

NOAA's National Marine Fisheries Service



Data Management Plan for the Fisheries Monitoring and Analysis Division at the AFSC 2015 (DRAFT)

Version: 1.0

Date published: April 2, 2015

Presented to: The Fisheries Information Management Advisory Committee (FIMAC)

Contents

Data Management Plan for the Fisheries Monitoring and Analysis Division at the AFSC 2015	1
I. Purpose and Context	3
II. Data Sets	4
III. Data Providers	6
IV. Data Consumers	7
V. Best Practices Currently in Use	8
VI. Existing Dataflow and Processing	12
VII. Effectiveness Goals	13
VIII. Short Term Best Practices to Attain Effectiveness Goals and Target State	13
IX. Long Term Best Practices to Attain Effectiveness Goals and Target State	14
X. Implementation Schedule	14
XI. Metrics Description	14
XII. Metrics Prior to Planned Best Practices Implementation	14
XIII. Plan Review and Approval	15
XIV. Metrics Following Planned Best Practices Implementation	15
XV. Analysis	15
Appendix	15
Acronyms	15
Engineering (ERD) diagrams	16
Legacy Engineering (ERD) diagrams	16

I. Purpose and Context

This document is a National Marine Fisheries Service (NMFS) data management plan (DMP). The purpose of a NMFS Data Management Planning is to establish a series of incremental improvements to data management, and thus improve NMFS' ability to achieve its mission. The NMFS DMP identifies its most critical effectiveness goals and the corresponding best practices to achieve those goals. DMPs are generally associated with data stewarding divisions, with flexibility to create plans at high or lower organizational levels as appropriate. Because NMFS DMPs support a culture of continuing improvement, DMPs are updated annually.

NMFS DMPs should address one or more of the following effectiveness goals for their data sets: capability, speed, efficiency, cost reduction and confidence of data. **Capability** refers to the ability to perform a particular analysis that is dependent on data management practices and cannot be assumed. **Speed** is the ability to provide data or analyses as fast as possible. **Efficiency** means doing a task with fewer resources, or doing more with the same resources. **Cost reduction** is closely tied to efficiency but deserves special emphasis in times of shrinking budgets. NOAA, industry, and taxpayers must all have **Confidence** in our analytical results.¹

The NMFS DMP is to be used in conjunction with a National Oceanic and Atmospheric Administration (NOAA) data management plan. Each plan type has its own content and purpose. The NOAA DMP is a checklist of essential data management information associated with the data in question, whereas the NMFS DMP is a plan for improving the effectiveness of an organization by improving the way it manages data. The NOAA DMP is organized around the idea of the *data life cycle*, whereas the NMFS DMP is organized around *effectiveness goals* and corresponding *best practices*. See

https://www.nosc.noaa.gov/EDMC/documents/NOAA_EDM_Framework_v1.0.pdf and

<https://www.nosc.noaa.gov/EDMC/documents/EDMC-PD-DMP.pdf>

for definition and application of the data life cycle. Information on NMFS's effectiveness goals can be found at

<https://www.st.nmfs.noaa.gov/confluence/display/edm/Effectiveness+Goals>

Information on NMFS best practices can be found at

<https://www.st.nmfs.noaa.gov/confluence/display/edm/NMFS+EDM+Best+Practices>.

NOAA DMPs are associated with data sets, whereas NMFS data management plans are associated with divisions or other organizations. The NOAA DMP does not have a particular update schedule, whereas the NMFS DMP is updated annually.

The NOAA DMP for AFSC/FMA data sets can be found at

[FMA Database Data Management Plan](#)

II. Data Sets

Table 1 provides a listing of the Fisheries Monitoring and Analysis Division (FMA) data sets. The purpose is to provide the basic identifying information and InPort metadata catalog references for each collection of data objects.

Table 1

Name	Data Type	Data Steward / Contact	Location	InPort Catalog Number
Norpac	Relational Database	Julie Blair	NMFS AFSC	7284
Observer Interface (OBSINT)	Secondary data set transformed and flattened for fishery management and scientific research. Stored in semi-relational tables.	Julie Blair	NMFS AFSC	10637
Domestic Observer Data (Archive)	Archived npgop * data. Collected prior to 2008	Glenn Campbell	NMFS AFSC	9555
Foreign Observer Data (Archive)	Archived foreign fishing fleet observer data Collected prior to 1983	Glenn Campbell	NMFS AFSC	9554
Observer Declare and Deploy (ODDS) (Observer Restructure)	Relational Database	Craig Faunce	NMFS AFSC	16688
Observer Logistics (OLS)	Relational Database (A subsystem Within Norpac)	Glenn Campbell	NMFS AFSC	12269
Observer Vessel Survey (SUR)	Relational Database (A subsystem Within Norpac)	Andy Kingham	NMFS AFSC	
Custom Authentication Management (CAM)	Relational Database	Martin Park	NMFS AFSC	17193
Vessel Assessment Logging System	Relational Database (Sub-	Martin Park	NMFS -	

(VALS)	system of Norpac)		AFSC	
Alternative Catch Monitoring of Alaskan Groundfish	Relational Database (Archive)	Jennifer Cahalan	NMFS AFSC	18960
Bycatch characterization of North Pacific Halibut	Archived Data Sheets	Jennifer Cahalan	NMFS AFSC	7288

III. Data Providers

This organization’s data is provided by other organizations or is collected by this organization as shown in Table 2. Data providers include those external organizations which are the sources of AFSC’s data. AFSC is its own data provider for data that it collects from an original, non-derived, source. <Supplemental information, such as how the data will be used may also be included in the table if appropriate.>

Table 2

Data	Target Data Set(s)	Data Providers
Current Model Observer Catch and Effort Statistics	Norpac VALS	Fishery Observers, both at sea and stations at processing plants
Legacy Observer Catch and Effort Statistics.	Domestic Foreign	Legacy observer data.
Vessel Trip Logging	ODDS	Trip logging by vessel captains and owners.
Extrapolated Observer Catch and Effort	OBSINT	These are secondary data objects derived from observer statistics (Norpac).
Custom Authentication and Validation	CAMS	This system enables seamless user validation from NOAA Lap combined with customized user populations.

<Provide further description if necessary.>

IV. Data Consumers

This organization's data provides data to other organizations as shown in Table 3. A Data Consumer is an organization such as a NMFS division, that receives your data. AFSC is its own data consumer for data that AFSC uses to create end products such as analyses and reports.

Table 3

Data	Source Data Table(s)	Data Consumers
Status of Stocks	Norpac	North Pacific Fishery Management Council
AKFIN	Obsint Debriefed Tables	Fishery Councils, Research Scientists, Population and Ecological Research.
.PacFin	Obsint Debriefed Tables	Fishery Councils, Research Scientists, Population and Ecological Research..
Industry	Obsint Current Tables	Confidential extrapolated data provided to vessel owners and managers.
AKRO	Obsint Current Tables	Confidential extrapolated data pulled by the Alaska Regional Office as in Input to the catch accounting system.

<Provide further description if necessary.>

V. Best Practices Currently in Use

NMFS Enterprise Data Management (DEM) has provided the following list of best practices that are currently being applied to data sets: data access, documentation, integrated systems, ergonomic data collection, reduced reporting burden, quality assurance and quality control, state of the art at sea communications, consistent definitions, streamlining master data stores, change tracking, separation of obsolete and working data sets, standardization use of data types, clean table design and naming conventions, same site storage and computation services for bulk data and automation or repetitive data processing

Table 4 provides a list of FMA data sets and best practices currently being used for each.

Data Set	Best Practice	Approach Used
Norpac (Including Subsystems)	Data access: <ul style="list-style-type: none"> • Restricted under Magnuson-Stevens 	(External and Internal) <ul style="list-style-type: none"> • Limited to staff and researchers who have confidentiality agreements on file • Provided to vessel owners for vessels they either own or for whom an agency agreement is in place
	Documentation	Metadata published in inPort Functional specifications and development notes preserved on a shared drive. Engineering drawings are published on the AFSC website.
	Quality assurance and quality control	Electronic data checking: procedural debriefing, observer training, Inseason advising, post debriefing auditing
	Consistent Definitions	The meaning of each element and table is defined within the metadata
	Change Tracking	All active production data value changes are logged in history tables attached to the production data table. In these the old record value is recorded along with the user. Every record is time stamped.
	Clean table design and naming conventions	The AFSC has a published document outlining naming conventions. The database design is 3NF and engineering documents are created and published
	Automation	CRON jobs periodically load incoming observer fishery statistics, and initialize the extrapolation and transformation processes
	State of the Art at Sea Communications	On board client connects using vessel communications to the internet, where compressed and encrypted data and observer communications are transmitted to custom email processors.

Data Set	Best Practice	Approach Used
Observer Interface (OBSINT)	Data access: <ul style="list-style-type: none"> Restricted under Magnuson-Stevens 	(External and Internal) <ul style="list-style-type: none"> Limited to staff and researchers who have confidentiality agreements on file Provided to vessel owners for vessels they either own or for whom an agency agreement is in place
	Documentation	Metadata published in inPort Functional specifications and development notes preserved on a shared drive. Engineering drawings are published on the AFSC website.
	Quality assurance and quality control	These are secondary data products used for Inseason management by AKRO, Status of stocks analysis, and providing a near real time estimation of catch and effort to vessel owners or their designee
	Consistent Definitions	The meaning of each element and table is defined within the metadata
	Change Tracking	These data are generated from the production observer database. Data changes are recorded in history tables in that space.
	Clean table design and naming conventions	The AFSC has a published document outlining naming conventions. The database design is 3NF and engineering documents are created and published
	Automation	CRON jobs periodically load incoming observer fishery statistics, and initialize the extrapolation and transformation processes

Data Set	Best Practice	Approach Used
Domestic Observer Data	Data access: <ul style="list-style-type: none"> Restricted under Magnuson-Stevens 	(External and Internal) <ul style="list-style-type: none"> Limited to staff and researchers who have confidentiality agreements on file Provided to vessel owners for vessels they either own or for whom an agency agreement is in place This is an archived and static data recording observer catch and effort statistics prior to 2008
	Documentation	Table and column metadata are published in inPort
	Clean table design and naming conventions	This is a 2NF data set and not fully relational
	Published Engineering Diagrams	http://www.afsc.noaa.gov/FMA/PDF_DOCS/Domestic%20Fishing%20Table%20Diagram.pdf

Data Set	Best Practice	Approach Used
Observer Declare and Deploy (ODDS)	Data access: <ul style="list-style-type: none"> Restricted under Magnuson-Stevens 	(External and Internal) <ul style="list-style-type: none"> Provided to vessel owners for vessels they either own or for whom an agency agreement is in place The system is public facing to authenticated users for the logging of commercial fishing trips.
	Documentation	Metadata published in inPort Functional specifications and development notes preserved on a shared drive. Engineering drawings are published on the AFSC website.
	Quality assurance and quality control	Electronic data checking: procedural debriefing, observer training, Inseason advising, post debriefing auditing
	Consistent Definitions	The meaning of each element and table is defined within the metadata
	Change Tracking	All active production data value changes are logged in history tables attached to the production data table. In these the old record value is recorded along with the user. Every record is time stamped.
	Clean table design and naming conventions	The AFSC has a published document outlining naming conventions. The database design is 3NF and engineering documents are created and published

Data Set	Best Practice	Approach Used
Observer Logistics (OLS)	Data access: <ul style="list-style-type: none"> Restricted under Magnuson-Stevens 	(External and Internal) <ul style="list-style-type: none"> Track employment history, training, deployment and debriefing of Fishery Observers Limited to Staff
	Documentation	Metadata published in inPort Functional specifications and development notes preserved on a shared drive. Engineering drawings are published on the AFSC website.
	Quality assurance and quality control	Electronic data checking: procedural debriefing
	Consistent Definitions	The meaning of each element and table is defined within the metadata
	Change Tracking	All active production data value changes are logged in history tables attached to the production data table. In these the old record value is recorded along with the user. Every record is time stamped.
	Clean table design and naming conventions	The AFSC has a published document outlining naming conventions. The database design is 3NF and engineering documents are created and published

Data Set	Best Practice	Approach Used
Alternative Catch Monitoring of Alaskan Groundfish	Data access: <ul style="list-style-type: none"> Restricted under Magnuson-Stevens 	(External and Internal) <ul style="list-style-type: none"> Internal use for, the data source for an NPRB project with papers published in 2012.
	Documentation	Metadata published in inPort Functional specifications and development notes preserved on a shared drive. Engineering drawings are published on the AFSC website.
	Quality assurance and quality control	These data were filtered by our standard error detection software and vetted by standard debriefings.
	Consistent Definitions	The meaning of each element and table is defined within the metadata
	Change Tracking	Archived, Static Data
	Clean table design and naming conventions	The AFSC has a published document outlining naming conventions. The database design is 3NF and engineering documents are created and published

Data Set	Best Practice	Approach Used
Bycatch Characterization of North Pacific Halibut	Data access: <ul style="list-style-type: none"> Restricted under Magnuson-Stevens 	(External and Internal) <ul style="list-style-type: none"> Internal use for, the data source for an NPRB project with papers published in 2009.
	Documentation	Metadata published in inPort Functional specifications and development notes preserved on a shared drive. Engineering drawings are published on the AFSC website.
	Quality assurance and quality control	These data were filtered by our standard error detection software and vetted by standard debriefings.
	Consistent Definitions	The meaning of each element and table is defined within the metadata
	Change Tracking	Data not archived in a database. Raw deck sheets are stored in NOAA Archives

Data Set	Best Practice	Approach Used
Foreign Fishery Observer Data	Data access: <ul style="list-style-type: none"> Restricted under Magnuson-Stevens 	(External and Internal) <ul style="list-style-type: none"> Limited to staff and researchers who have confidentiality agreements on file Provided to vessel owners for vessels they either own or for whom an agency agreement is in place This is an archived and static data recording observer catch and effort statistics prior to 2008
	Documentation	Table and column metadata are published in inPort
	Clean table design and naming conventions	This is a 2NF data set and not fully relational
	Published Engineering Diagrams	http://www.afsc.noaa.gov/FMA/PDF_DOCS/Foreign%20Fishing%20Table%20Diagram.pdf

VI. Existing Dataflow and Processing

The following data flow diagrams for individual subsystems that comprise the Observer data assets are published on the FMA website and referenced by the AFSC library in inPort.

The following diagram depicts the transmission of data via satellite from vessels at sea to loading into the working table data store prior to loading the production observer statistics tables

[ATLAS Application Flow Diagram](#)

The following diagrams demonstrate the movement of data during the production data load, the deployment of observers to the field and debriefing at the end of an observer contract.

[Move Data to Production](#)

[Observer Deployment](#)

[Debrief Observers](#)

The following dataflow diagrams reflect the creation of data for Inseason management by AKRO, AKFIN, PACFIN, and analysis authors of stock assessments.

[Observer Interface Current Haul Data Transformation](#)

[Observer Interface Debriefed Data Transformation](#)

The following dataflow diagram reflects the movement of data during the validation of user accounts for all applications.

[Custom Authentication Data Flow](#)

The following dataflow diagram reflects the movement of data during the vessel assessment for temporary observer release system

[Vessel Analysis Logging System](#)

The training of North Pacific fishery observers is defined and described in the Observer Manual

[2015 Observer Sampling Manual](#)

VII. Effectiveness Goals

Below is a description of each effectiveness goal, and how it is anticipated they will be realized through the use of best practices.

Effectiveness Goals:

- **Capability** – The data collected by observers is a critical input for a great deal of what the AFSC accomplishes in realms of stock assessment and fisheries science and to the mission of AKRO in terms of Inseason fishery management.
- **Efficiency** – The availability of fishery dependent data compliments the direct fishery surveys and significantly extends the available biological data for fishery scientists and managers. The goal of FMA is to provide those services at the lowest possible cost.
- **Timeliness** – These data are a primary input for the managing of Inseason quotas of both targeted species and bycatch, including prohibited species. As such, the data must be error checked transformed and the presentation layer (OBSINT) made available to fishery managers in near real time.
- **Confidence** – The protocols of data collection by observers and the subsequent vetting and debriefing procedures are published and available both from InPort and the AFSC web site. The data structure and data flows are fully documented and published on the AFSC web site.

VIII. Short Term Best Practices to Attain Effectiveness Goals and Target State

The NMFS Enterprise Data Management Program categorizes best practices as listed above in Section V. The following table contains data sets that FMA continues to work to optimize. Primarily with re-engineering of the user interfaces, extensions to the data model and continuous improvement of the training and debriefing activities through Value Stream Mapping and adaptation of changes to statute and/or regulation.

Table 5

Data Set	Best Practice	Approach	Effectiveness Goal
Norpac	Integrated Systems	Migration from Oracle Forms to APEX	Efficiency Timeliness Confidence
Observer Interface	Integrated Systems,	Revising catch and effort extrapolation algorithms to reflect the current requirements of the fishery management council	Confidence Timeliness
	Quality Assurance	Revising catch and effort extrapolation algorithms	Confidence Timeliness
Observer Declare and Deploy	Integrated systems,	Adapting program to changing regulatory requirements	Confidence Capability
Observer Vessel Survey	Reduced reporting burden	Adapting this summary of observer experience to shifting regulatory and environmental conditions	Efficiency Confidence
All	Integrated systems,	Adapting program to changing regulatory requirements	Confidence

The target state is to have fully documented and publically published software and data object metadata, complete with data, and, flow and use case diagrams. The following is the current state.

Data Dictionary Definitions – published at the attribute level in InPort.

Software and data transformation summaries – In Process

Data Flow and Use Case diagrams – In Process

IX. Long Term Best Practices to Attain Effectiveness Goals and Target State

Although we have listed several best practices above in Table 5 that are currently in use, the focus of the Observer program going forward for the next several years will be integrating electronic monitoring with existing and revised observer programs. As those data sets, data acquisition, storage, and analysis products are developed, new measures of efficiency and quality will be applied. How those requirements will be implemented are not yet fully defined.

X. Implementation Schedule

Table 6

Date	Description
2015	Signed Data Management Plan
2016	Pre implementation of Electronic Monitoring Data Model and documentation
2017	Revision of Data Management Plan Reflecting Electronic Monitoring System Rollout

XI. Metrics Description

Metrics used in evaluating the attainment of effectiveness goals are detailed below. Although numeric measures are preferred, they are not mandatory. Proxy measures may be utilized when necessary, including discussion that the best practices improve effectiveness.

Effectiveness Goal	Metric Description
Capability	Number of forms converted, tested and deployed on APEX
Efficiency	The time spent by observers on vessel surveys, and the number completed
Confidence	Number of ERD, DFD, and Use Case Diagrams published and properly referenced by their InPort catalog Item. Completion and Publication of all detail level metadata for all observer systems.

XII. Metrics Prior to Planned Best Practices Implementation

XIII. Plan Review and Approval

This plan was reviewed and recommendations provided by:

Information Management Coordinator: _____ Date: _____

FIMAC Member: _____ Date: _____

FIMAC Member: _____ Date: _____

This plan was approved by:

Division <or other organization> Director: _____ Date: _____

< Plan can only be approved after completion of review process. >

XIV. Metrics Following Planned Best Practices Implementation

<Provide metrics associated with improved data management. >

XV. Analysis

<Compare expectation to results, note lessons learned, and provide possible explanations where appropriate, for use in subsequent data management planning. >

Appendix

Acronyms

AFSC – Alaska Fisheries Science Center

Atlas – Java client for data entry by observers and submission to the AFSC via satellite.

FMA – Fisheries Monitoring And Analysis Division;

The North Pacific Groundfish Observer Program

NPGOP – North Pacific groundfish observer program

Norpac – North Pacific observer statistics database

OBSINT – Observer interface database

VALS –Vessel assessment and logging subsystem

Engineering (ERD) diagrams

[Norpac - Current Active Observer Data Model](#)

[Custom User Authentication Data Model](#)

Legacy Engineering (ERD) diagrams

[Archived \(pre 2008\) Observer Data Model](#)

[Archived \(Foreign Fishery pre 1983\) Observer Data Model](#)