



NOAA
FISHERIES

August 7,
2013

DMIS: Data Matching & Imputation System

J. Michael Lanning, PhD
Chief, Monitoring and Analysis Section
APSD
Northeast Regional Office
J.Michael.Lanning@noaa.gov

Purpose

Match data sources to provide timely fishery trip level data in support of

- Weekly public catch and landing reports
- Analysis to support, improve and evaluate the region's fishery management programs and decisions

Purpose

System required to meet more detailed monitoring needed for groundfish sectors

- Moving from weekly and yearly aggregate catch estimates to individual vessel trip level accounting
- Resolution of current data may or may not support future needed analysis
 - One location point for fishing , i.e. stat area
 - Different data source – different collection rules

The combination of some data and an aching desire for an answer does not ensure that a reasonable answer can be extracted from a given body of data ~ John Tukey

Example for the following Terms of Reference

4) To what extent do fishery dependent data quality, statistical precision, and timeliness issues impact overall assessment accuracy, precision and timeliness?

6) What are the major fishery dependent data limitations/weaknesses and how could they be resolved? Define potential improvements and priorities for recommended improvements.

Primary Data Sources

Vessel trip reports

Dealer reports

Species catch reports via VMS

Observer reports

Allocation Management System

- (AMS - Compilation of Trip Notifications/Declarations)

Permits / MQRS

vTrack (Functionality to be added in future version)

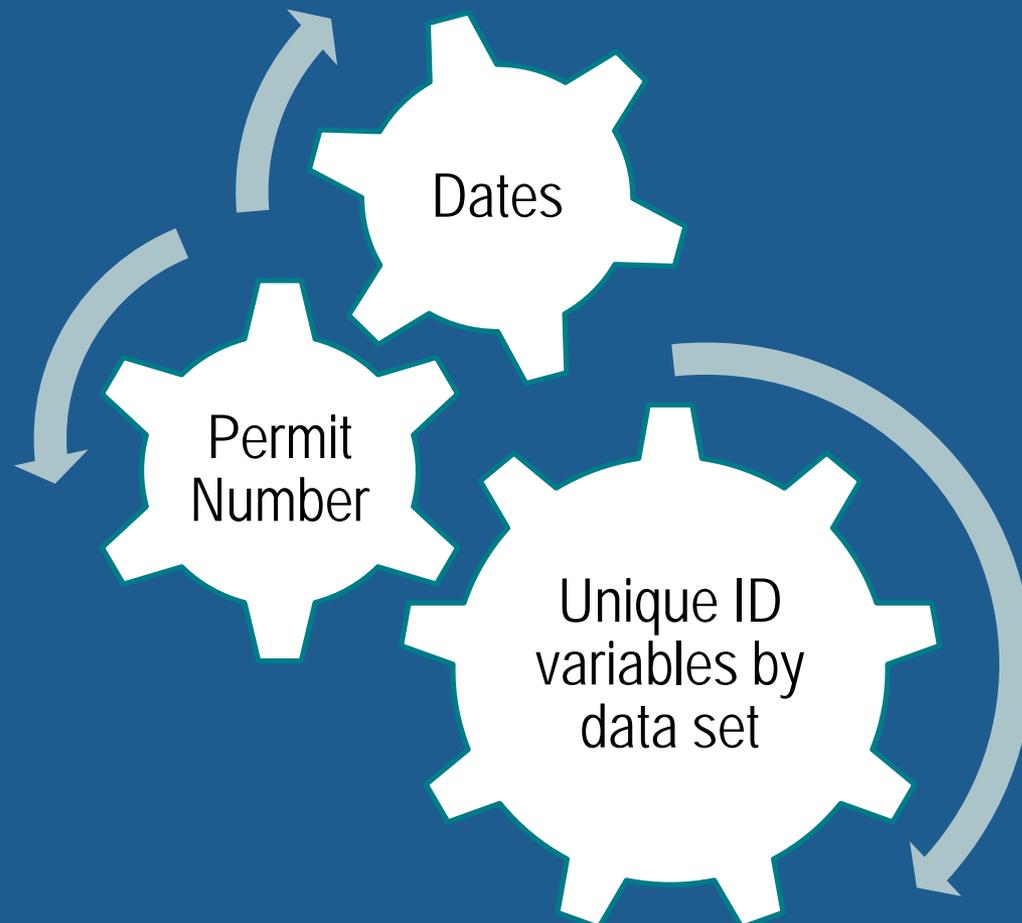
DMIS: Major Parts

Build map of matching data sets

Divide and classify matched data into individual trips

Apply fishery management plan decision and business rules

Build Map of Matching Data Sets



Build Map of Matching Data Sets

General Order of Matching

- AMS Declaration ↔ VTR / Catch Reports ↔ Observer Reports ↔ Dealer Reports ↔ Informational Data (Permits, Home Ports, etc.)

Matching Between Record Sources is not One to One

- Multiple declarations per VTR
- Multiple VTRs per declaration
- Multiple observer reports per VTR
- Multiple dealer reports per VTR
- Dealer reports rolling up two or more trips
- Dealer reports with just one species belonging to multiple trips (Functionality to be added in future version)

Build Map of Matching Data Sets

Orphan Records exist

Updated once a week, moving toward daily

- Updated information depends on updating of source data
- Dealer records are updated weekly
- Observer reports can be weekly to monthly depending on fishery
- Ongoing record corrections

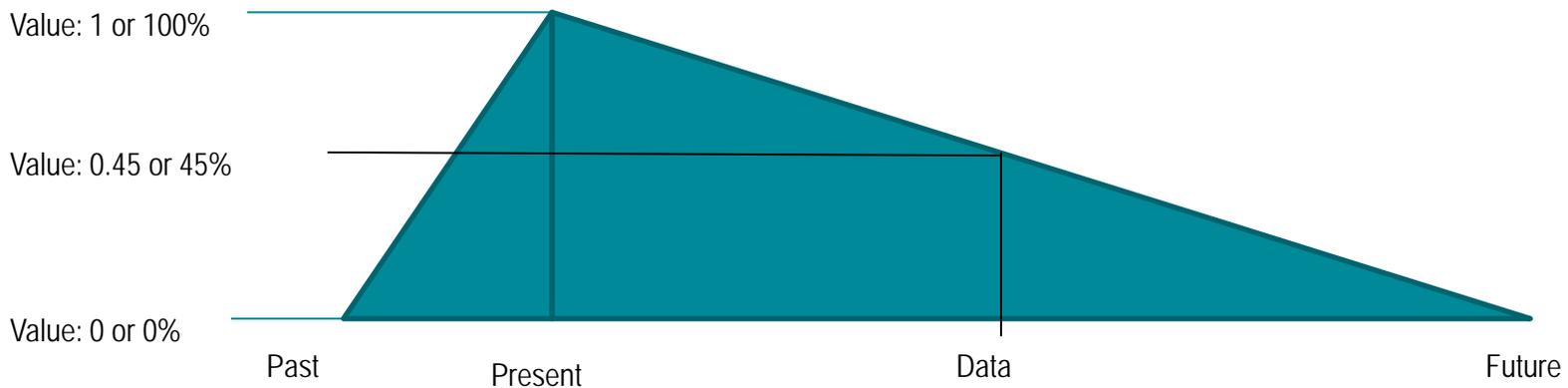
Three Primary Matching Criteria

- Permit number (hard match)
- VTRSERNO (hard when exists, limited to past month - duplicate usage)
- Dates (DMIS evaluates strength of match through membership functions)

Build Map of Matching Data Sets

Date Matching

- Restricted: Set Only Trips, Transport, Party/Charter
- Based on Membership Functions



Classify Matched Data into Trips

VTR or VMS catch report defines a trip

- 1 AMS Declaration matched to 3 VTRs = 3 Trips
- 2 AMS Declarations matched to 1 VTR = 1 Trip

Declarations with no VTR or catch report

Orphan dealer reports group by date and port

- (Definition to be added in future versions)

Apply trip-specific information Trip (Ports, Permits, Etc.)

Fishery Management Plan Decision and Business Rules

Fishery Dependent Data Collection Working Group

Examples:

- FMP and attribution of catch
- Groundfish dealer to VTR apportionment
- Area (reported, calculated, vTrack) determination
- Port assignment (Home, Landing, Dealer)
- Discard calculation by stock

Fishery Management Plan Decision and Business Rules

Missing Value Estimation (Imputation)

- Based on matched DMIS data
- Restricted to trips with VTR and dealer records
- Information imputed
 - Area: Restricted to permit history (vTrack information to be added in future)
 - Gear: Restricted to permit history
 - Sector landings (Place holder for the purpose of quota monitoring until VTR and Dealer Reports are submitted)

Questions?