

# DATA RESCUE ACTIVITIES AT ICPAC (2014-2015)

# 1. INTRODUCTION

The implementation plan for the Global Framework for Climate Services (GFCS) includes large scale Data Rescue as its priority area through WMO and Regional Climate Center (RCCs) like IGAD Climate Prediction and Application Center (ICPAC), which are in close proximity with the NMHSs. A substantial amount of the climate archives that go back to the nineteenth century are on paper and has been stored under poor conditions in the national met services from which they need to be recovered, imaged and digitized before they deteriorate beyond use. ICPAC has the opportunity to work with development partners to implement some its functions, The USAID project planning for Resilience in East Africa through Policy Adaptation, Research, and Economic Development (PREPARED) is funded by USAID is a five-year, multi-organization, comprehensive program aimed at mainstreaming climate-resilient development planning and program implementation into the East African Community (EAC). PREPARED has been working closely with the East Africa Community (EAC) Partner States, (ICPAC), and the World Meteorological Organization (WMO) on strengthening climate information generation in the region. Directors of Meteorological and Hydrological Services (DMHS) from each EAC Partner State requested the PREPARED Project's assistance in rescuing and digitizing historical meteorological data that will eventually be integrated into the GeoCLIM datasets to improve products generation. This initiative fits in the objective of Global Framework for Climate Services (GFCS) for strengthening climate information generation where Data Rescue (DARE) is number 3 on WMO priority list.

# 2. ICPAC DATA RESCUE BACKGROUND

According to the needs assessment done by the DARE team in December 2014, Data Rescue was to start in Burundi but due to political turmoil, it was not possible to implement data rescue in that country. Directors of Meteorological Services in the EAC Partner States meeting held in Arusha 28-30 April 2015, recommended implementing a DARE Pilot Project in Tanzania as an alternative with ICPAC technical support. It was hoped that, the DARE pilot would provide additional data to fill gaps in the missing records, and improve access to longer-term data series that in turn will advance service delivery in the provision of climate information. The DARE also has the potential to expand the data coverage and inputs into GeoCLIM. a spatial visualization software. The objective of the ICPAC\PREPARED Project DARE initiative is to support each EAC Partner State in establishing sustainable operational systems for gathering data; rescuing, preserving, digitizing, archiving and improving data quality; analyzing records; and disseminating the information to the public. ICPAC in collaboration and support of the PREPARED Project planned implementing the Piloting of Data Rescue activities at Tanzania Meteorological Agency (TMA) and at the same time build capacity for the Data Managers (DMs) in the EAC region in basics of Climate Data Management. It is anticipated that capturing and cataloging historical data in Tanzania will demonstrate best practices for DARE that can be replicated in the other National Meteorological and Hydrological Services (NMHS) within the EA.

## 3. NEEDS ASSESSMENT

A needs assessment was carried out in each of the 5 EAC countries from Nov to Dec 2014 under the support of USAID. The purpose of which was to establish the status of the records, the donor support on the ground and also to get a buy in and commitment by the top management of the NMHSs, can click here for full report.

#### 3.1 Other donor support available

This assessment helped to identify the kind of donor support available in each NMHSs as shown in the table below:

No	Donor Support	Country	Items
1		Kenya	30 Pcs ,30 Data clerks
	Bill Clinton Foundation (2014-2015)		and desks
2		Africa Climate Center	Imaging Upper Air data
	IEDRO NOAA USA supported (2005)		
3		Uganda	Non-Inflammable
	Norwegian Govt. NORAAD (2015)		Archive and imaging
4		Rwanda	
	UNECA DARE Support (2014-15)		Data digitization
5		Uganda	10 computers and one
	GIZ is support (2012- 2014)		server
6		Kenya	Data Entry clerks
	Kenya Government supports DARE On going		digitizing data
7		Tanzania	Record inventory,
	DFID DARE support (2014-16)		Imaging & Digitization
8	Tanzania Government supports DARE	Tanzania	General Data
	activities (Continuous)		management
9		Tanzania	10 PCs, Data clerks,
	USAID- ICPAC - PREPARED		Imaging , digitisation of
			LVB stations

#### Table:1

## 4. INSTITION INVOLVED IN DARE PILOT

Following are the institutions involved in the data rescue exercise and their roles:

**ICPAC:** Given its institutional mandate as an RCC (Regional Climate Center), ICPAC is playing the implementation role of providing technical support.

<u>WMO:</u> through its regional office, (RAI) is doing the oversight role to ensure that parity of the institutions involved exists and also DARE standards and quality are adhered to.

**PREPARED:** facilitates the DARE pilot activities with financial support for Data Rescue and information generation.

**TMA:** Routine supervision of the DARE-Pilot activities such as imaging, indexing, quality control and climate database management to increase on usable records.

## 5. IMPLEMENTATION

TMA had agreed to be the proverbial guinea pig of DARE in the EAC region, everything emanating from the TMA DARE pilot would be an experience for the betterment of practices in the other NMHSs of the region. The DARE team comprises of Ms Lubega Fortunata (ICPAC), Mr. Joseph Kimani (WMO ET DARE), Dr Indenje Matayo (PREPARED), Mr. Evans Mukolwe as an Oversight Consultant and Ms Janet Loningo, (TMA data manager). This team was charged with the responsibility of implementing this TMA DARE pilot exercise. It should be clearly noted that those who are engaged in the Data Rescue venture, should bear in mind that:

## Data Rescue is a slow pain staking activity, requiring a lot of



Following below are a series of activities as they were being implemented under the TMA DARE pilot Project:

- The consultative meeting with the TMA top management to chart out the DARE pilot project approach;
- The TMA management requested that they be trained in the DARE techniques and be in position to do the work themselves rather than outsourcing outside companies,
- The stations inventory and gaps identification was done as a next step to demarcate areas of interest (Fig 1a) below;

- An MOU between PREPARED, ICPAC and TMA was drafted, agreed upon with TMA harmonizing their input before the DARE could kick off.
- Procurement of the equipment (computers and other accessories) supported by PREPARED have been accomplished and ICPAC donated Server and 3 work station were delivered to TMA as part of DARE tools.
- Capacity building in DARE Concepts, and climate data management techniques, including available use of CDMSs in the region (CLIDATA and CLIMSOFT, GeoCLIM) was done in the training workshop.
- Data Clerks theoretical training was intended to teach them their work ethics, the key role they were playing in the TMA pilot exercise and in the global climate change issues,
- Data Clerks were taken through the key tasks of Data rescue namely: sorting the paper record, Imaging Indexing the images digitising records using CLIDATA, proof reading.

Actual implementation of the key TMA DARE pilot activities are explain in details below:

#### 5.1 Station inventory

In order to establish synergy and avoid duplication of work, this exercised started by liaising with DFID, from UK MET Office who were doing data Rescue in Tanzania around the same time. From the discussion, it was established that DFID's area of interest was on the coast and central Tanzania. DFID also supported TMA in records sorting and archive inventory as part of data rescue (**Fig: 1b**). While records are well sorted and packed neatly in boxes at TMA archive, they are not all imaged or digitized in usable formats. This was followed by stations inventory and gaps analysis together with an agreement on the selection of 257 rainfall stations to be rescued from the Lake Victoria Basin (LVB) on the Tanzanian side as can be seen in **Fig:1a** below. These involved 3 synoptic station, 18 climatological, 6 Hydromet and 230 rainfall stations for elements of Rainfall, temperature maximum and minimum.



Fig1a: 257 Rainfall and 18 temperature stations



Fig1b: Sorted and stacked records to be rescued

#### **5.2 Procurement of DARE tools**

The 3rd visit to Tanzania was centered on procurement of DARE equipment and tools, installation and networking of facilities prior to the regional training for the Data Managers (DMs).

> PREPARED Operations Manager, Ms. Nancy Ngugi, had joined the team to guide on the procurement process.

- The idea that procurement was done locally within Tanzania (in consultation with the local staff) rather than having to source the DARE equipment from out of the country was highly commended by the TMA top management.
- ICPAC provided 1 Server with Server OS, 9KVA UPS and 2 Workstation for data rescue at TMA. The computers were cleared through customs and securely arrived at TMA on 12th October, 2015.
- > The items which were not locally available were outsourced from Nairobi Kenya.
- > The Computers were installed and networked at the TMA Lab as it can be seen in the **Fig2a** below.

No	Item	Qty
1	Desktop Computers	10
2	A3 scanners	2
3	MS Office 2010; professional	1
4	ICPAC serve + 9KVAUPS	1
5	Digital Cameras	2
6	Networking cards and cabling wires	
7	Desks	10
8	Diskette Reader	2
	Consumables; (dust coats, gloves, masks, CDs Flush disks,	
9	stationery, etc	

#### 3.3 Dare Tools acquired

Table: 2

5.3 Training Workshop for DMs

Training of Climate Data Managers (DMs) is considered one of the key activity because they are charged with the full responsibility of data in their services and should have the basic knowledge required in the subject matter. The method of teaching this course to them was situational, such that it can be adopted as part of the best practices for data rescue and management. As different managers are better versed with some aspects of the subjects than the others and also by virtue of the fact that some managers are better facilitated by their services to carry out some aspects of DARE than the others, they will be asked to share such information with the other participants so as to bring out good practices in the areas of concern. The course was mainly theoretical apart from the climate data quality control which was a hands on exercise, using their own dataset. This DARE training that was carried out from November 16th-20th, 2015 provided participants with the knowledge on the importance of data rescue concepts, climate data management, tools and methodologies of using CDMS such as CLIDATA and CLIMSOFT.

They were also trained on generating products because they are expected to serve the user needs. The training material and course program can be obtained <u>here</u>. Following were the key objective of this training, enabling DMs:

- To supervise the DARE Activities in their own NMHS with clear understanding of each of the procedures;
- To rescue climatological data from paper records at risk of loss and from obsolete electronic media;
- To Identify and acquire the necessary data readers for the case of electronic media;
- To understand key concepts of Climate Data management and gain operational skills expected at their level;
- To Use the Image Indexing tool for cataloguing the imaged paper for easy retrieval;
- To Ensure migration of electronic records to match changing technologies;
- To participate in and Supervise Climate Data quality Control at all levels;
- To know and guide the QC criteria pertaining to their country NMHS;
- To gain knowledge of available CDMS such that they can advise and guide their NMHS on which ones are suitable for their services;
- To have an insight of WMO guidelines on Climate data management and sharing;



Fig2a: Newly purchased computers installed at TMA



Fig2b: EAC DMs Workshop participants

#### 5.4 Actual implementation of the DARE

The actual implementation of DARE is a core activity transforming paper manuscripts into digital usablerecords and electronic images at TMA. This DARE seeding activity involved the following importantfunctions:DigitizationImagingIndexingQuality controlArchiving

Data Entry Clerks were first given a theoretical training on the importance of the work they are doing and the key role they were to play in the Data Rescue. This training is different from that of the DMs as it focus on actual hands data rescue work. In order to ensure that the Data Rescue activities are implanted successfully, in the second week, the DARE team had to test if all the required tools\software are up and running in all the computers. The DARE team, therefore walked the Data clerks through each of the exercise and resolved all hitches encountered. The **Fig3a** shows the careful movement of the records from and back into the TMA archives.



## Fig:3a

Fig3b below, shows the data clerks taking pictures of the rainfall cards using the 2 mounted cameras shown by the arrow. It was deceided to image all available paper records of the rainfall cards from 1961 to 2015. Fig 3c shows one of the data clerks digitising data and also indexing the photogrphed images.







## **6. TECHNICAL ISSUES**

The ICPAC provided server was targeted to keep the images captured in the Data Rescue at TMA;

TMA IT personnel to finish the connectivity for the data entry personnel to be. This should include setting up of a user account for each key-entry staff.

One of the ICPAC provided workstation to act as an intermediary Client Server for purpose of backstopping and quality controlling images before finally archiving them into the server.

The catalogue of Lake Victoria Basin stations IDs to be updated to have 8 digits to avoid mismatch with the ones in the CLIDATA at TMA (should be done by experienced Supervisor).

On the Quality Control limit issue, where global limits were set into the CLIDATA database instead of the local limits, allowing errors to pass undetected; it was recommended to be handled at a higher level.

As part of the understanding that easy to use and locally available cheap DARE tools, the DARE team developed an Image Indexing tool which was being tested at the TMA DARE project;

The images are indexed using the Image Indexing tool storing the file location paths rather than the image itself and this reduces on computer space.

All the PCs using wireless networking could connect to the CLIDATA server but had a problem communicating with the CLIDATA CDMS itself. The TMA IT officer, Victor Massam, to work on this later;

Images folders uploaded to individual PCs should act as a backup and ensure that they are not removed to disturb the link until the next visit by the DARE Team;

It was agreed that 2 experienced staff be selected to proof read or validate the data before loading it into CLIDATA to ensure errors introduced at key entry don't sail through to the final database;

## 7. CONCLUSION

The main objective of the DARE Pilot with for Tanzania Meteorological Agency (TMA) is to draw practical lessons from implementing DARE that will inform feasibility of up-scaling at Regional level in an efficient and cost effective manner driving, analysis and dissemination. The desired outcome is the generation of easy to access usable quality controlled climate records and the electronic imaged archives. It is hoped that the Tanzania DARE pilot project will provide a learning experience which can be replicated or avoided in other areas. The project is now at the monitoring phase and the updates to include the results of monitoring phase will be proided.