

Determining a minimum video review rate to estimate discards in New England groundfish

Daniel W. Linden, Ph.D.

Analysis and Program Support Division

Greater Atlantic Regional Fisheries Office

Data from Groundfish EM projects

- 25 vessels
- **2**016–2019
- 567 trips
- 2,104 hauls





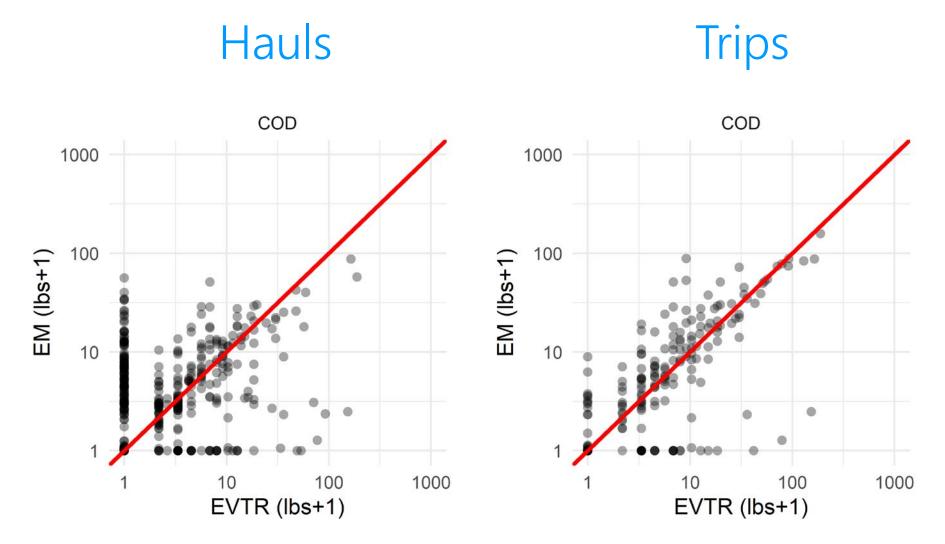






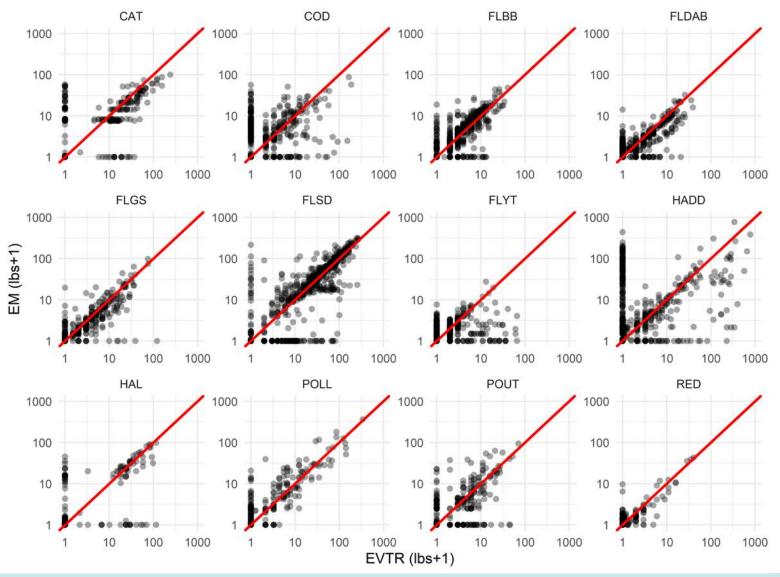


Discard comparison



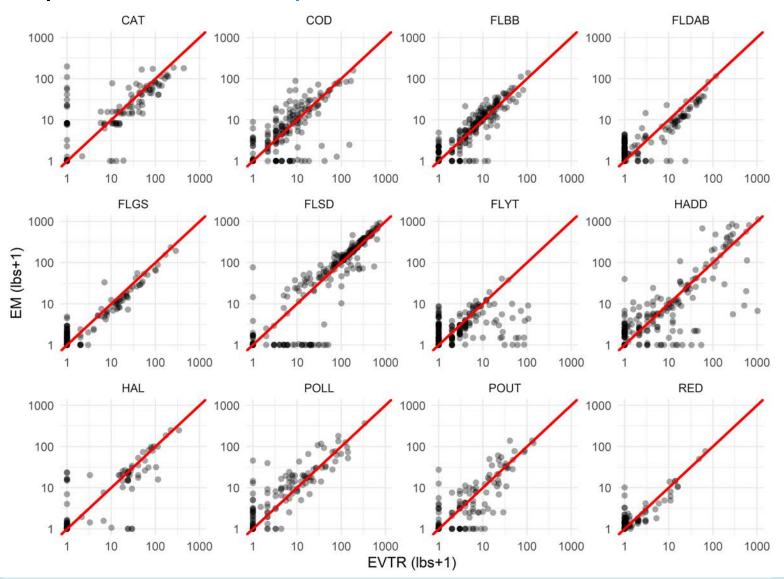


Discard comparison – Haul level





Discard comparison - Trip level





Model-based approach to discard estimation using audit review

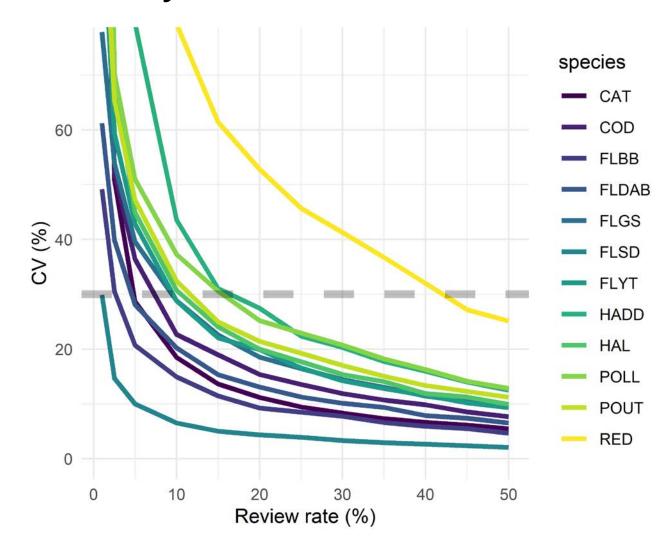
- Review random X% of fishing effort
- Fit model to predict EM using VTR
- Estimate adjustment to non-reviewed effort
- Similar to design-based SBRM



Model-based approach to discard estimation using audit review

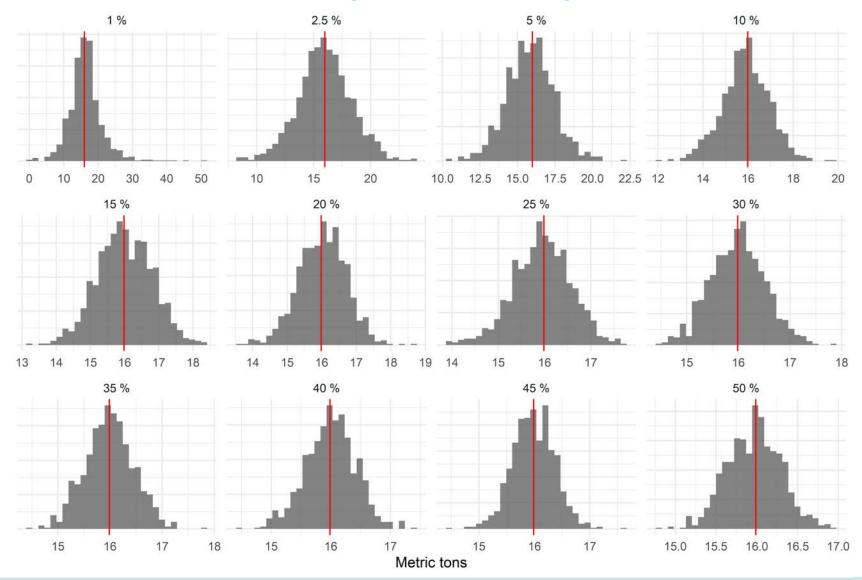
$$\begin{aligned} d_{\text{EM}} &\sim \text{Poisson}(\lambda) \\ \log(\lambda) &= \beta_{species} + \beta_{\text{VTR}} \log_{10}(d_{\text{VTR}} + 1) \\ \beta_{species} &\sim N(\mu, \sigma) \end{aligned}$$

RESULTS: Uncertainty across review rates



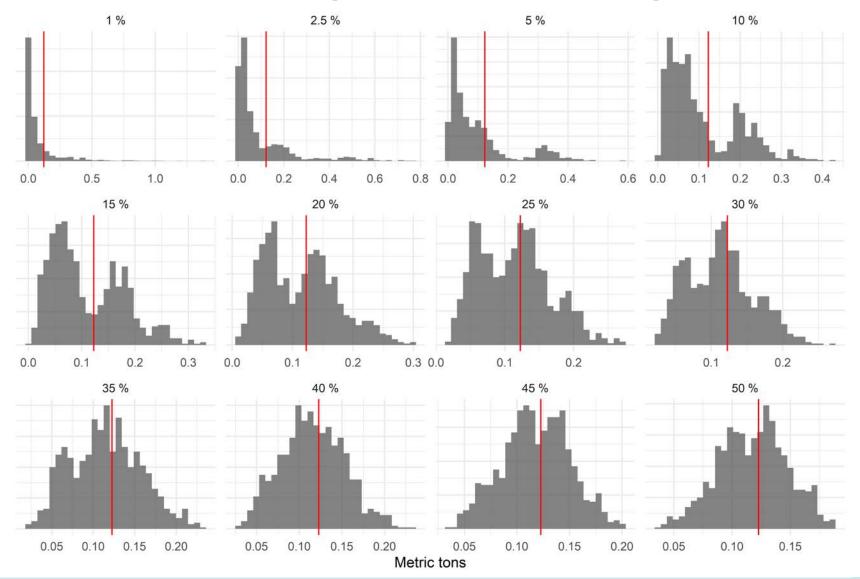


Discard estimation: FLSD (sand dabs)



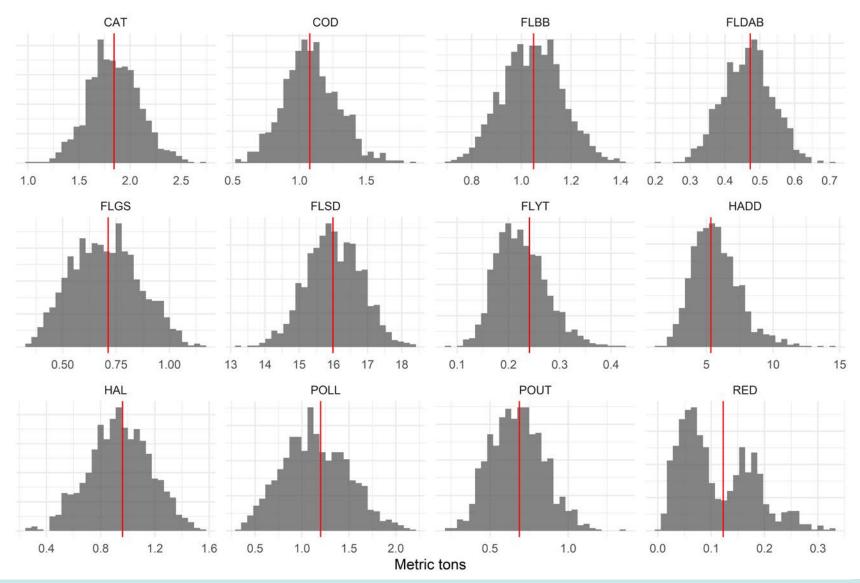


Discard estimation: RED (Acadian redfish)





Discard estimation under 15% review





Considerations for the audit approach

- Additional variables to model
 - Gear
 - Permit
- Vessel performance across fleet
 - High in EFP projects
- Haul monitoring is best
- Estimation vs. reporting



Questions?



Electronic Monitoring Program Design: Video Review and Data Processing Cost Considerations in a Pilot Reef Fish Monitoring Program in the Gulf of Mexico

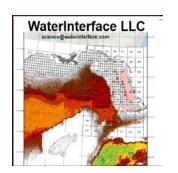
Authors: Roberts, Daniel¹, C. Neidig², M. Lee², J. Steinwachs², T. Taccardi², R. Taccardi² and R. Schloesser²

 ¹WaterInterface LLC, 1620 Surrey Trail, Wimauma, FL. 33598
 ²Mote Marine Laboratory, Directorate for Fisheries and Aquaculture, 1600 Ken Thompson Parkway, Sarasota, FL 34236

National Electronic Monitoring Workshop – East Coast, November, 13-14, 2019













PRESENTATION OBJECTIVES:

- 1. Sampling Design, Statistical basis and Theoretical Validation
- 2. Explain the MML CFEMM EM review and data processing approach.
 - A. Describe Review Process Steps.
 - B. Describe Error Checking and Quality Control.
 - C. Identify Outputs.
- 3. Provide a brief preliminary description of staffing costs.
 - A. Capital Outlay.
 - B. Expense.
 - C. Staffing.
- 4. Identify Cost Drivers as Fulcrums for Cost Control.
 - A. Staffing.
 - **B.** Data Storage

2c: Review Process Outputs Spatial and Temporal framework

- 1. Species Identification
- 2. Discards
- 3. Condition on Arrival
- 4. Disposition
- 5. Shark Sex Determination
- 6. Shark Size Estimates
- 7. Catch Per Unit Effort (CPUE)
- 8. Bycatch

Program is grant-based; each has a set of specific objectives, the above are common to all grants

Random subsampling of Trip Set-Haul-Events

1.Post-trip subsampling 25% of set-haul-events.

2.Culling and validation of subsample location continuity by comparing set locations of subsamples to set locations of a sample frame containing 1357 set-haul-events.

Experimental Unit = Set-Haul-Event

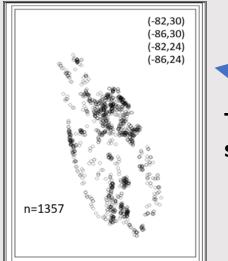
We assess the relevance of the subsample to the fishing area and historical dataset

Fishing Area = 84,000 square kilometers or about 32,402 square miles.

Depths = about -40 to -300 m although some fishers fish in deeper water.

Basically the
West Florida Shelf
from the Florida Middle
Grounds to Pulley Ridge
from the escarpment to
about 40 m.





The Fishing Area Site Map from which sampling validation was constructed.



Data Flow and Post-trip Sample Frame

Experimental Unit = Set-Haul-Event (SHE)

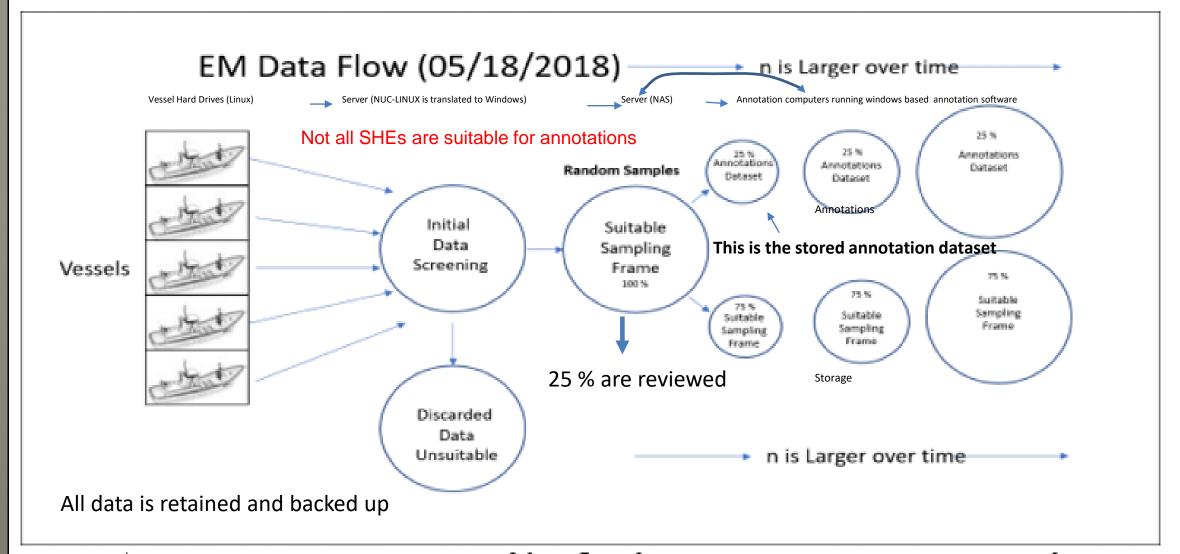


Figure 25. Diagrammatic representation of data flow for MML EM annotation process as of May 2018.

VALIDATION: Test the null hypothesis that post-trip random sampling of fleet Set-Haul-Events was biased using a 3 paradigmatic statistical approach Centrography, density-intensity, geographic)

In every case the **null hypothesis was rejected**: subsamples and Monte Carlo simulations seeded from the sample frame of fleet SHE locations were statistically homogeneous.

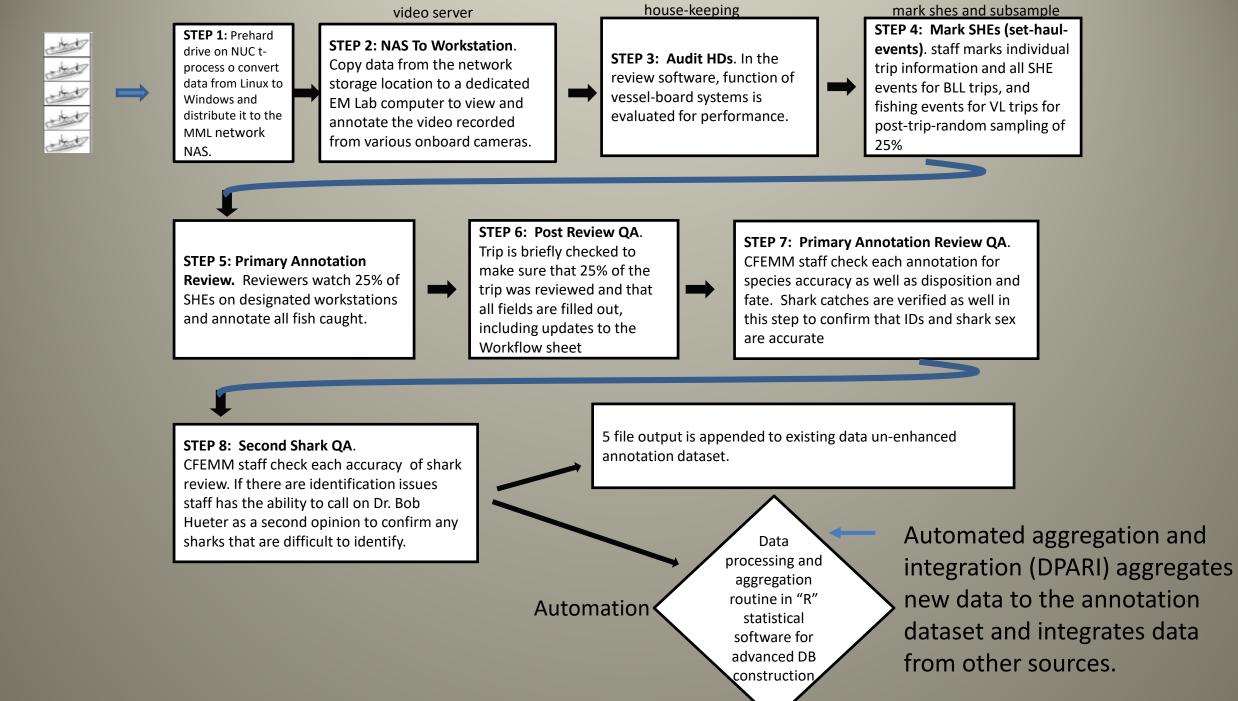
15 methods to evaluate sample frame, random sample, and 12 simulation continuity.

- 1. Centrography ANOVA; Distribution Analysis; Parametric; non-parametric.
- 2. Density/intensity-Nearest neighbor; Quadrat; kernel density; K, L-functions.
- 3. Geography-comparison of convex hulls.

Method	Test	Statistics	Hypothesis (H0)	Dataset	Variable	Result	Probability
Centrography							
	2-Sample comparison	K-S Test Lat/Lon	Samples same distribution	Annotations, AllSets	Lat/Lon	Accept	95%
		t-test	Means are equal	Annotations, AllSets	Lat/Lon	Accept	95%
		Mann-Whitney W	Medians are equal	Annotations, AllSets	Lat/Lon	Accept	95%
		Graphic Comparison	Samples same distribution	Annotations, AllSets	Lat/Lon	Accept	
		Centroid Plot	Centroid Locations Equal	Annotations, AllSets	Lat/Lon	Accept	
		Graphic Comparison	Density Traces Equal	Annotations, AllSets	Lat/Lon	Accept	
	4-sample comparison	Smooth Kernel Density	Graphically Similar	Annotations, AllSets, 2 simulations	Point Pattern	Accept	
	4-Sample comparison	Goodness-of-Fit;K-S, CVM, and AD	Samples from Normal Dist.	Annotations, AllSets	Lat/Lon	Reject	95%
	14-Sample comparison	Homogeneity Variance; Levene	Variances All Equal	14 Datasets; All Sims, Ann. AllS.	Lat/Lon	Accept	95%
	14-Sample comparison	ANOVA;Decomp.Var. Equal Means	Means are equal	14 Datasets; All Sims, Ann. AllS.	Lat/Lon	Accept	95%
	14-Sample comparison	Kruskal-Wallace Equal Medians	Medians are equal	14 Datasets; All Sims, Ann. AllS.	Lat/Lon	Accept	95%
DensityIntensity							
	4-sample comparison	Quadrat Analysis (R, Spatstat)	Set Locations Random	Annotations, AllSets, 2 simulations	Point Pattern	Reject	95%
	3-sample comparison	Near.Neigh. ECDF Comparison;	Set Locations Random	Annotations, AllSets, 2 simulations	Point Pattern	Reject	
	3-sample comparison	Mood's Median NNDist	No Difference Medians	Annotatins, 2 simulations	Point Pattern	Accept	95%
	3-sample comparison	K-W NNDist Equal Medians	No Difference Medians	Annotatins, 2 simulations	Point Pattern	Accept	95%
	4-sample comparison	NNDist;nndiff<-simulation (nndist(ANN)- (nndist(simulation)	Coerced linear trends equal	Annotations,3 simulations	Point Pattern	Accept	
	5-sample comparison	Average NN Z-test	Nndist Random	Annotations, 4 simulations	Point Pattern	Reject	All p-value = 0.01 z- score<2.58; clustered
	5-sample comparison	Ripley K test comparison; comparison of linear trends	Observed K-values similar;Similar Slopes	Annotations, 4 simulations	Point Pattern	Accept	
Geographic							
	14-Sample comparison	Comparison of planar scatter plots	Graphically Similar	Annotations, AllSets, 12 simulations	Point Pattern	Accept	
	6-Sample comparison	Summary StatisticsArea/Perimeter comparison	All Convex Hulls are similar	Annotations, AllSets, 4 simulations	Polygon Comp.	Accept	

2. STEPS IN THE REVIEW PROCESS; DATA PROCESSING





DPARI: DATA PROCESSING, AGGREGATION AND INTEGRATION ROUTINE

DPARI

Data Processing Aggregation and Integration Routine

Aggregates output from annotation review

Ecological, environmental, oceanographic and other GIS data

Spatial join in R

Data Joined to annotation locations: Rugosity, depth, temperature, current velocity, current direction, geomorphic features, etc.

Final dataset for analysis and reporting

2. COSTS OF THE REVIEW PROCESS; DATA PROCESSING

Not ignoring capital outlay, depreciation, or expense, just not enough time: Those costs are fairly constant from program to program (workstations, servers, etc.).

Concentrate on **staffing costs**.

Staffing costs are based on number of staff, wages, time, number of trips, number of vessels, sample size etc.

Time is how long it takes in minutes to complete each step of the review process.

					the NAS	into the CFEMM databases.
3	Audit HDs	After a trip is opened up in the Review Software, it is checked to find any issues with the onboard system (ie sensor problem, broken magnet, camera damage). If it is concluded that there is some form of system malfunction, automatically generated audit files from the trip are sent to SWI technicians to	7	Primary review QA	CFEMM staff check <u>each</u> annotation for species accuracy as well as disposition and fate. Shark catches are verified as well in this step to confirm that IDs and shark sex are accurate.	
4	Mark SHEs	CFEMM staff marks individual trip information and all SHE events for BLL trips, and fishing events for VL trips	8	2nd Shark QA	Staff has the ability to call on Dr. Bob Hueter as a second opinion to confirm any sharks that staff feels are difficult to	
		ps of the EM data review SWI output for advance	•			DPARI is a Value-added

STEP

5

6

TITLE

Primary

Review

Annotation

Post review QA

DESCRIPTION

Reviewers watch 25% of SHEs

on designated work stations

and annotate all fish caught

Trip is briefly checked to make

filled out, including updates to

the Workflow sheet prior to

workstation to a OA folder on

moving the trip off the

sure that 25% of the trip was reviewed and that all fields are REVIEW OUTPUT

STANDARD SWI OUTPUT

processing and aggregation routine

Team, 2017) that enables integration

of catch data from vessels and other

sources to be seamlessly integrated

in "R" statistical software (R Core

DPARI: DPARI is a data

STEP

TITLE

Process HDs on

NUC

NAS to

Workstation

DESCRIPTION

Hard Drives are loaded using one of

three NUCs (NUCs are linked to NAS).

platform to Windows platform and

Trips are copied from the NAS to one

of six open work stations for steps 4

Data is translated from LINUX

uploaded to a NAS

and 5.

A. STAFFING COSTS 5 BLL VESSELS

	Α	В	С	D	Е	F	G	Н	1
1	Dataset ID	Vessel	# of Trips	Processed on NUC #	Date	Ву	Short Drive #	Captain Log Forms Retrieved	User Time (min)
			Ст. трс			DL,		11011101101	()
2	0623869_20181008	No Bull	1	1	10/10/2008	СМС	S1015	No	10
3	1057895_20171117	Miss Ruby	1	1	01/09/2018	JPS		Yes	10
4	0623869_20171211	No Bull	1	3	01/09/2018	JPS		Yes	10
5	1057895_20171211	Miss Ruby	1	2	01/09/2018	BH		Yes	10
6	0565290_20170814	Miss Donna	1	2	01/16/2018	JPS		Yes	10
7	0623869_20171226	No Bull	1	1	01/19/2018	MN		Yes	10
8	0544027_20180126	Midnight Sun	1	2	01/27/2018	CN		No	10
9	1057895_20180129	Miss Ruby	1	1	01/27/2018	CN		Yes	10
10	0544027_20171218	Midnight Sun	1	3	01/30/2018	DL		No	15
11	0544027_20171218	Midnight Sun	2	3	01/30/2018	DL		No	0

The dataset used was compiled of 118 trips from January, 2018 Through October, 2019 $^{\sim}$ 22 months. Data represents the efforts of 5 BLL vessels fishing the WFS.

STEPS 1-4	Process_HDs_NUC (min)	NAS_WorkStation (min)	Audit_HDs (min)	Mark_SHEs (min)	Total Time (min)
Mean (Per Trip)	8.36	5.71	19.64	52.55	86.26
Standard Error	0.64	0.37	0.97	5.54	
Standard Deviation	6.99	3.86	8.06	55.67	
Sample Variance	48.81	14.88	64.94	3099.35	
Range	60.00	30.00	30.00	430.00	
Minimum (Per Trip)	0.00	0.00	10.00	5.00	
Maximum (Per Trip)	60.00	30.00	40.00	435.00	
Total (118 Trips)	986.00	634.00	1355.00	5308.00	8283.00
	Total Time (hrs)	Estimated Labor Cost (22 months)	Estimated Labor Cost (per month)	Estimated Labor Cost (per year)	Labor Cost Basis = \$20.00 per hour
Mean (Per Trip)	1.44	\$28.75	\$1.31	♦ \$15.68	Number of
Total (118 Trips)	138.05	\$2,761.00	\$125.50	\$1,506.00	Trips =118

Table_. Time required and estimated labor costs to process review steps 1 through 4 for 118 BLL trips WFS, GOM at MML CFFEM. Labor cost basis \$20.00 per hour.

STEP 5	Primary	Species Annotation	Species Annotation Review			
	Useable Hauls	Hauls Reviewed (~25 %)	View Time (min)	View Time (hrs)		
Mean (Per Trip)	21.60	7.21	925.46	15.42		
Standard Error	1.92	0.78	117.05			
Standard Deviation	19.38	7.85	1158.77			
Sample Variance	375.75	61.57	1342758.11			
Range	109.00	45.00	7418.00			
Minimum (Per Trip)	0.00	0.00	0.00			
Maximum (Per Trip)	109.00	45.00	7418.00			
Total Hauls	2203.00	735.00	90695.00	1511.58		
	Cost (22 months)	Cost (per month)	Cost (per year)			
Mean (Per Trip)	\$308.49	\$14.02	\$168.27			
Total Hauls	\$30,231.67	\$1,374.17	\$16,490.00			

Table_. Useable hauls, hauls reviewed, view time and labor costs for species annotation video review for 118 BLL trips WFS, GOM at MML CFFEM. Labor cost basis \$20.00 per hour.

Top table is hauls reviewed, processing time for Revew Team: Bottom table is Costs

STEP 6	Post Revew QA		
	Post_Review_QA_Rev _Time (min)	Post_Review_QA _Rev_Time (hrs)	
Mean Per Trip	8.25	0.14	
Standard Error	0.55		
Standard Deviation	5.28		
Sample Variance	27.92		
Range	25.00		
Minimum Per Trip	5.00		
Maximum Per Trip	30.00		
Total	751.00	12.52	
	Cost (22 months)	Cost Per Month	Cost Per Year
Mean Per Trip	2.75	\$0.13	* \$1.50
Total	250.33	\$11.38	\$136.55

Table_. Cursory post review QA view time and labor costs for species annotation video review for 118 BLL trips WFS, GOM at MML CFFEM. Labor cost basis \$20.00 per hour.

Top table is processing time for Revew Team: Bottom table is Costs

STEP 7		Primary Revew QA		
	# Species Corrected	# Annotations	Primary QA Time (min)	Primary QA Time (hrs)
Mean	31.13	444.20	94.02	1.57
Standard Error	6.11	63.26	12.97	
Standard Deviation	59.55	596.84	122.32	
Sample Variance	3546.01	356214.89	14961.75	
Range	400.00	3791.00	600.00	
Minimum	0.00	7.00	5.00	
Maximum	400.00	3798.00	605.00	
Total	2957.00	39534.00	8368.00	139.47
	Cost (22 months)	Cost (per month)	Cost (per year)	
Mean	\$31.34	\$1.42	\$17.09	
Total	\$2,789.33	\$126.79	\$1,521.45	

Table_. Primary review species ID corrected, view time and labor costs for species annotation video review for 118 BLL trips WFS, GOM at MML CFFEM. Labor cost basis \$20.00 per hour.

Top table is corrections, processing time for Revew Team: Bottom table is Costs

		2nd SHARK Q A		
STEP 8	# Species Corrected	# Sharks Total	Shark QA Time (min)	Shark QA Time (hrs)
Mean (Per Trip)	0.01	6.39	2.97	0.05
Standard Error	0.01	2.47	0.52	
Standard Deviation	0.11	22.78	5.09	
Sample Variance	0.01	519.07	25.86	
Range	1.00	200.00	25.00	
Minimum (Per Trip)	0.00	0.00	0.00	
Maximum (Per Trip)	1.00	200.00	25.00	
Total	1.00	543.00	279.00	4.65
	Cost (22 months)	Cost (per month)	Cost (per year)	
Mean (Per Trip)	\$0.99	\$0.04	\$0.54	
Total	\$93.00	\$4.23	\$50.73	

Table_. Second shark QA: ID corrected, view time and labor costs for shark annotation video review for 118 BLL trips WFS, GOM at MML CFFEM. Labor cost basis \$20.00 per hour.

Top table is corrections, processing time for Revew Team: Bottom table is Costs

Staffing Cost Summary for Annotation Review Based on Actual Data with Training and Research Activities

1 Staff	Mean Cost/Trip Per Year	Mean Cost/Trip Per Vessel/Year	Mean Cost/Vessel/Sea Day	Data Analysis Per Year	Data Storage Per TB/Month
Step 1-4	\$15.68	\$3.14	\$0.31	\$13,000.00	\$4.0000
Step 5	\$168.27	\$33.65	\$3.37		
Step 6	\$1.50	\$0.30	\$0.03		Data Retrieval Per TB
Step 7	\$17.09	\$3.42	\$0.34		\$10.00
Step 8	\$0.54	\$0.11	\$0.01		
Total	\$203.08	\$40.62	\$4.06		
10 Staff					
Step 1-4	\$156.80	\$31.40	\$3.10		
Step 5	\$1,682.70	\$336.50	\$33.70		
Step 6	\$15.00	\$3.00	\$0.30		
Step 7	\$170.90	\$34.20	\$3.40		
Step 8	\$540.00	\$1.10	\$0.10		
Total	\$2,565.40	\$406.20	\$40.60	Volunteer Review Staff = no staff salary at MML	

Current Annotation Review Production Elements

Currently reviewing 60 Trips per year

5 BLL Vessels WFS (largest dataset)

Review Time: 1000 minutes to review 1 Trip with 10 Set-Haul-Events

Each Haul takes about 100 minutes; with no training, just production

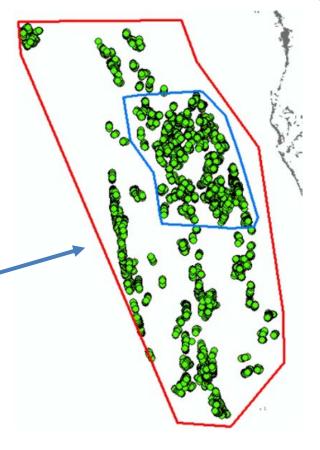
Current Data Archived = 120 TB; Rate = 30 TB per year

Current data storage is network storage (2 NAS, 8 TB Seagate drives

Managing Costs of Review Processes: Drivers Personnel, Data Storage

Sylvia et al.: 1. video review costs depend on the level of mandated video sampling (5-100%) as well as 2. the goals of the review (estimating discard volumes versus species identification, size of individual fish etc.)

- 1. Species Identification
- 2. Discards
- 3. Condition on Arrival
- 4. Disposition
- 5. Shark Sex Determination
- 6. Shark Size Estimates
- 7. Catch Per Unit Effort (CPUE)
- 8. Bycatch



- 1. Modify current subsample strategy
 - A. Stratify by location
 - B. Stratify by vessel
- 2. Statistically investigate subsample alternatives to reduce 25% subsample size
- 3. Modify objectives (outputs)

20 % of fishing area; 64 % of catch; highest CPUEs, highest species diversity; highest catch of red grouper and red snapper.

Summary

This was a summary

West Coast EM Program Design: Video Review and Data Processing

Courtney Paiva

Pacific States Marine Fisheries Commission

National Electronic Monitoring Workshop – East Coast
New Castle, New Hampshire
November 14, 2019



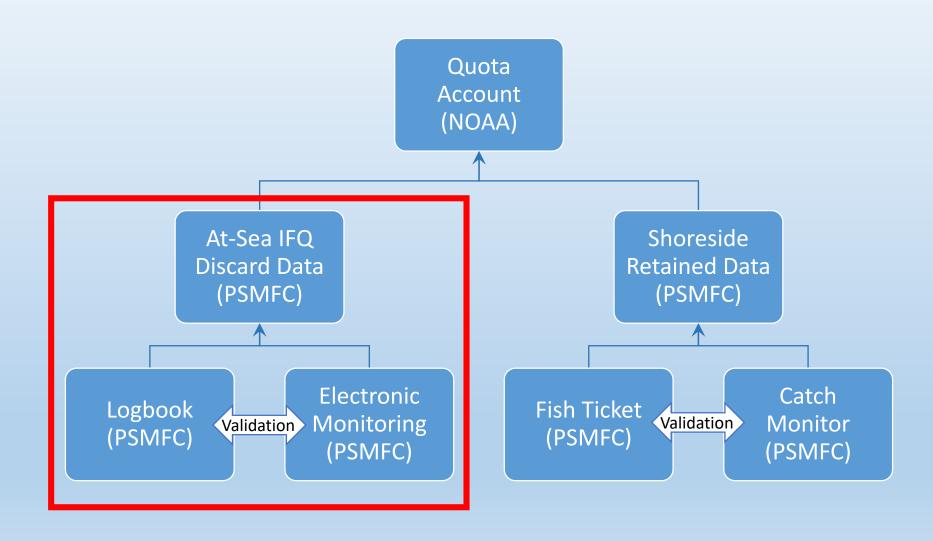
West Coast EM Program Overview

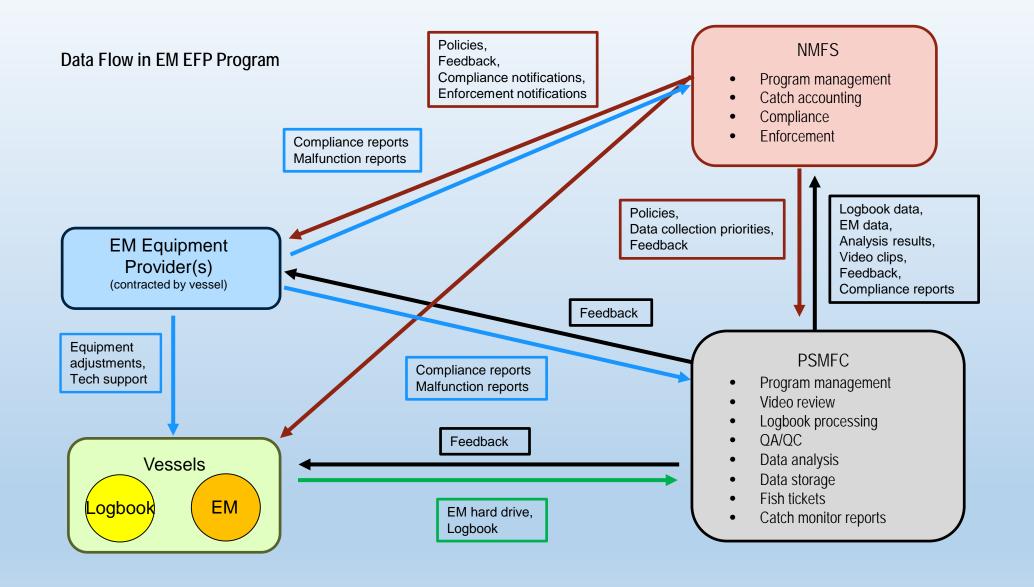
- 100% observer/EM coverage
- EM for compliance monitoring of quota species
 - Providing data on IFQ discards
- Logbook is primary source for discards
 - EM used to audit logbook discards
- EM Review: IFQ discards and prohibited species





West Coast Catch Accounting





How are West Coast EM data used in compliance?

- EM vs. Logbook data
 - Comparison of species weights
 - EM data used to audit LB data
- When EM \neq LB:
 - 10% allowable discrepancy between LB & EM
 - When >10% difference → use the larger estimate
 - No allowable discrepancy for overfished species
 - Use the larger estimate
 - If LB estimate missing → use EM estimate
 - Occurs when EM sees a discard that's not in the LB
 - If **EM estimate** missing → use LB estimate
 - Occurs when the LB has a discard that EM does not, or in the rare event that the EM data cannot be reviewed (system malfunction, drive malfunction, etc.)

Species Retention Rules: 2015 vs. Current

- 2015:
 - All fisheries were Maximized Retention (1 bottom trawler was optimized retention)
- 2016 to Present:
 - Whiting = Maximized Retention
 - Bottom Trawl & Fixed Gear = Optimized Retention (different retention rules)
 - Driven by fishers wanting the ability to discard bycatch at-sea

Total Deck Length = 40m Total Deck Width = 8m 1.80m 1.80m 122" 124" 124" 124" 124" 110" Trawl Alley Length = 31m Trawl Alley Width = 2.4m Discard Control Point 105" 35" 110" 110" Discard Control Point

Maximized Retention

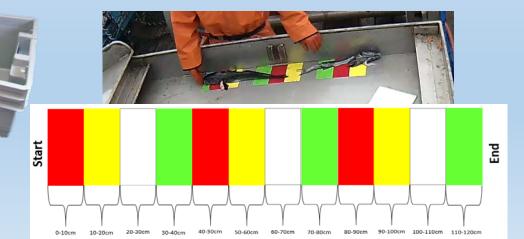
- Allowed to discard
- 1 tote/haul operational discards
- Animals >6ft
- Invertebrates
- Debris
- Unavoidable discards
- Prohibited Species

Optimized Retention (Bottom Trawl)

- Allowed to discard:
- Everything in Maximized Retention
- 6 IFQ species:
- Arrowtooth flounder
- English Sole
- Dover Sole
- Pacific Sanddab
- Pacific Hake
- Lingcod

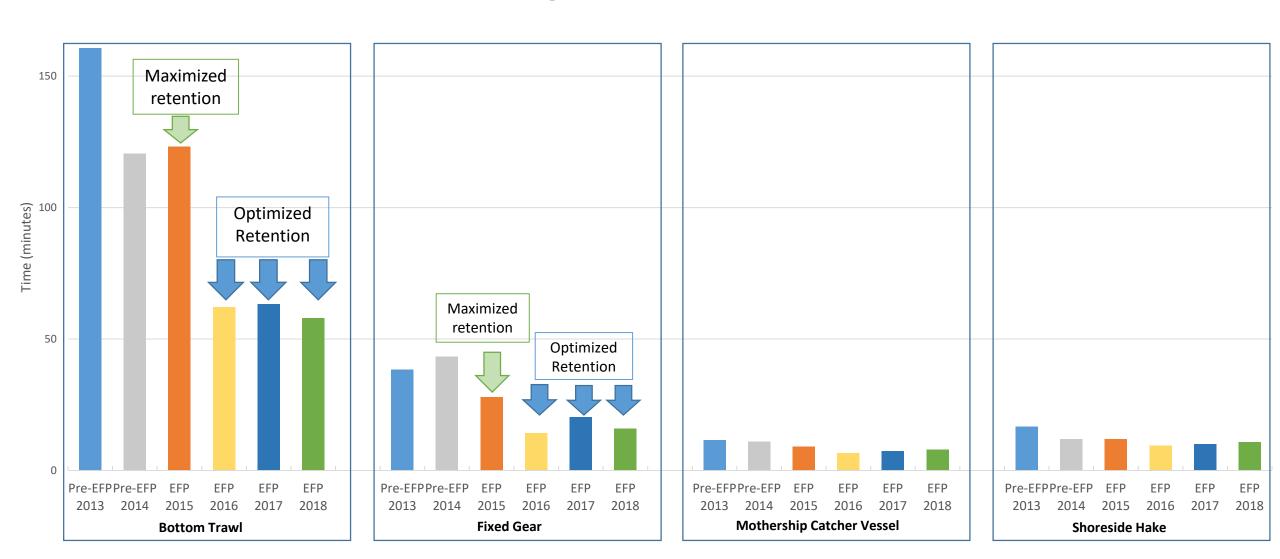
Optimized Retention (Fixed Gear)

- Allowed to discard:
- Everything in Maximized Retention
- All IFQ and non-IFQ species
- Must place all discards on length strip prior to discard
- Must retain salmon and eulachon



West Coast Review Rate Progression

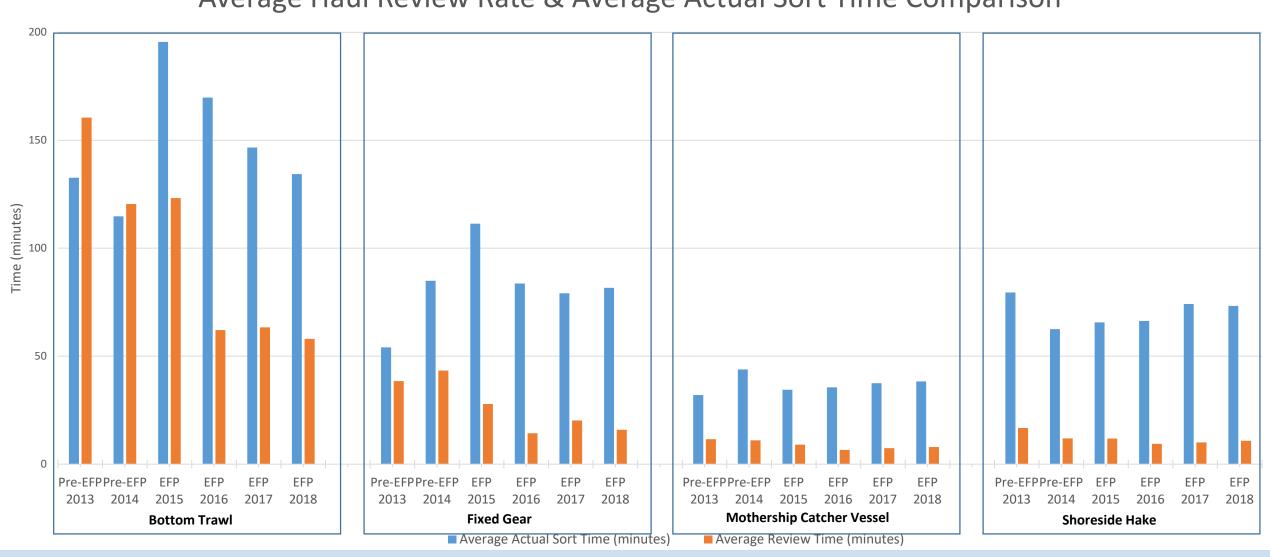
EFP vs. Pre-EFP: Average Haul Review Rate

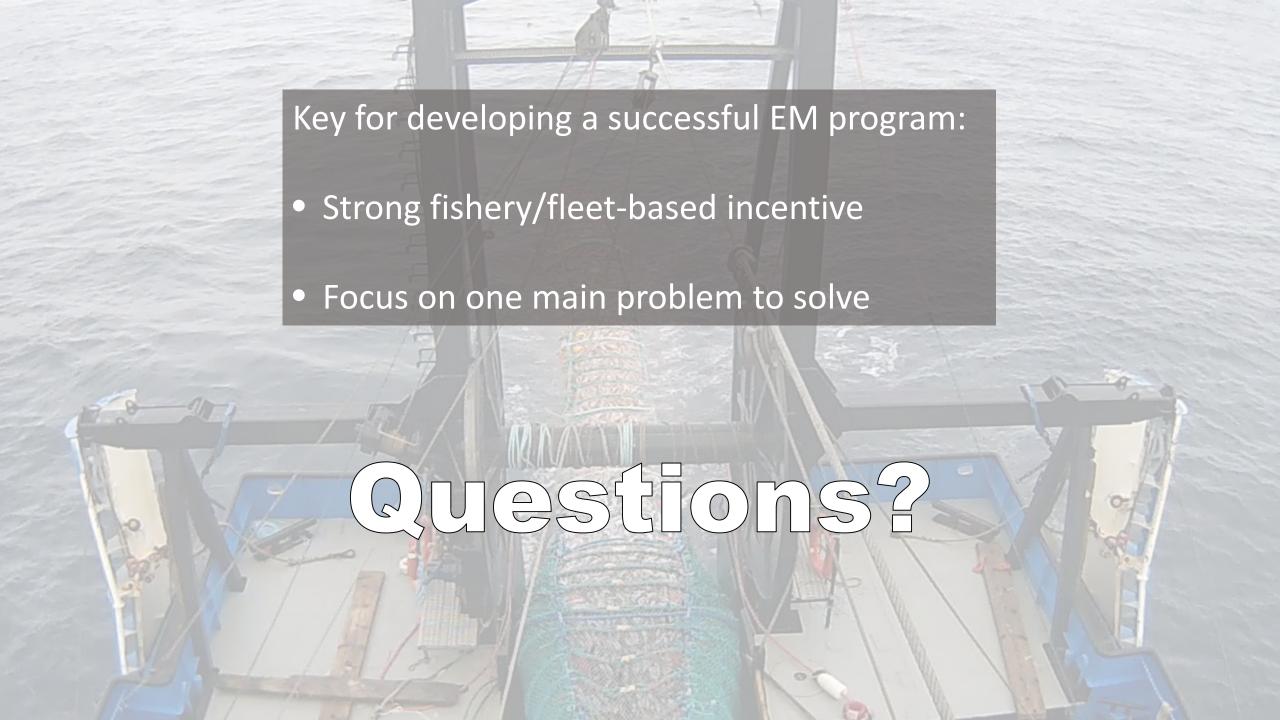


West Coast Review Rate Progression

EFP vs. Pre-EFP:

Average Haul Review Rate & Average Actual Sort Time Comparison





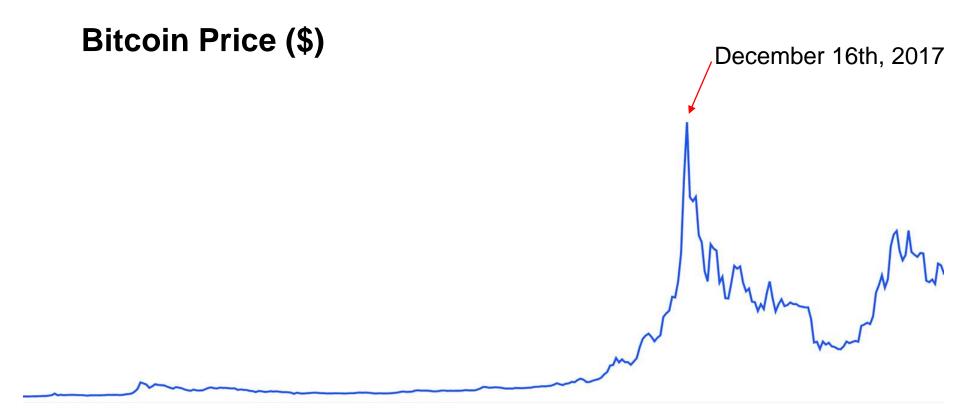


EM program design: Video review and data processing

Eric Pennaz ericpennaz@google.com

EM Video Storage Costs

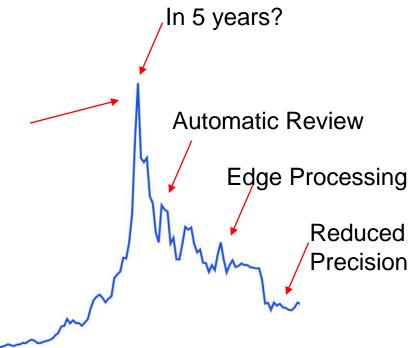


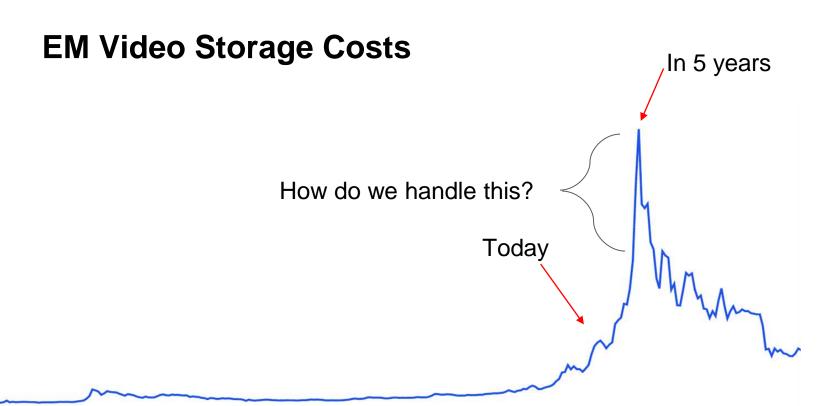




EM Video Storage Costs

- Label Everything
- Metadata is important
- Ultimately one federated model

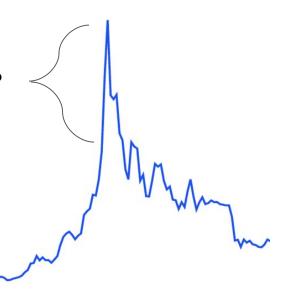




EM Video Storage Costs

How do we handle this?

- Economies of Scale / Type
- Cloud Price Reduction
- Collective Bargaining for Storage
- Data Lake



fishnet.ai

```
img_id,bbox_id,x_min,x_max,y_min,y_max,label_name
94a95d88-23f0-11e9-aef8-6377a4c78e35,1,392,607,197,405,Human
94a95d88-23f0-11e9-aef8-6377a4c78e35,2,315,444,132,260,Human
94a95d88-23f0-11e9-aef8-6377a4c78e35,3,228,353,36,183,Human
94a95d88-23f0-11e9-aef8-6377a4c78e35,4,166,331,286,470,Human
94a95d88-23f0-11e9-aef8-6377a4c78e35,5,331,377,359,513,Yellowfin tuna
94a95d88-23f0-11e9-aef8-6377a4c78e35,6,156,191,513,589,Yellowfin tuna
94a95d88-23f0-11e9-aef8-6377a4c78e35,7,531,605,0,71,Human
94a95d88-23f0-11e9-aef8-6377a4c78e35,8,50,150,95,217,Human
94a95d88-23f0-11e9-aef8-6377a4c78e35,9,0,51,226,306,Human
94ac6e38-23f0-11e9-b635-cb6efd3bfc7b,1,318,625,367,488,Black marlin
94ac6e38-23f0-11e9-b635-cb6efd3bfc7b,2,120,291,361,472,Human
```

