

Immerzeel, W., Petersen, L., Ragettli, S., and Pellicciotti, F. (2014). The importance of observed gradients of air temperature and precipitation for modeling runoff from a glacierized watershed in the Nepalese Himalayas. *Water Resour. Res.* 50, 2212–2226. doi: 10.1002/2013wr014506

Klein Tank, A. M. G., Zwiers, F. W., and Zhang, X. (2009). "Guidelines on analysis of extremes in a changing climate in support of informed decisions for adaptation" in WCDMP No. 72, WMO/TD No. 1500 (Geneva, Switzerland: World Meteorological Organization).

Luterbacher, J. (2024). The importance and scientific value of long weather and climate records; examples of historical marine data efforts across the globe. *Climate* 12:39. doi: 10.3390/cli12030039

Masenya, T., and Ngulube, P. (2019). Digital preservation practices in academic libraries in South Africa in the wake of the digital revolution. *SA J. Inf. Manag.* 21:a1011. doi: 10.4102/sajim.v21i1.1011

Mateus, C., Potito, A., and Curley, M. (2020). Engaging secondary school students in climate data rescue through service-learning partnerships. *Weather* 76, 113–118. doi: 10.1002/wea.3841

Odunayo, O., Sookoo, N., Bathla, G., Cavallin, A., Persaud, B., Szigeti, K., et al. (2021). Rescuing historical climate observations to support hydrological research. Proceedings of the 21st ACM Symposium on Document Engineering

Ośródka, K., Otop, I., and Szturc, J. (2022). Automatic quality control of telemetric rain gauge data providing quantitative quality information (RainGaugeQC). *Atmos. Meas. Tech.* 15, 5581–5597. doi: 10.5194/amt-15-5581-2022

Rahman, M. (2020). Review of digital record management needs for academic libraries. *Libr. Hi Tech News* 37, 21–22. doi: 10.1108/lhtn-11-2019-0083

Rakemane, D., and Mosweu, O. (2020). Challenges of managing and preserving audio-visual archives in archival institutions in sub Saharan Africa: a literature review. *Collect. Curation* 40, 42–50. doi: 10.1108/cc-04-2020-0011

Rasaki, O., Egbedokun, A., and Adedimeji, A. (2023). Preservation of audiovisual collections at Albert Ilemobade Library, Federal University of Technology Akure,

Ondo State, Nigeria. *Collections* 19, 69–87. doi: 10.1177/15501906231159036

Ryan, C., Duffy, C., Broderick, C., Thorne, P., Curley, M., Walsh, S., et al. (2018). Integrating data rescue into the classroom. *Bull. Am. Meteorol. Soc.* 99, 1757–1764. doi: 10.1175/bams-d-17-0147.1

Salonen, J., Korpela, M., Williams, J., and Luoto, M. (2019). Machine-learning based reconstructions of primary and secondary climate variables from North American and European fossil pollen data. *Sci. Rep.* 9:15805. doi: 10.1038/s41598-019-52293-4

Seshaba, P. (2024). Analysis of extreme temperature variability in Rwanda. *J. Atmos. Sci. Res.* 7, 74–89. doi: 10.30564/jasr.v7i1.6107

Sfîcă, L., Ichim, P, Ion, C., Baltag, Ș., and Ignat, A. (2021). Filling the gap of meteorological data along the Prut River Valley, Romania—Cârja Experimental Weather Station. Air and Water—Components of the Environment 2021 Conference Proceedings

Wilkinson, C., Woodruff, S., Brohan, P., Claesson, S., Freeman, E., Koek, F., et al. (2011). Recovery of logbooks and international marine data: the reclaim project. *Int. J. Climatol.* 31, 968–979. doi: 10.1002/joc.2102

Williamson, F., Allan, R., Switzer, A., Chan, J., Wasson, R., D'Arrigo, R., et al. (2015). New directions in hydro-climatic histories: observational data recovery, proxy records and the atmospheric circulation reconstructions over the earth (ACRE) initiative in Southeast Asia. *Geosci. Lett.* 2:2. doi: 10.1186/s40562-015-0018-z

Yaseen, A., and Abdulkareem, A. (2022). Treatment missing data of daily and monthly air temperature in some Iraqi cities by using curve fitting. *Al-Mustansiriyah J. Sci.* 33, 34–41. doi: 10.23851/mjs.v33i4.1202

Zhou, C., Azorín-Molina, C., Engström, E., Minola, L., Wern, L., Hellström, S., et al. (2022). HomogWS-se: a century-long homogenized dataset of near-surface wind speed observations since 1925 rescued in Sweden. *Earth Syst. Sci. Data* 14, 2167–2177. doi: 10.5194/essd-14-2167-2022