NOAA Technical Report NOS CO-OPS

NWLON/DMS Data Replacement Criteria: Functional Requirements Document

Joann Nault
Information Systems Division

Silver Spring, Maryland
August 2001

noaa **National Oceanic And Atmospheric Administration**

U.S. DEPARTMENT OF COMMERCE
National Ocean Service
Center for Operational Oceanographic Products and Services (CO-OPS)

TABLE OF CONTENTS

1.0 General Information. 3
- 1.1 Summary 3
- 1.2 Environment 3

2.0 Scope and Overview 4
- 2.1 Background 4
- 2.2 Summary of Current System Deficiencies 4
- 2.3 Resource Impacts 5

3.0 Requirements 5
- 3.1 Replacement Hierarchy Architecture 5
- 3.2 Replacement Criteria 6
- 3.3 Performance 7
- 3.4 Error and Recovery 7
- 3.5 Testing 7

4.0 Operating Environment 8
- 4.1 Hardware Requirements 8
- 4.2 Support Software 8
- 4.3 Security 8

Appendix A - Replacement Requirements Team. A.1
Appendix B - NWLON DMS Modifications. B.1

Glossary C.1

1.0 General Information

The National Ocean Service (NOS) operates and maintains the National Water Level Observation Network (NWLON) to accomplish its mission requirement to collect and disseminate water level data. A key component of the NWLON is data ingestion into a central data repository. All incoming data currently is stored as first come first stored, which means that if the stored data is of poor quality and incoming is better, it is not replaced. A team was established to determine the criteria for data replacement. Since the development of Physical Oceanographic Real-Time Systems (PORTS), there is now a need to change the first come first stored policy. Participants can be found in Appendix A.

This document defines the general functional requirements for the data replacement hierarchy and the criteria for the ingestion of all data sources into the NWLON Database Management System (DMS).

The remainder of this section contains a brief summary of the flow of data from an NWLON water level gauge to CO-OPS. Section 2 contains a brief summary of the current policy for data ingestion. The functional requirements of the NWLON data replacement hierarchy and criteria for such are outlined in Section 3. The operating environment is described in Section 4, and includes equipment, support software, security and privacy, and operational control requirements.

1.1 Summary

This document contains the functional requirements for the application software associated with the quality control process executed during the ingestion of data into Operational Platform for Acquistion of Water Levels (OPAWL) and the process for placing data into the NWLON DMS. The task of preparing these requirements is the responsibility of the Information Systems Division (ISD), with guidance from the Products and Services Division (PSD) and the Requirements and Development Division (RDD).

1.2 Environment

- NOS responsibilities for collecting, processing, and analyzing tidal and Great Lakes data are held by the Center for Operational Oceanographic Products and Services (CO-OPS) located in Silver Spring, MD. The Center directs the establishment, operation, and maintenance of the NWLON, which provides oceanographic data and produces tide and current predictions. These products are for direct use in marine navigation, civil engineering, environmental protection, and the correlation and analysis in scientific research.

- The NWLON network for collecting, analyzing, and disseminating water level data and associated hardware is operated and maintained by ISD. Primarily, ISD collects water level gauge data via the Internet from the National Weather Service's (NWS),

Office of System Operations (OSO). OSO receives its data directly from Wallops Island, Virginia. The NWLON water level gauges have transmitters, which transmit via a GOES data collection system to a receiver at Wallops. ISD also collects water level gauge data via telnet to Wallops, direct phone calls to gauges, Mitroned Analog to Digital Recorder (ADR), and diskettes from CO-OPS system personnel.

- The PORTS network for collecting, analyzing, and disseminating water level data and associated hardware is operated and maintained by ISD. Primarily, ISD collects water level gauge data via a PC-based Data Acquisition System (DAS), installed at each of the local PORTS sites.

2.0 Scope and Overview

This document will limit the scope of the requirements to only the applications associated with data replacement. Gauge transmission and decoding requirements will not be addressed. This document is for the data replacement requirements only.

This section provides background information on the existing CO-OPS NWLON data ingestion process and the purpose and scope of the proposed replacement process. Staffing requirements for the proposed software and hardware are discussed in Section 2.3.

2.1 Background

The CO-OPS NWLON ingests water level data from four different types of gauges as well as ancillary data (wind, air temperature, barometric pressure, etc.). The five different types of gauges are Acoustic, Shaft Angle Encoder (SAE), Single Pressure, Dual Pressure, and Ancillary. This data is transmitted via GOES data collection system either every hour or every three hours and then processed, quality controlled, and ingested. Currently PORTS data is ingested into the PORTS DMS every six minutes. In order to make the PORTS data available through the NWLON DMS, ingestion software needs to be modified.

2.2 Summary of Current System Deficiencies

The current NWLON DMS ingestion process does not allow for multiple sources of data for Satellite and Redundant. Diskette and Phone Call data are used to fill gaps where no data exists for the Satellite and/or Redundant. Storm Surge and Tsunami are stored in separate database tables. The ingested data is not tested for the validity of the same data from a different source. For instance, storm surge data, since transmitted over a random GOES satellite channel, tends to have a lot of parity errors which make this data bad. With a replacement hierarchy and applied criteria, a bad value would be replaced with more reliable data.

The current incoming data source field values [] are as follows:

[A] Primary data from ADR
[D] Primary data from Diskette (Manual Phone Call, or from Data Collection Platform(DCP))
[P] Primary data Phone Call from DCP calls
[Q] QC Spreadsheet (other)
[R] Redundant water level data from Satellite
[S] Primary data from Satellite (NWS and/or Wallops)
[T] Tsunami data from Satellite
[X] Storm Surge data from Satellite
[Z] PORTS data from DAS

2.3 Resource Impacts

- All ancillary tables will need to be modified to include the DATA_SOURCE field.

- New software will be written in Structured Query Language (SQL).

- New software will be run in parallel with operational software for four weeks to verify results of new software. A test plan will be written and test actually conducted before any application is placed in production.

- It is anticipated that 240 staff hours will be required for development, testing, and documentation. CO-OPS is short of qualified analysts and this will impact other tasks.

- The operational database will need to be disabled to all users and ingestion during conversion. This conversion can take up to two weeks to complete.

- No additional costs associated with capital expenditures are anticipated.

3.0 Requirements

3.1 Replacement hierarchy architecture

In order to determine what data could replace what, we determined which data was of better quality than the other. Storm surge data is transmitted on a random channel, so this data can contain parity errors and there is not a parity flag set by the receiver at Wallops. Redundant water level data, although good quality, does not contain important information like water level sigma and outliers. PORTS data can have radio transmission failures, while satellite data may have transmission errors the problem area is flagged by the receiver at Wallops. Phone call data has the next to best quality, this data is collected

over a modem. If the during the collection, the phone has problem collecting data, the data is not saved. The best quality is the data collected directory from the DCP itself. When an engineer works on a DCP, the data is download to diskette before the work is started.

The following table identifies what incoming source can or can not replace the currently existing data for that given station at the given date/time:

	Existing					
Incoming	*Storm Surge*	*Redundant*	*PORTS*	*Satellite*	*Phone*	*Diskette*
Storm Surge	--	☹	☹	☹	☹	☹
Redundant	®	--	☹	☹	☹	☹
PORTS	®	®	--	☹	☹	☹
Satellite	®	®	☹	--	☹	☹
Phone	®	®	☻	☻	--	☹
Diskette	®	®	☻	☻	☹	--

® Replace, no testing ☻ Replace, testing with flag testing ☹ Do not Replace

3.2 Replacement criteria:

Tsunami and ADR data will not be included in the data replacement hierarchy. Storm surge water level data will be put in the ACOUSTIC table and the SS_ACOUSTIC table. The storm surge water level data in the ACOUSTIC table will be replaced with the next incoming data for that time. Storm surge ancillary data will be put in the individual ancillary tables and in the SS_MET table. The storm surge ancillary data in the individual ancillary tables will be replaced with the next incoming data for that time.

Four tables will be generated to contain all ingestion corrections and errors. The two ingestion correction tables are INGESTION_WL_CORRECTION and INGESTION_ANCILLARY_CORRECTION. The two ingestion error tables are INGESTION_WL_ERROR and INGESTION_ANCILLARY_ERROR. The error tables will be used for the daily report. Other reports can be written later to make use of the correction tables.

® If existing **Storm Surge or Redundant** value in the database for the given station at the given date/time, **no testing** is done, it is **replaced** with first incoming value from PORTS, Satellite, Phone Call, or Diskette. Replacement is noted in the dpas ingestion correction and error table. If replaced, then redo the flat and rofc flags for the 4 records before and the 2 records after the incoming date/time.

☻ If existing **PORTS or Satellite** value in the database for the given station at the given date/time that **differs** from the incoming Phone Call, and Diskette value, **replace** data record if all quality control flags match. Replacement is noted in the ingestion correction and error table. If replaced, then redo the flat and rofc flags for the 4 records before and the 2 records after the incoming date/time.

☹ If existing **Phone Call or Diskette** value in the database for the given station at the given date/time, **test** for differences from incoming Phone Call and Diskette, **no records are replaced**. If existing **Satellite** value in the database for the given station at the given date/time, **test** for differences from the incoming PORTS value, **no records are replaced**. If different from the incoming Phone Call, Diskette, or PORTS value, note in the ingestion error table.

Daily Report:

A daily report will be generated which reports the stations that have matched the above criteria for ☻, and ☹.

3.3 Performance

The primary performance requirement is to have data, at a preliminary level of quality, available as soon as possible from any incoming source (NWLON or PORTS).

3.4 Error and Recovery

The new program will have error handling for a number of items where appropriate.

3.5 Testing

The Acoustic, Backup Water Level, Pressure, Wind, Water Temperature, Barometric Pressure, Water Conductivity, Ancillary tables will be verified for a four week time period. Selected values for a given time period in each table will be evaluated. A test plan will be written and used prior to operation, that will explain testing methodology.

4.0 Operating Environment

ISD has chosen SGI as the data ingestion workstation platform. This allows for a homogeneous development environment within the office.

4.1 Hardware Requirements

Processor:	200 MHz MIPS R5000 (IP32)
Memory:	192 MB
Disk:	9 GB
Peripherals:	Mouse; Monitor
Operating System:	SGI IRIX UNIX

4.2 Support Software

- The major software components will be written in the Structured Query Language (SQL).

- Some software components will use the IRIX/UNIX Korn shell scripting language.

4.3 Security

- Since the product will be the property of the federal government, there should not be any proprietary regulation imposed on the product.

- Access to the OPAWL platform and the NWLON DMS is restricted and controlled via user names and passwords. Both platforms are behind the CO-OPS firewall

Appendix A - Replacement Requirements Team

Name	CO-OPS Division
Joann Nault (project manager)	Information Systems Division
Janet Burton	Information Systems Division, Deputy
Geoff French	Information Systems Division
Steve Gill	CO-OPS Staff
Leonard Hickman	Products and Services Division, Deputy
Thomas Huppmann	Products and Services Division
Thomas Landon	Requirements and Development Division
Phil Libraro	Field Operations Division
Thomas Mero	Requirements and Development Division, Chief
Mickey Moss	Field Operations Division
Mike O'Hargan	Field Operations Division, Chief
Bruce Servary	Requirements and Development Division
William Stoney	Products and Services Division, Chief
Brooks Widder	Requirements and Development Division

Appendix B - NWLON DMS Modifications

New Tables

REPLACEMENT_REASON

REASON_ID	small int
DESCRIPTION	varchar 80

INGESTION_WL_ERROR

STATION_ID	varchar 7
DCP#	varchar 1
SENSOR_ID	varchar 2
DATE_TIME	datetime
INCOMING_SOURCE	varchar 1
EXISTING_SOURCE	varchar 1
INFERRED	bit
WATER_LEVEL_VALUE	float
PRESSURE	float
FREQUENCY	float
WATER_TEMP	float
AIR_TEMP1	float
AIR_TEMP2	float
TEMP_FLAG	FLAG
WL_VALUE_FLAG	FLAG
WL_ROFC_FLAG	FLAG
WL_FLAT_FLAG	FLAG
WL_SIGMA_FLAG	FLAG
WL_OUTLIERS_FLAG	FLAG
WL_VS_BCK_FLG	FLAG
WL_VS_PRED_FLG	FLAG
REASON_ID	small int

INGESTION_WL_CORRECTION

STATION_ID	varchar 7
DCP#	varchar 1
SENSOR_ID	varchar 2
DATE_TIME	datetime
INCOMING_SOURCE	varchar 1
EXISTING_SOURCE	varchar 1
INFERRED	bit
WATER_LEVEL_VALUE	float
PRESSURE	float
FREQUENCY	float
WATER_TEMP	float
AIR_TEMP1	float

AIR_TEMP2	float
TEMP_FLAG	FLAG
WL_VALUE_FLAG	FLAG
WL_ROFC_FLAG	FLAG
WL_FLAT_FLAG	FLAG
WL_SIGMA_FLAG	FLAG
WL_OUTLIERS_FLAG	FLAG
WL_VS_BCK_FLG	FLAG
WL_VS_PRED_FLG	FLAG
REASON_ID	small int

INGESTION_ANCILLARY_ERROR

STATION_ID	varchar 7
DCP#	varchar 1
SENSOR_ID	varchar 2
DATE_TIME	datetime
INCOMING_SOURCE	varchar 1
EXISTING_SOURCE	varchar 1
INFERRED	bit
ANC_VALUE	float
DIRECTION	float
SPEED	float
GUST	float
MIN_FLAG	FLAG
MAX_FLAG	FLAG
GUST_MAX_FLAG	FLAG
ROFC_FLAG	FLAG
REASON_ID	small int

INGESTION_ANCILLARY_CORRECTION

STATION_ID	varchar 7
DCP#	varchar 1
SENSOR_ID	varchar 2
DATE_TIME	datetime
INCOMING_SOURCE	varchar 1
EXISTING_SOURCE	varchar 1
INFERRED	bit
ANC_VALUE	float
DIRECTION	float
SPEED	float
GUST	float
MIN_FLAG	FLAG
MAX_FLAG	FLAG
GUST_MAX_FLAG	FLAG
ROFC_FLAG	FLAG
REASON_ID	smallint

Table Modifications

REASON_ID	DESCRIPTION
1	Existing Redundant data replaced, no testing
2	Existing Storm Surge data replaced, no testing
3	Replaced data with no flag testing
4	Incoming flags were all zero
5	Incoming flags matched existing
6	Data has been verified, no replacement
7	Incoming Phone data does not match Diskette data

GLOSSARY

ADR	Analog to Digital Recorder
CO-OPS	Center for Operational Oceanographic Products and Services
DAS	Data Acquisition System
DCP	Data Collection Platform
DMS	Database Management System
FLAG	Represents a Sybase tinyint field, which is 8 bits
GOES	Geostationary Operational Environmental Satellite
ISD	Information Systems Division
NOAA	National Oceanic And Atmospheric Administration
NWLON	National Water Level Observation Network
NWS	National Weather Service
OPAWL	Operational Platform for Acquisition of Water Levels
OSO	Office of System Operations
PORTS	Physical Oceanographic Real-Time Systems
PSD	Products and Services Division
RDD	Requirements and Development Division
SAE	Shaft Angle Encoder
SGI	Silicon Graphics Inc.
SYBCHAR	Sybase character field