

# **Executive Summary**

Belize's HydrometDB is a Climatic Database Management System (CDMS) that allows easy integration of multiple sources of automatic and manual stations, data quality control procedures, and provides powerful visualisation and report tools.

The system with the requirements of the Scope of Reference (SoR) in terms of platforms, the use of open-source software and standards, can be easily extended by users through the development of plug-ins since it has Application Program Interface (API) that allows access to the data for 3 rd party or further developments.

The system provide an easy way to capture data, it provides a powerful generator of capture and visualization forms that gives the user unlimited possibilities to capture data and provides an easy to use tool to set the order of the capture based on the source paper forms. The mapping features of HydrometDB not only let the user locate the stations but can also show and map different elements from the database on comprehensive maps with a few clicks. There is also a charting and reporting tool to present and extract data from the system for easy analysis.

### **Table of Contents**

**Executive Summary** 

Table of Contents

Project Background

<u>Methodology</u>

Modules Developed

<u>Technology Architecture</u>

Station Metadata Management Module

**Data Assimilation Module** 

Data Entry Module

Quality Control Module

Data Inventory Module

Reporting Module

Charts and Mapping Module

**CLICOM's Data Migration** 

CDMS Checklist based on SoR

Glossary of Terms



# **Project Background**

<u>Project</u> <u>Information:</u> National Meteorological Service of Belize, Climate Data Management System.

Over the past 20 years the National Meteorological Service of Belize has been using the World Climate Data and Monitoring Program (WCDMP) Climate Computing data management software, CLICOM as its main tool for storing and analyzing climate data.

Since the implementation of this software there have been great improvements in database technologies such as relational databases, query languages and web applications, links with Geographical Information Systems (GIS) and more efficient data capture methods through the increased use of automatic weather station and internet services.

With these and other advancements in technology the National Meteorological Service of Belize desires to upgrade its current database software with a web based application, whose role will be to interact with an open source database, to input and output data with the help of an open source web server so that the National Meteorological Service can more effectively collect, analyze and output data to its staff and clients.

The reasons for the NMS choosing web application software are many. This type of software is advantageous because of the ubiquity of web browsers, and the convenience of using a web browser as a end user client. They have the ability to be updated and maintained without distributing and installing software on multiple client computers and supports cross-platform compatibility. There are also a number of open source solutions for DBMS and vast resources of online support when troubleshooting becomes necessary or when the need arises to improve/expand the system.

The need to improve the current database management system in Belize was further magnified in October of 2011 when the export functionality of CLICOM ceased operating and online support for the software was virtually non existent.

This document seeks to outline the development of HydrometDB exposing the



implementation of this new system in The National Meteorological Service of Belize.

## Methodology

Based on the Scope of Reference, and the requirements exposed, we proposed an <u>Incremental methodology</u>, where the project execution is approached via incremental functional prototypes (versions of functional software pieces based on requirements), after an Initial Requirements Meeting to define the baseline.

Based on this development methodology we presented during the development process incremental prototypes of the features and tasks been developed and milestones being achieved.

### **Modules Developed**

The Scope of reference (SoR) exposed the need for a new CDMS and explain the observations network layout, the different source of data, and the different frequency of data capture. I shows examples of data records for Daily, and Hourly data entered manually. And Hourly and less that hour data from AWS (Automatic Weather Stations),

The SoR also propose a checklist of features the system should have about specific topics regarding data interfaces, technology stack, security, disaster recovery, data types and data management.

Based on this requirements we proposed a systems that comply with that criteria in terms of management of data and in terms of technology architecture. Also identified 5 main modules for the system, that we are presenting as part of the modules developed.

### **Technology Architecture**

- Open Source Platform
- LAPP technology stack
- Linux with Apache as Application server
- PostgreSQL as database management system
- PHP as server side scripting.
- Client-side (web browser) scripting will be Javascript, HTML, CSS.
- All end user access to data will be via a web browser.



The system is optimized to be used with Firefox and Chrome web browsers.

### **Station Metadata Management Module**

This module provides all the functionality required to effectively manage the station network metadata following the best practices recommended by WMO.

Besides the usual metadata such as identification codes and geographical location and characteristics, data acquisition and transmission, provides features to enable the management of all information about the station and its instruments and elements history.

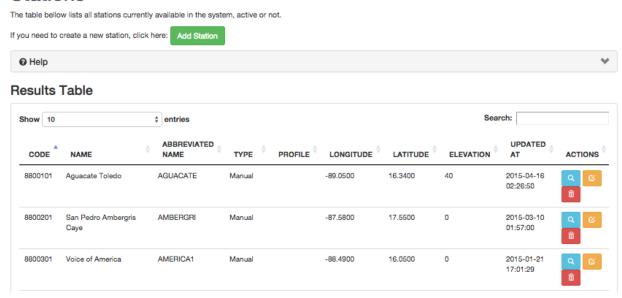
The Station Metadata Management Module also cover the storage of reports, documents, technical documentation, users manuals, maps and other digital media that can be related to the station like a station's File System.

# Settings 79 WEATHER STATIONS View details

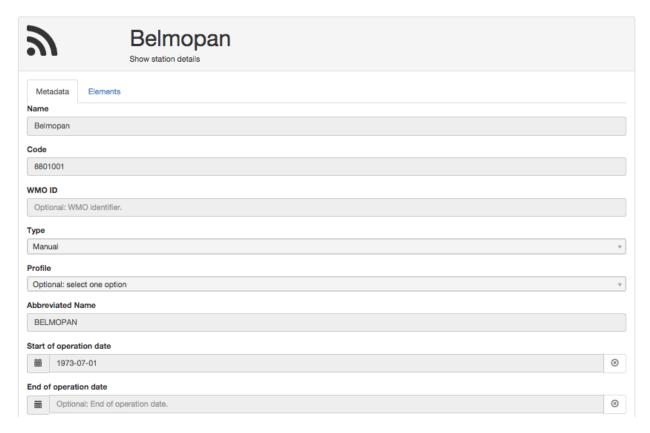
The Station Metadata Management Module is accessible from the Settings Screen, it shows the number of stations in the System and give you access to the management of stations to add a new station, set variables or elements to an station, define ranges of climatic values accepted by the stations, instrumentation associated to the station and media files, like map, charts, photos, or other documents relevant to each station, having an Electronic File System for the station.



# **Stations**

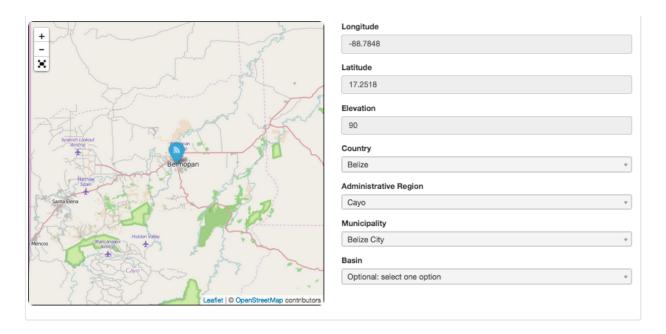


After accessing the module, the system shows the list of stations, where user can search, list, select or delete an station.

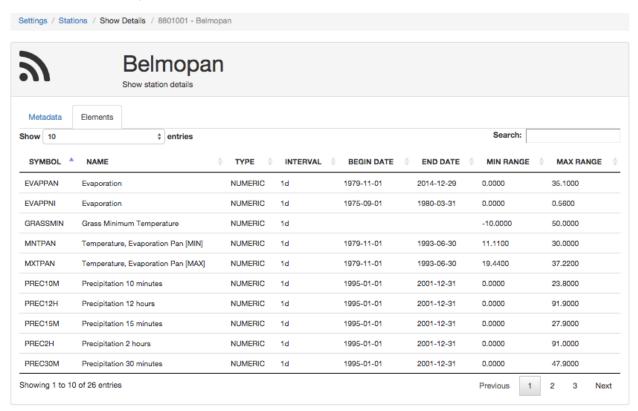


Stations Metadata Sample





### Stations Metadata, related to location.



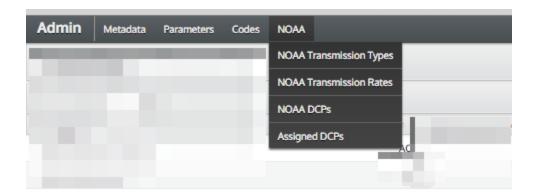
From the Station's metadata, users can manage and associate related elements, climatic variables, frequency of capture, range of dates, and range of values to do proper quality control upon data entry.

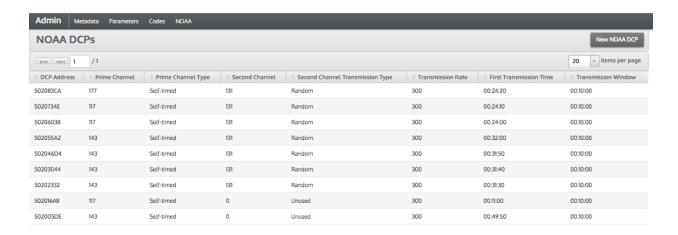


### **Data Assimilation Module**

The data assimilation module processes different sources of data from different collecting equipments and transmission technologies, this data is then transferred into the database to be queried and sent to reports.

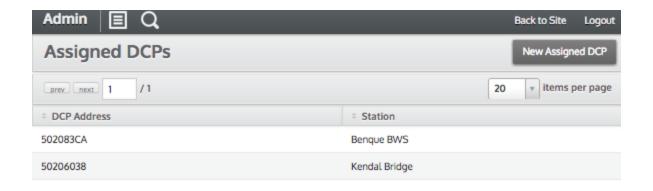
The system provides a section for all parametrization of values to be able to ingest the data from different sources



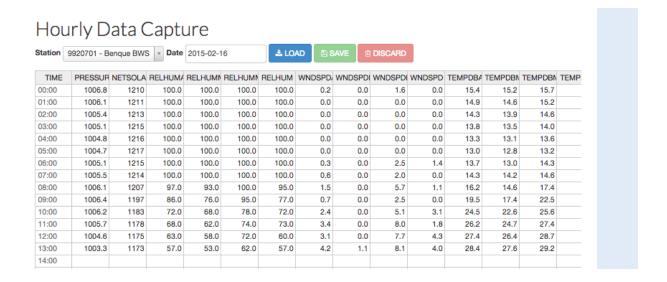




### NOAA DCP's



Assignment of DCP's to Stations in order to download data from Satellite AWS's



NESA Station. Transmission done via Satellite ingested to the system.

### **Data Entry Module**

This module process Data entry (manual and batch mode) to be able to add, select, update or delete daily and hourly climatic data.



For the Data Entry the initial approach and design based on current data entry form as WMO Guidelines where modified and enhanced. A custom form generator was build that enable the user to define its own capture forms in terms of layout, stations that will be used in the form, elements to be captured, frequency and range of dates and custom order of fields and stations.

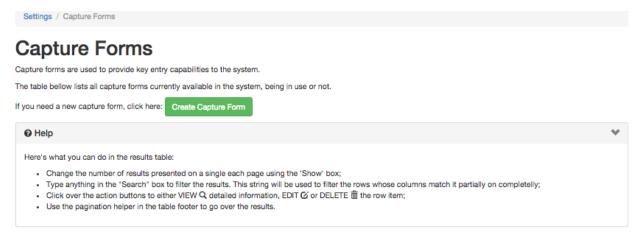
This new approach enables the user to have unlimited combinations of capture forms, so the system can adapt to new forms, or emulate current paper forms.

This module integrates the quality control validation in real time when capturing data, it enforces validation of data based on range, step, and persistence validation. To create and configure the data captures forms the system provides a module name Forms, that is accessible by administrator of the system in the Settings.



Once entering the module, it enables the user to define the forms, if any forms are already created it presents the current forms enabled in the system.

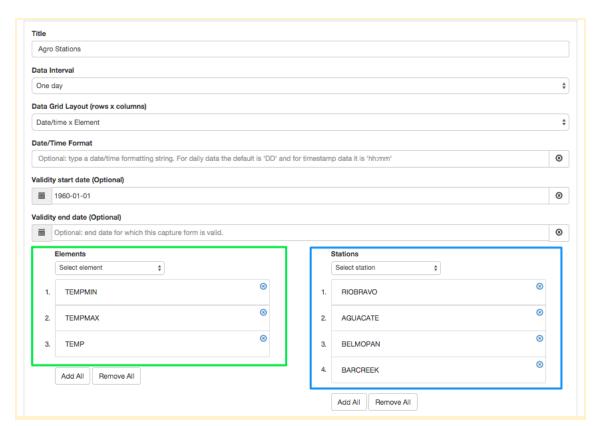




### Results Table



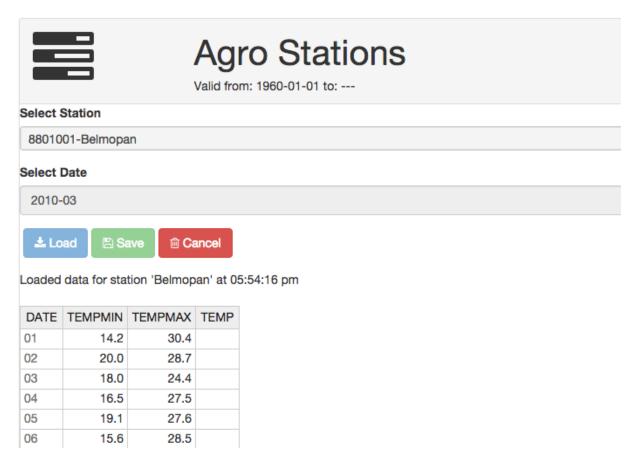
# List of capture forms



The Data forms module defines the way the user can capture, edit, update or delete

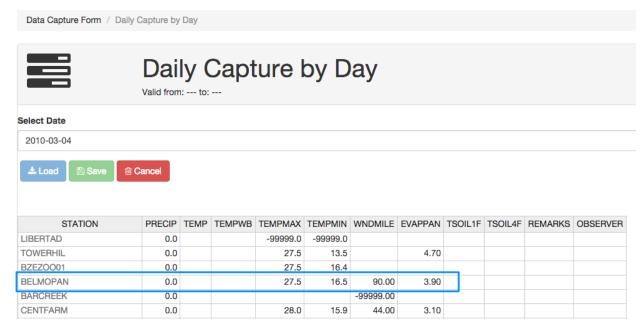


data. It can also let the user view data from certain stations selection the elements, and the stations.



Sample layout capture by month, daily data.





Sample layout of same data from above viewed by day (March 3, 2010)

The Capture module also validates and does quaility control check directly on data input for verification purposes, advicing the user about possible error on capture, minimizing data capture errors.

Even the inconsistency is detected the systems can accept the error flagging the inconsistent data for post-verification purposes.

Date: 2010-03-04, Station: Belmopan, Element: Precipitation- Enter a value between 0.0 and 334.0

STATION	PRECIP	TEMP	TEMPWB	TEMPMAX	TEMPMIN	WNDMILE	EVAPPAN
LIBERTAD	0.0			-99999.0	-99999.0		
TOWERHIL	0.0			27.5	13.5		4.70
BZEZOO01	0.0			27.5	16.4		
BELMOPAN	0.0			27.5	16.5	90.00	3.90

Sample: System informs user about the acceptable value for the station.



Date: 2010-03-04, Station: Melinda Forest Station, Element: Air Temperature - Dry Bulb [MAX]- Enter a value between 18.9 and 38.5 Errors:

Range: value 17.0 is below 18.9

STATION	PRECIP	TEMP	TEMPWB	TEMPMAX	TEMPMIN	WNDMILE	EVAPPAN	TSOIL1F	TSOIL4F	REMA
LIBERTAD	0.0			-99999.0	-99999.0					
TOWERHIL	0.0			27.5	13.5		4.70			
BZEZOO01	0.0			27.5	16.4					
BELMOPAN	0.0			27.5	16.5	90.00	3.90			
BARCREEK	0.0					-99999.00				
CENTFARM	0.0			28.0	15.9	44.00	3.10			
CHACREEK	0.0			27.7	16.7					
BALDYBCN	0.0			19.0	18.8					
HERSHEY1	0.0			26.5		12.00				
POMONA01	0.0			27.5	17.0					
MIDDLESE	0.0									
MELINDA1	0.9			17.0	16.4	3.00	3.50			
14144141416						- 00				

Sample: The system detects and value out of range and inform the user about the error.

Date: 2010-03-04, Station: Central Farm, Element: Air Temperature - Dry Bulb [MIN]- Enter a value between 7.8 and 27.1

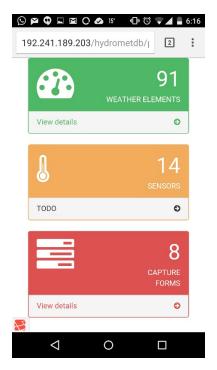
STATION	PRECIP	TEMP	TEMPWB	TEMPMAX	TEMPMIN	WNDMILE	EVAPPAN	TSOIL1F	TSC
LIBERTAD	0.0			-99999.0	-99999.0				
TOWERHIL	0.0			27.5	13.5		4.70		
BZEZOO01	0.0			27.5	16.4				
BELMOPAN	0.0			27.5	16.5	90.00	3.90		
BARCREEK	0.0					-99999.00			
CENTFARM	0.0			15.0	15.9	44.00	3.10		

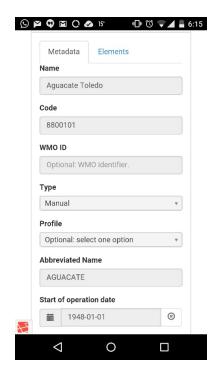
Sample: The data entry module can detect inconsistencies between variables and inform the user accordingly

This module calculate automatically based in data insertion the different aggregated and derived variables applicable for daily, monthly, yearly and decadal data calculation.

The Capture Forms follows a responsive design approach to adapt the content and data entry optimized for mobile **web** to encourage the use of mobile technology to capture data from an smartphone for the daily reports. The Form management allow creation of single station and basic elements to allow basic data capture.







### **Quality Control Module**

All data ingested by the system, either manually or automatically, is validated by the system and assigned quality control flags.

Data validation and quality control procedures are fully parameterized (from the elements associated to the station). The parameters and validation tests are be defined per element as exposed in the station and data capture module.



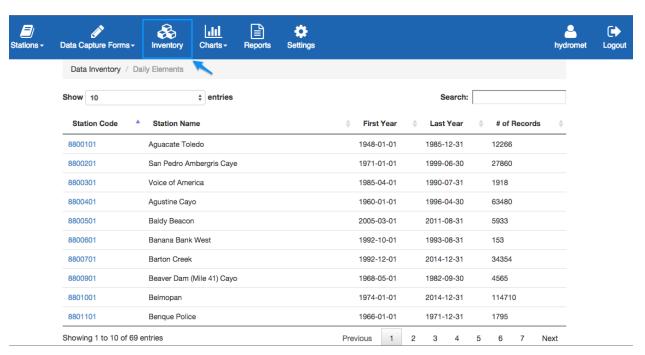
### **Data Inventory Module**

This module allow users to analyze data availability on a station and variable basis.

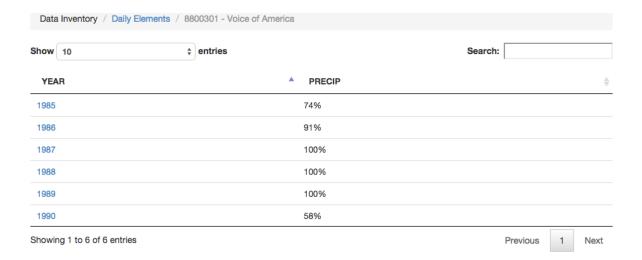
It is composed of special reports and charts that allow data users to quickly check and



assess data availability conditions per station or variable by month and year.

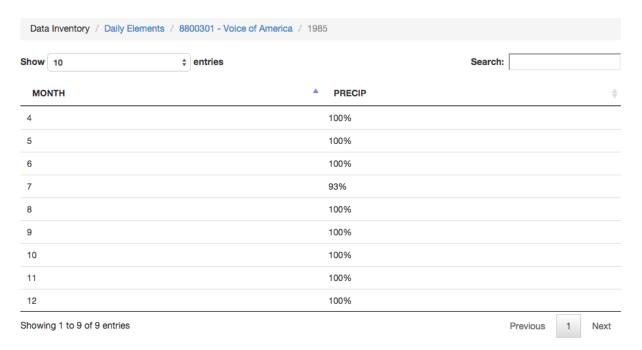


After accessing the Inventory module, by default shows the amount of data available per station, first and last date of data.



The system allows to drill down inside the inventory to determine in more detail the completeness of the data.





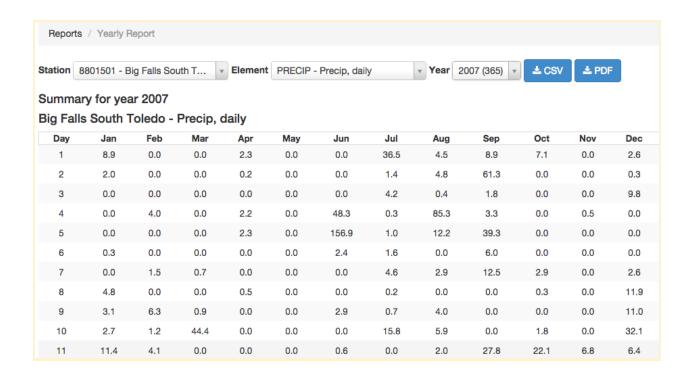
Up to the level to determine exactly where data is missing in this case 1985 for this station for precipitation is missing on July, and from January to March. That is why the completeness is 74%.

### **Reporting Module**

The reporting module have the options to generate tabular report-style data, some reports includes, Variable per station, daily, monthly and yearly, most reports are based on a range of date and a station, and have CSV (Excel Compatible), HTML and PDF outputs.





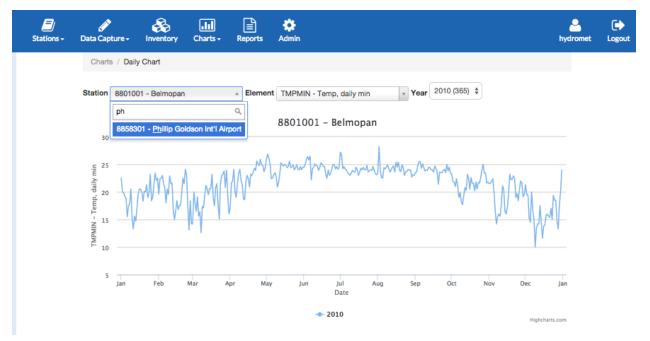


### **Charts and Mapping Module**



This module present charts based on data in different ranges, timeframes and variables, charting data compare visually different year of a same station per variable or different stations of a same variable for an specific year among other data analysis charts like monthly and yearly. Also be able to generate time series.



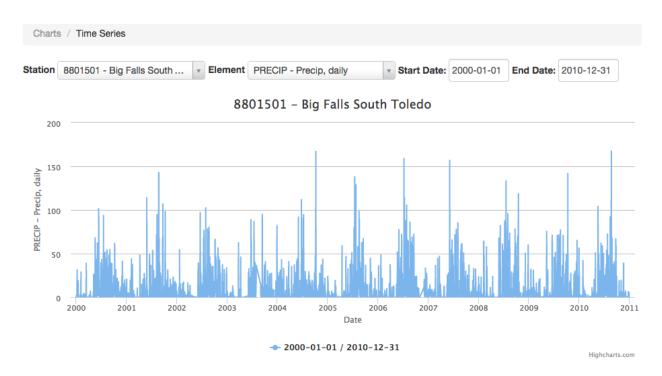


Example: Full text search by ID or name in Charts, reports and Data capture, let user search by ID or name in same control. Yearly chart of TempMin.





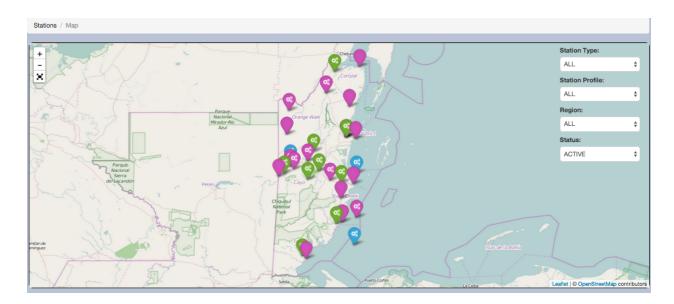
Sample: Monthly chart calculated in realtime of 15 years of data, showing Min. Temperature climatology



A time series chart let the user analyse the behavior of a climatic element in a long



time, be able to determine if there are periods of time where data is missing, or maybe should be subject to quality control for example.



A map representing the location of Belize's Stations that can be selected by type or profile.

The mapping module is 100% web based using a common access online mapping service like Google Maps, OpenStreetMaps. All geospatial representation is done online to present Belize's CDMS data in form of layers of data geospatial interpolation overlays and point data layers.

### **CLICOM's Data Migration**

On the Initial visit to NMS of Belize, we received a plain text with all the information that was able to be rescued from CLICOM. This data file was used in combination with custom made programs and processes to extract the data and placed in an structured and relational way.



# **CDMS Checklist based on SoR**

Process	Level of Importance	Developer Response	Description of Compliance
Conform to WMO(appendix ii)	М	С	We comply with the requirement of flagging suspicious data based on data validation on common climatic validation criteria. To be able to be revised, the system provides user access level based on groups of users.
CDMS written in PHP	М	С	The system is developed on PHP 5.3+ using Laravel Framework
Source Code and Data Dictionary	М	С	All source code will be installed on server, will not use any code encoders, will provide technical documentation and data dictionary definition
Installation and User Documentation	М	С	The system will be implemented, full installed and operational and a user manual will be provided
Ingest of all data formats	М	С	The Systemo already provide the processes to ingest the format described in this document.
CRUD Station, parameters, units etc	М	С	The application has all backend processes and modules to do CRUD of data.
Station Meta Data	М	С	Stations Metada is implemented.
Maintenance/Instrume nt Meta Data	М	С	Instruments and Maintenance Metadata is related to stations
Facilitate Multiple reading	М	С	The system is a multiuser network based system. Multiple users can be accessing data.
Automatic Calculations (Page 15, x,xi)	М	С	The system calculates aggregated data, monthly, yearly, and other calculated fields.
Production of chart	М	С	Charts are generated as show in this document.
Spatial Distribution Maps	D	С	Data visualization is implemented as map leyers using open source Javascript mapping



			libraries.
Generation of reports (page 15, xii)	М	С	The system has a reporting module exporting to HTML , PDF and CSV (Excel compatible) formats.
Data Validation	М	С	The system validates data on manual key entry and on ingest processes.
			We comply with the requirement of flagging suspicious data based on data validation on common climatic validation criteria.
Output into BDCAC	D	С	The system has an option to generate the file to be ingested to BDCAC system in the required format
Network Configuration	М	С	In the implementation process the system will be installed in the servers and configured to be available to the networks users.
System Security (Application, Database and Web Servers)	М	С	The system has user level validation security access, also all the configurations will be done to Web and database servers to have user validated access of root for server and databases. Proper database backup training is done to ensure data is secured and protected.
Hardware Redundancy	М	С	In the technical training we provide the guidelines and desired configurations for enabling data replication and automatic backup (to cloud services for example) to lower downtime in case of failure
Certificate of registration for webserver	М	С	We can provide a domain name in accordance with Belize's Met Service for registration of name for 2 years after implementation
99.998% Accuracy of Data retrieved	М	С	Based on consistency source data files the process should ingest data accurately. The data has been migrated and processed.
Network Configuration, Installation, User	М	С	Upon completion of support for the project will have technical workshops covering network configuration, data structure, technology



Training			stacks, server architecture. Also a User Training workshop will be done for end users.
Database recovery after failures from Back files	М	С	The system does scheduled automatic backups (on site and offsite / other server /cloud server). We also provide an offsite development server to have a copy of the system and implement scheduled backups to this service.
New data format integration by CDMS user as described in section 5.0 under Revision/Update	D	P	The system propose an standard format of ingestion of data to CDMS based on the current data model, with this schema and scope. The system currently supports hourly, (less that a day), daily and monthly structures. Basic data field mapping is implemented to move data from AWS's to the system.



# **Glossary of Terms**

API - An abbreviation of application program interface, is a set of routines, protocols, and tools for building software applications. The API specifies how software components should interact and are used when programming graphical user interface (GUI) components.

BDCAC- Base de Datos Climáticos de América Central, Central American Climatic Database management System, managed by CRRH - SICA

CDMS - Climatic Database Management System

NMS - Refers to National Meteorological Service of Belize

SoR - Scope of Reference, the terms and requirements presented by NMS of Belize as baseline for the development of HydrometDB

HydrometDB - Refers to the name of the Climatic Database management system developed.