Science, Service, Stewardship

Estimating in-season discards from the Northeast United States groundfish fishery: an investigation of the separate ratio method (Part II)

Michael C. Palmer
Northeast Fisheries Science Center

A working paper in support of the Discard Estimation Methodology Review
Working Paper #3
February 22, 2010

NOAA FISHERIES SERVICE

Terms of Reference - Review of the Northeast Region Discard Estimation Methods

- Compare and contrast the performance with alternative estimators of total discards (November WS).

- Evaluate impacts of trimming observations (e.g. large discard events) on the magnitude of bias and measures of precision (P. Nitschke).

- **Examine the cumulative [yearly separate ratio] method and temporally stratified cumulative [separate ratio] method (with various time steps and discarding patterns) and recommend a preferred method with consideration of the following:**
  - **Within-year pattern of precision and bias, including a review of historical data used to estimate landings and discards (precision/accuracy).**
  - **Feasibility/practical aspects, particularly implications of stratum size for small sectors and the ability to obtain fine-scale measures of total discards and its uncertainty.**
  - **Estimate the probability of premature closure.**

- **Provide guidance of methods to measure uncertainty of the preferred method (e.g. asymptotic versus bootstrap estimates of variance).**

- **Provide guidance on risks of alternative management actions given the uncertainty and/or behavior of the preferred discard estimation method. In particular, consider the costs to the industry/sector from premature cessation of fishing and the risk to the resource from excess harvests.**

- **Consider implications of finer-scale stratification on performance of estimators that might be required for sector-specific discard rates and will be required for multi-stock species.**

How important are discards in overall quota monitoring?

- Discards are a minor component of the total species catch for most stocks.
  - Major component for those species with zero retention (ocean pout, windowpane flounder, SNE/MA winter flounder, etc.)
    - But…no in-season discard estimation (assumed rates only?)

Stock		Gear		
		Otter trawl dt ratio	**Gillent dt ratio**	**Longline dt ratio**
Atlantic cod	Gulf of Maine	0.22	0.09	0.30
	Georges Bank	0.16	0.10	0.15
Haddock	Gulf of Maine	0.03	0.08	0.05
	Georges Bank	0.12	0.07	0.06
Pollock		0.00	0.03	0.12
White hake		0.02	0.03	0.14
Acadian redfish		0.14	0.09	0.07
Yellowtail flounder	Cape Cod/Gulf of Maine	0.31	0.10	
	Georges Bank	0.10	0.82	
	Southern New England/mid-Atlantic	0.26	0.12	
Winter flounder	Gulf of Maine	0.12	0.03	
	George Bank	0.04	0.14	
American plaice		0.16	0.19	
Witch flounder		0.06	0.04	

WP#3, Table 5

- All of the analyses have assumed that landings are known with certainty.
  - Landings will not be known with certainty; there will be compliance problems, problems matching trips across data sources and the precautionary approach used by the stock apportionment methodology could lead to the double counting of species landings (and K_{all}) in these circumstances.
  - The magnitude of this problem and its effect on the precision of the landings and contribution to discards estimates and total catch is unknown.

How large will each discard strata be under Amendment 16 (stock, gear, sector)?

Sector Name	VTR gear code	Stock/region	Number of trips	Number of vessels
Fixed Gear Sector	OTF	SNE	1	1
NEFS 11	OTF	GBK	1	1
Port Clyde Community Groundfish Sector	OTF	GBK	1	1
Port Clyde Community Groundfish Sector	OTF	SNE	1	1
NCCS	OTF	GBK	3	1
NEFS 4	OTF	SNE	4	2
Tri-State Sector	OTF	SNE	5	2
NEFS 13	OTF	GOM	6	2
NEFS 2	OTF	SNE	6	5
NEFS 8	OTF	SNE	9	6
NEFS 4	OTF	GBK	18	4
NEFS 9	OTF	GOM	21	1
NEFS 8	OTF	GOM	22	1
Sustainable Harvest Sector	OTF	SNE	23	5
NEFS 5	OTF	GBK	24	10
NEFS 9	OTF	SNE	28	7
NEFS 10	OTF	GBK	29	8
NCCS	OTF	SNE	35	1
Common	OTF	GBK	48	8
NEFS 6	OTF	GOM	65	8
NEFS 6	OTF	GBK	71	8
NEFS 4	OTF	GOM	105	2
NEFS 10	OTF	SNE	105	5
Tri-State Sector	OTF	GBK	105	8
NEFS 2	OTF	GBK	111	17
Tri-State Sector	OTF	GOM	149	7
Port Clyde Community Groundfish Sector	OTF	GOM	149	12
NEFS 13	OTF	GBK	153	18
NEFS 7	OTF	GBK	178	13
NEFS 8	OTF	GBK	178	13
NEFS 12	OTF	GOM	230	4
NEFS 7	OTF	SNE	237	11
Common	OTF	GOM	278	13
Sustainable Harvest Sector	OTF	GBK	284	23
NEFS 9	OTF	GBK	286	19
NEFS 13	OTF	SNE	313	19
NEFS 11	OTF	GOM	487	11
NEFS 10	OTF	GOM	547	10
Sustainable Harvest Sector	OTF	GOM	630	26
NEFS 5	OTF	SNE	1436	30
Common	OTF	SNE	2149	68
NEFS 2	OTF	GOM	2917	41

WP#3, Table 4a

A discard simulator was developed in SAS based on Paul Nitschke's MS Excel work to further investigate the separate ratio method.

Capable of estimating discards using both the cumulative (yearly separate ratio) method and the temporally stratified separate ratio method on identical populations using identical set of observed trips.

Supports a variety of temporal stratifications (weekly, biweekly, monthly, quarterly) and computation frequencies (how frequently discards are estimated; daily, weekly, etc.).

Can be run for all groundfish stocks and major gear types (trawl, gillnet and longline).

Can control either the fleet size or the number of trips used in the simulation to examine estimator performance across the entire range of discard strata sizes likely to be observed under sectors/A16.

Can be run at any observer coverage rate.

Each simulation can be run x iterations.

Collects summary statistics for each iteration of the simulation and outputs a single table summarizing all iterations.

Seed values from each simulation/iteration are archived so runs can be reproduced at a later date for further investigation.

Can produce plots of individual simulation iterations to examine temporal trends in dk rates, discards and estimated catch.

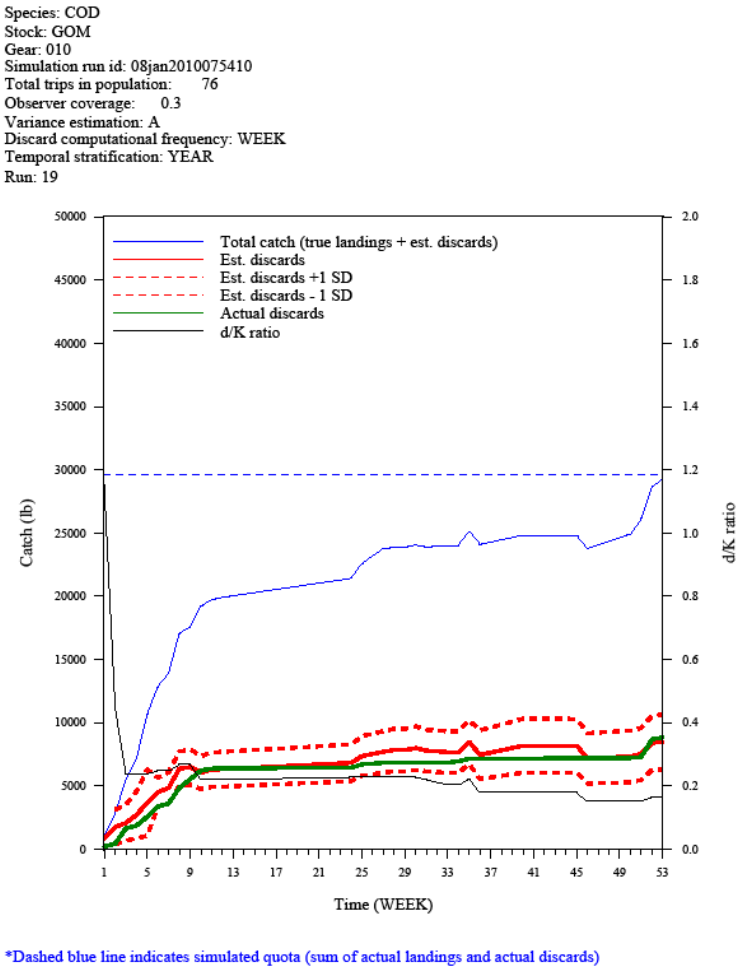
Each simulation can be reproduced using a bootstrap estimation of the discard point estimate and associated variance. For each run within a simulation, the bootstrap and analytical estimates can be compared.

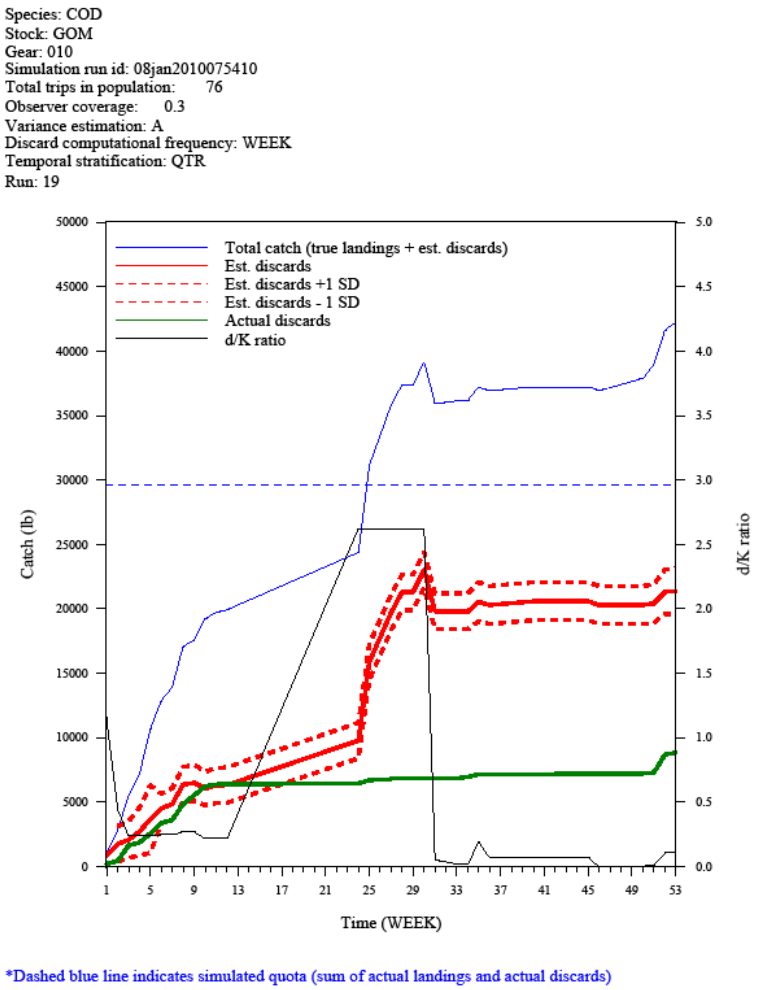
Discard estimates and variances from identical analytic and bootstrap simulations can be compared for a variety of metrics.

One of the concerns with the yearly separate ratio method was the temporal instability of the dk rate.

Is this a significant issue?

Would a temporally stratified estimator resolve concerns without compromising performance?





WP#3, Figure 1

Overview of data inputs to the simulation exercises

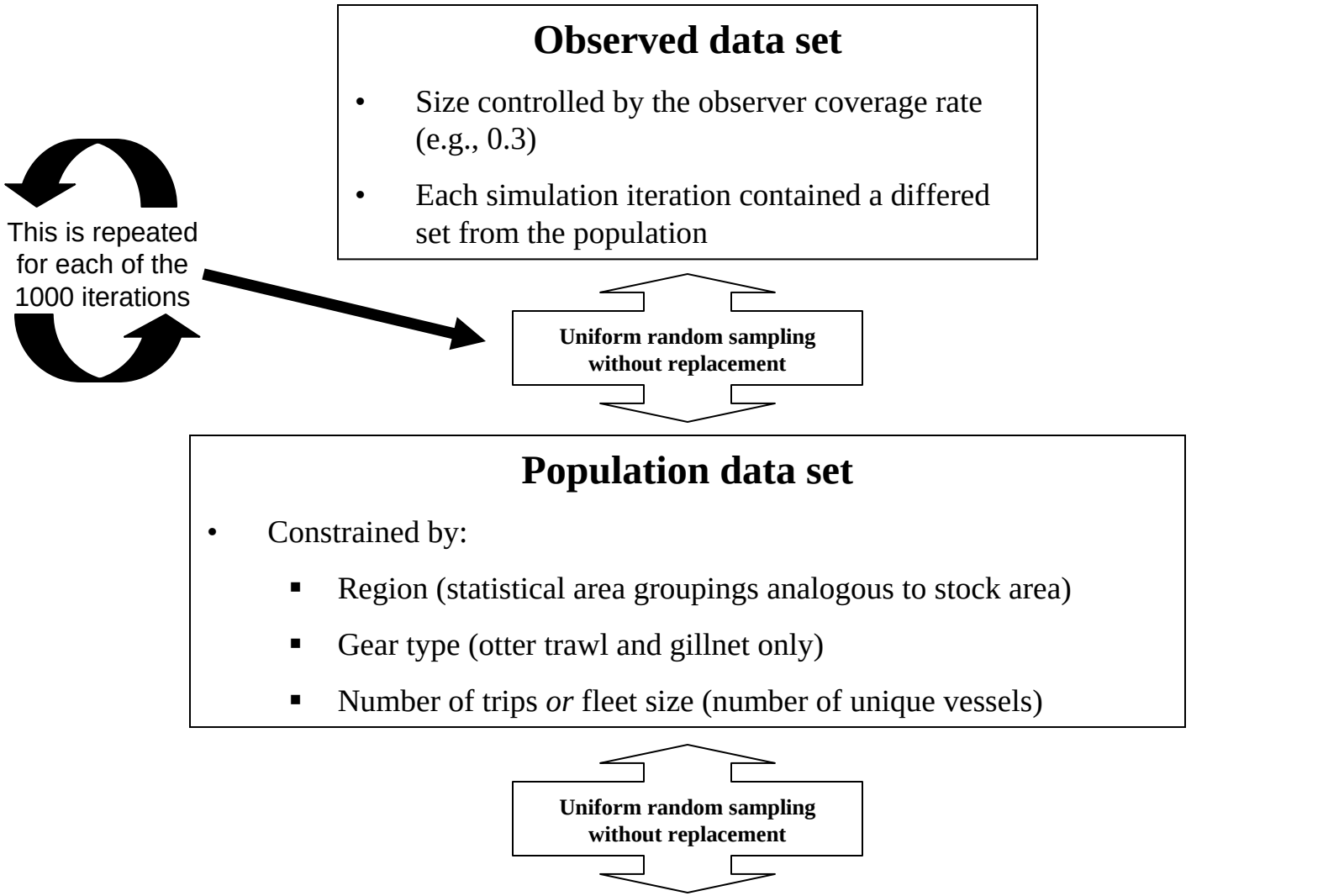
- Relied on conditioned SAS sets constructed by Wigley et al. to support SBRM and GARM discard estimation.
    - Contain haul-level discards for all species and major gear types (otter trawl, gillnet, longline, etc.) that can be rolled up to trip-level discards.
    - Simulations focused on:
        - The 14 stocks and 3 gear types for which in-season discards will need to be estimated in 2010.
        - Large mesh otter trawl (>6.0”), large mesh sink gillnet (>6.0”), benthic longline.

- Aggregated 5 years of observer data (2004 to 2008) and collapsed to a single year to construct a single year base data set (e.g. dropped year and reassigned to 2010).

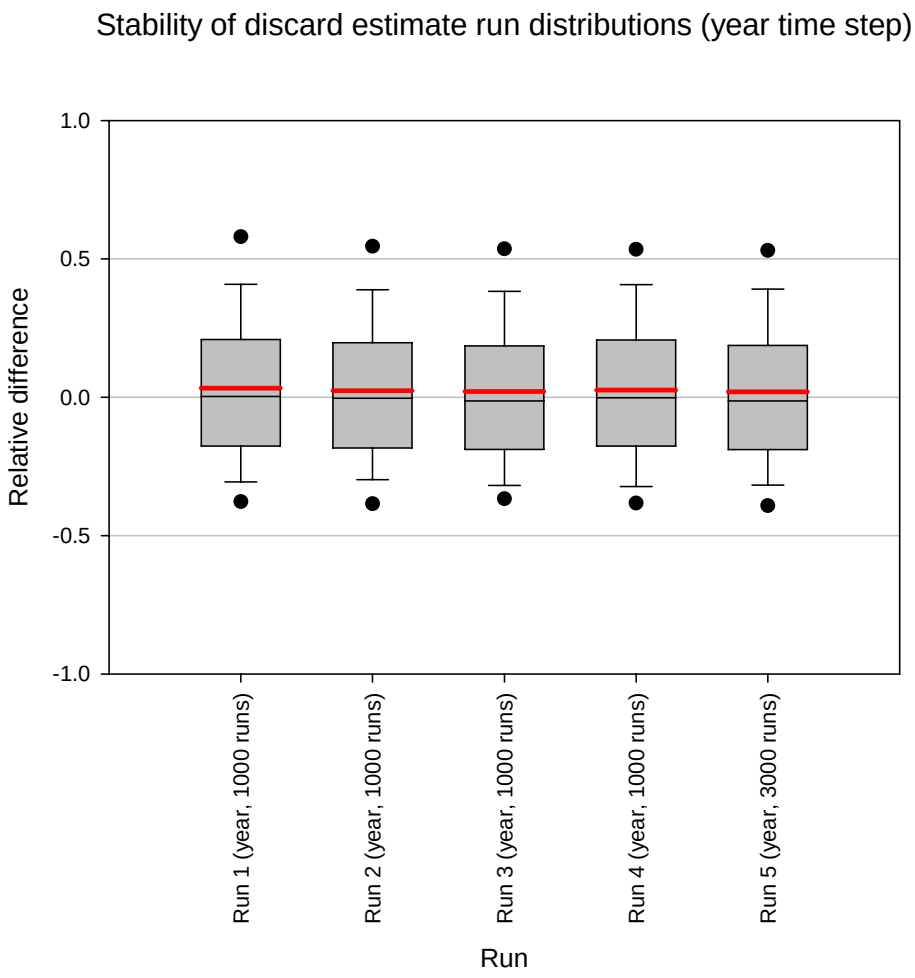
- For each simulation a population was constructed from the base data set and constrained by either the number of trips (fixed number of trips, number of vessels variable) or fleet size (fixed number of vessels, number of trips variable) and selecting a region (aggregation of statistical areas) and gear type.
    - Pulled from the base set using uniform random sampling without replacement.

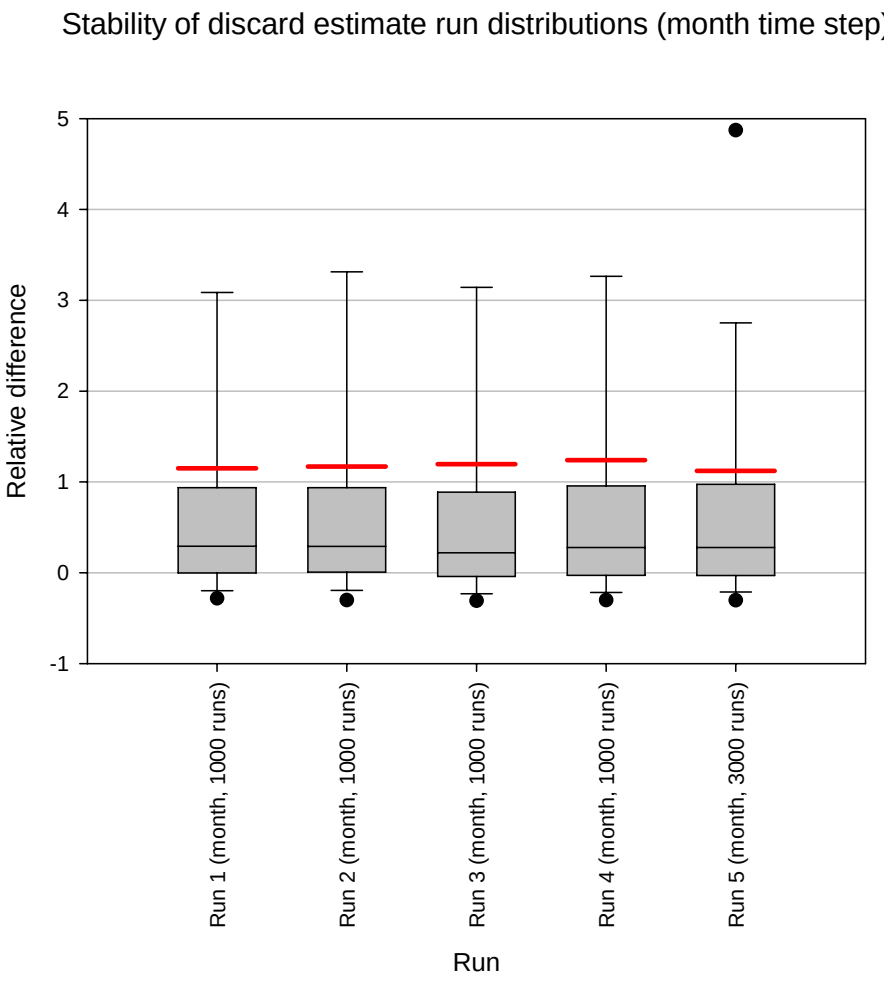
- An observed set was created from the population using a specified observed coverage level.
    - Held constant for most simulations at 0.3.
    - Uniform random sampling without replacement.
    - Each run within the simulation contained a different observed set from the population.

Overview of data inputs to the simulation exercises



Reproducibility of the simulations when run on the same population – testing the sufficiency of 1000 simulation iterations



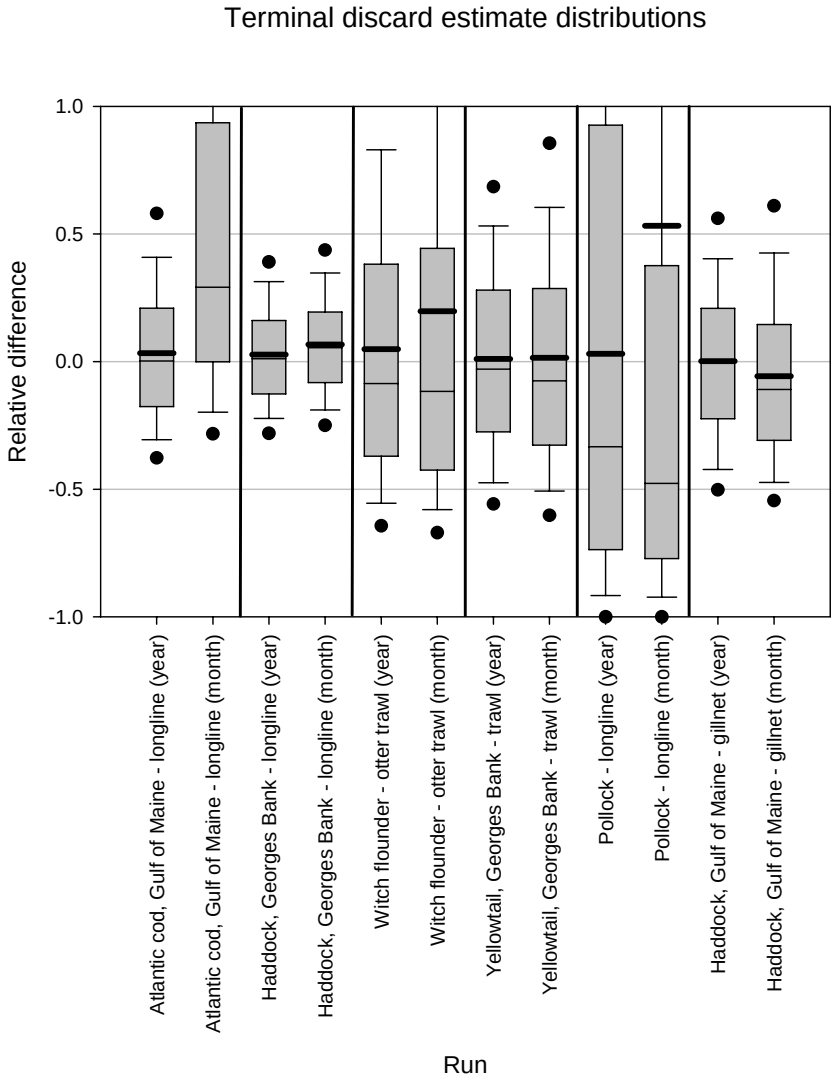


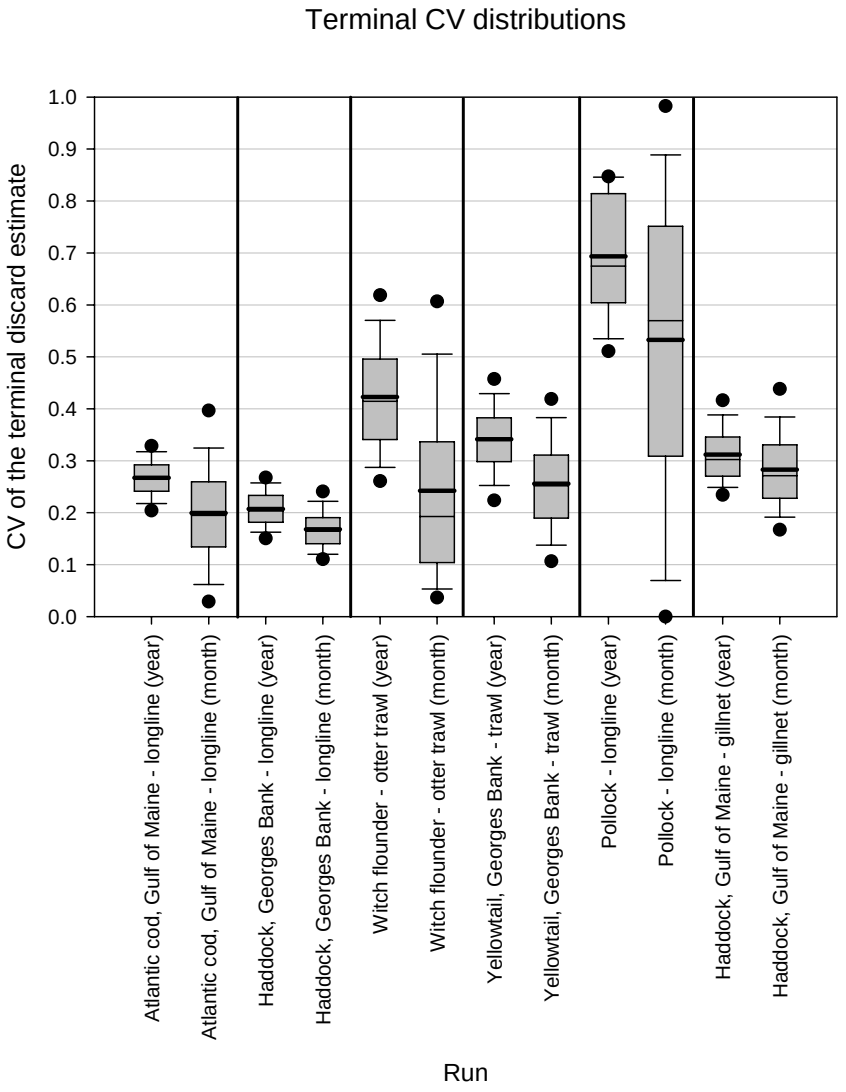
(cod estimation in the Gulf of Maine longline fishery)

WP#3, Figures 2 & 3

Impact of dk rate variability on stability of discard estimates

**Investigative runs*





fleet size of 15 vessels, population trips variable. 1000 simulations

WP#3, Figures 4, 5

Glossary of terms

- **Median relative difference** – the median value from 1000 simulation runs of the relative difference between the estimated discards and the true discard value (relative measure of median-bias).
- **Mean relative difference** - the mean value from 1000 simulation runs of the relative difference between the estimated discards and the true discard value (relative measure of mean-bias).
- **Fraction of runs +/- 5.0% difference** – The fraction of runs within +/- 5% relative difference of the true discards (measure of the probability that any single realization is correct).
- **Interquartile range of the terminal discard estimate** – The interquartile range of the relative differences from the 1000 simulation runs (measure of the spread of realizations – how wrong can the estimator be?).
- **Mean terminal CV** – The mean of the CV at the end of the year from each of the 1000 simulation runs (measure of the level of precision associated with the end of the year discard estimate).
- **Mean number of weeks with null dk** – The mean number of weeks when a dk ratio could not be computed (measure of how much imputation will be required).
- **Mean number of weeks with null variance** - The mean number of weeks when a variance could not be computed (measure of how much of the time series will be without 'real' estimates of precision).
- **Mean week when quota was first exceeded** – The mean week when the estimated discards exceeded the quota (assumes no uncertainty in the landings; measure of how bias/variability in the discard estimates will lead to premature closure of a fishery).
- **Mean number of weeks when discard estimates were adjusted down** – The mean number of weeks when the discard estimates in week t+1 < week t (measure how many 'down' corrections occurred).
- **Average change when discard estimate was adjusted down** – The average change relative to the terminal discard estimate when 'down' corrections occurred (measure of the average magnitude of the 'down' corrections).

Effect of the dk variability in the population and fleet size on estimator performance

**Note: grey shaded cells indicate optimal performance across methods (w/in fleet sizes), blue cells indicate ties*

dK CV level	Temporal stratification	Summary statistic	Fleet size		
			Small	Medium	Large
Low (haddock in the Georges Bank longline fishery)	Run summaries	Simulation runs	1000	1000	1000
		Total vessels	5	15	25
		Total trips	62	109	197
		Vessels observed	4.1	12.7	18.8
		Trips observed	19	33	59
		Days observed (dates of landing)	17.7	26.6	44.8
	Monthly	Median discard relative difference	-0.032	0.054	-0.013
		Mean discard relative difference	0.011	0.067	0.001
		Fraction of runs +/- 5.0% of true discards	0.151	0.188	0.253
		Interquartile range of the terminal relative difference	0.347	0.276	0.195
		Mean terminal CV	0.197	0.168	0.112
		Mean number of weeks with null DK	8.4	4.1	10.3
		Mean number of weeks with null variance	14.8	7.0	19.4
		Mean week when the 'quota' was first exceed	52.0	53.0	53.0
		Mean number of weeks when discard estimates were	1.6	0.6	1.9
		Average change when discard estimates were adjusted down (relative to terminal Discard estimate)	-0.0524	-0.0214	-0.0142
	Quarterly	Median discard relative difference	-0.061	0.054	-0.025
		Mean discard relative difference	-0.027	0.067	-0.011
		Fraction of runs +/- 5.0% of true discards	0.157	0.323	0.278
		Interquartile range of the terminal relative difference	0.309	0.264	0.183
		Mean terminal CV	0.203	0.186	0.123
		Mean number of weeks with null DK	4.9	3.5	4.9
		Mean number of weeks with null variance	10.2	5.7	10.5
		Mean week when the 'quota' was first exceed	52.1	53.0	53.0
		Mean number of weeks when discard estimates were adjusted down	2.4	1.2	2.4
		Average change when discard estimates were adjusted down (relative to terminal Discard estimate)	-0.040	-0.030	-0.012
	Year	Median discard relative difference	-0.010	0.011	-0.021
		Mean discard relative difference	0.018	0.027	-0.010
		Fraction of runs +/- 5.0% of true discards	0.153	0.190	0.278
		Interquartile range of the terminal relative difference	0.323	0.288	0.175
		Mean terminal CV	0.223	0.207	0.129
		Mean number of weeks with null DK	1.2	2.1	0.8
		Mean number of weeks with null variance	3.1	4.3	2.3
		Mean week when the 'quota' was first exceed	52.1	53.0	53.0
		Mean number of weeks when discard estimates were adjusted down	3.5	1.8	3.7
		Average change when discard estimates were adjusted down (relative to terminal Discard estimate)	-0.029	-0.044	-0.012

WP#3, Table 8a

Effect of the dk variability in the population and fleet size on estimator performance

**Note: grey shaded cells indicate optimal performance across methods (w/in fleet sizes), blue cells indicate ties*

dK CV level	Temporal stratification	Summary statistic	Fleet size		
			Small	Medium	Large
High (haddock in the Gulf of Maine gillnet fishery)	Run summaries	Simulation runs	1000	1000	1000
		Total vessels	5	15	25
		Total trips	35	145	365
		Vessels observed	3.8	11.0	20.2
		Trips observed	11	44	110
		Days observed (dates of landing)	10.8	40.3	89.0
	Monthly	Median discard relative difference	-0.399	0.026	0.006
		Mean discard relative difference	-0.222	0.127	0.046
		Fraction of runs +/- 5.0% of true discards	0.037	0.085	0.140
		Interquartile range of the terminal relative difference	0.845	0.687	0.420
		Mean terminal CV	0.250	0.325	0.274
		Mean number of weeks with null DK	6.0	5.9	3.1
		Mean number of weeks with null variance	12.9	14.2	7.4
		Mean week when the 'quota' was first exceed	46.2	51.1	50.3
		Mean number of weeks when discard estimates were adjusted down	0.8	3.7	6.1
		Average change when discard estimates were adjusted down (relative to terminal Discard estimate)	-0.285	-0.048	-0.025
	Quarterly	Median discard relative difference	-0.195	0.037	0.005
		Mean discard relative difference	-0.021	0.085	0.021
		Fraction of runs +/- 5.0% of true discards	0.038	0.084	0.145
		Interquartile range of the terminal relative difference	1.041	0.584	0.378
		Mean terminal CV	0.574	0.369	0.266
		Mean number of weeks with null DK	3.0	2.0	1.2
		Mean number of weeks with null variance	7.1	5.2	3.2
		Mean week when the 'quota' was first exceed	45.4	51.4	50.6
		Mean number of weeks when discard estimates were adjusted down	1.7	6.8	8.5
		Average change when discard estimates were adjusted down (relative to terminal Discard estimate)	-0.230	-0.045	-0.017
	Year	Median discard relative difference	-0.026	-0.001	0.004
		Mean discard relative difference	-0.018	0.033	0.014
		Fraction of runs +/- 5.0% of true discards	0.063	0.099	0.130
		Interquartile range of the terminal relative difference	1.048	0.531	0.378
		Mean terminal CV	0.631	0.378	0.277
		Mean number of weeks with null DK	0.4	0.3	0.1
		Mean number of weeks with null variance	1.6	0.9	0.5
		Mean week when the 'quota' was first exceed	44.6	51.5	50.6
		Mean number of weeks when discard estimates were adjusted down	1.9	8.3	10.8
		Average change when discard estimates were adjusted down (relative to terminal Discard estimate)	-0.082	-0.034	-0.020

WP#3, Table 8b

Effect of the dk variability in the population and fleet size on estimator performance

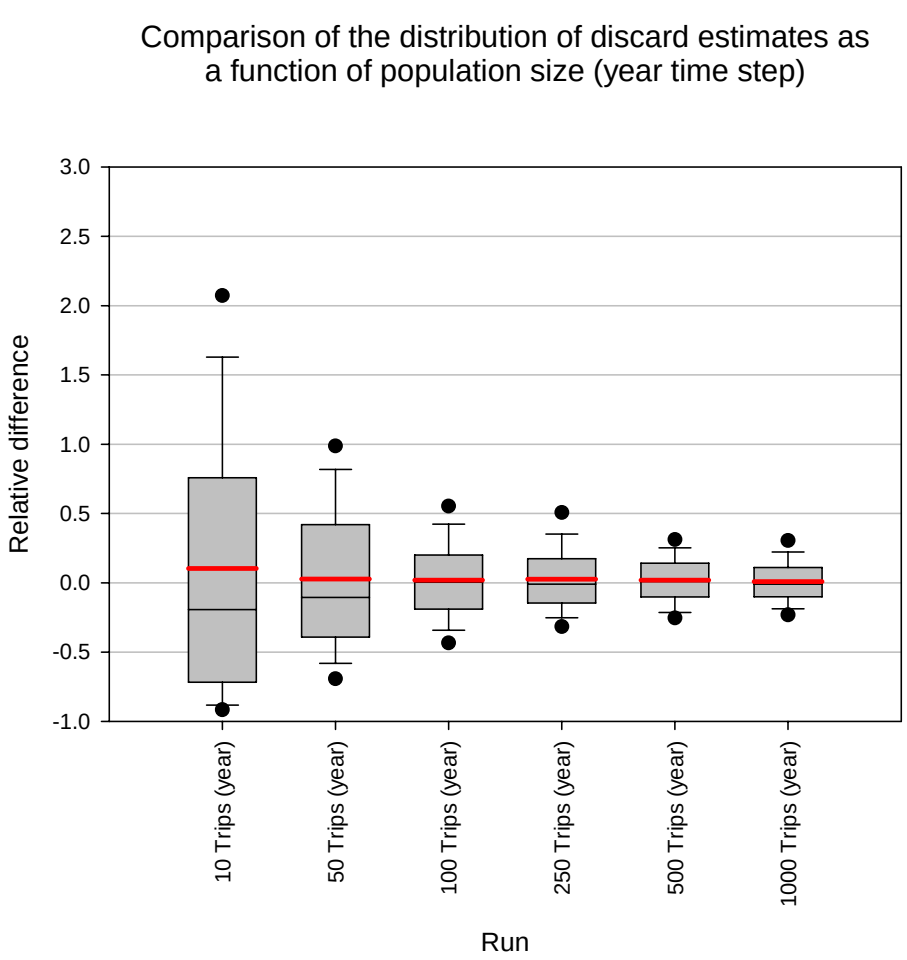
**Note: grey shaded cells indicate optimal performance across methods (w/in fleet sizes), blue cells indicate ties*

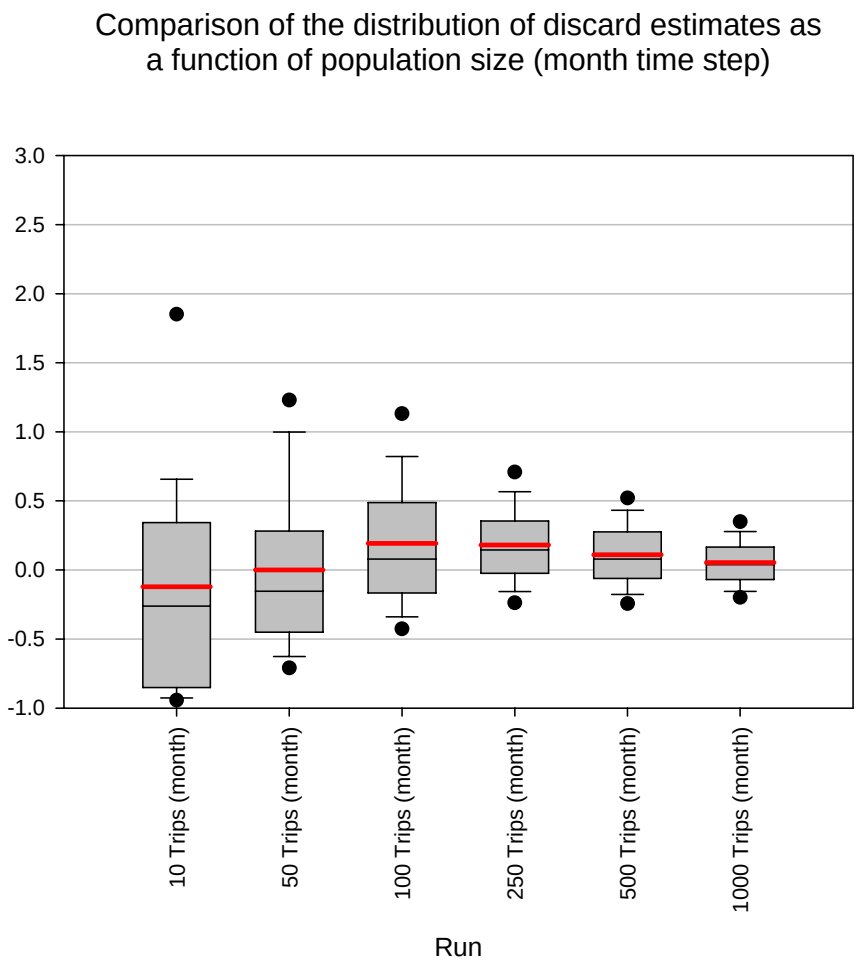
dK CV level	Temporal stratification	Summary statistic		Fleet size		
				Small	Medium	Large
High (witch flounder in the trawl fishery)	Run summaries	Simulation runs		1000	1000	1000
		Total vessels		5	15	25
		Total trips		16	100	170
		Vessels observed		3.2	11.2	16.2
		Trips observed		5	30	51
		Days observed (dates of landing)		4.9	29.1	47.3
	Monthly	Median discard relative difference		0.764	-0.011	0.003
		Mean discard relative difference		3.224	0.083	0.037
		Fraction of runs +/- 5.0% of true discards		0.012	0.097	0.116
		Interquartile range of the terminal relative difference		5.686	0.554	0.396
		Mean terminal CV	n/a		0.186	0.202
		Mean number of weeks with null DK		6.8	14.2	7.3
		Mean number of weeks with null variance		11.1	29.9	17.8
		Mean week when the 'quota' was first exceed		39.4	51.1	49.9
		Mean number of weeks when discard estimates were adjusted down		0.3	5.1	6.3
		Average change when discard estimates were adjusted down (relative to terminal Discard estimate)		-3.759	-0.092	-0.029
	Quarterly	Median discard relative difference		0.831	-0.005	-0.017
		Mean discard relative difference		2.970	0.006	0.008
		Fraction of runs +/- 5.0% of true discards		0.015	0.100	0.144
		Interquartile range of the terminal relative difference		5.226	0.488	0.369
		Mean terminal CV		0.307	0.288	0.233
		Mean number of weeks with null DK		4.8	5.0	2.6
		Mean number of weeks with null variance		9.7	11.7	6.7
		Mean week when the 'quota' was first exceed		39.6	51.4	50.2
		Mean number of weeks when discard estimates were adjusted down		0.5	7.6	9.0
		Average change when discard estimates were adjusted down (relative to terminal Discard estimate)		-0.142	-0.061	-0.024
	Year	Median discard relative difference		-0.030	-0.017	-0.010
		Mean discard relative difference		0.726	-0.004	0.009
		Fraction of runs +/- 5.0% of true discards		0.077	0.128	0.152
		Interquartile range of the terminal relative difference		1.115	0.481	0.348
		Mean terminal CV		0.603	0.310	0.238
		Mean number of weeks with null DK		1.4	0.6	0.1
		Mean number of weeks with null variance		3.6	1.7	0.4
		Mean week when the 'quota' was first exceed		40.3	51.4	50.3
		Mean number of weeks when discard estimates were adjusted down		0.7	9.1	9.1
		Average change when discard estimates were adjusted down (relative to terminal Discard estimate)		-0.133	-0.031	-0.015

WP#3, Table 8c

How does the number of trips in each population (sector strata) affect estimator performance?

**Note: each pull represents a different population*





(haddock estimation in the Gulf of Maine gillnet fishery)

WP#3, Figures 6 & 7

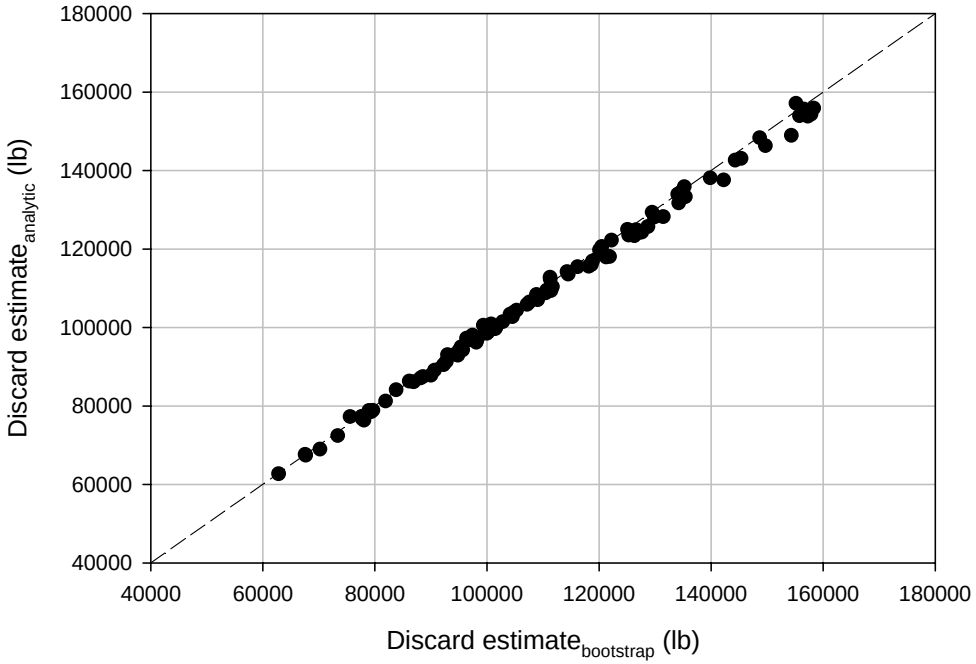
How does the computational frequency (daily vs. weekly) affect estimator performance?

Stock	Computational frequency of the discard estimate	Summary statistic	Temporal stratification		
			Month	Quarter	Year
White hake longline fishery	Run summaries	Simulation runs			1000
		Total vessels			50
		Total trips			100
		Vessels observed			22.8
		Trips observed			30
		Days observed (dates of landing)			27.3
	Daily	Median discard relative difference	0.034	-0.005	-0.018
		Mean discard relative difference	0.090	0.021	0.003
		Fraction of runs +/- 5.0% of true discards	0.083	0.094	0.108
		SD discard relative difference	0.500	0.404	0.408
		Mean terminal CV	0.358	0.366	0.385
		Mean number of days with null DK	15.4	7.5	2.0
		Mean number of days with null variance	31.6	16.8	4.5
		Number of runs when the 'quota' was exceeded	597	510	488
		Mean day when the 'quota' was first exceed	354.1	354.9	355.0
		Mean number of days when discard estimates were adjusted down	7.4	8.6	10.0
		Average change when discard estimates were adjusted down (relative to terminal Discard estimate)	-0.066	-0.023	-0.014
	Weekly	Median discard relative difference	0.034	-0.005	-0.018
		Mean discard relative difference	0.090	0.021	0.003
		Fraction of runs +/- 5.0% of true discards	0.083	0.094	0.108
		SD discard relative difference	0.500	0.404	0.408
		Mean terminal CV	0.358	0.366	0.385
		Mean number of weeks with null DK	8.9	4.3	0.6
		Mean number of weeks with null variance	16.8	9.7	1.4
		Number of runs when the 'quota' was exceeded	570	497	475
		Mean week when the 'quota' was first exceed	51.9	51.9	51.9
		Mean number of weeks when discard estimates were adjusted down	2.1	3.0	3.9
		Average change when discard estimates were adjusted down (relative to terminal Discard estimate)	-0.117	-0.039	-0.023

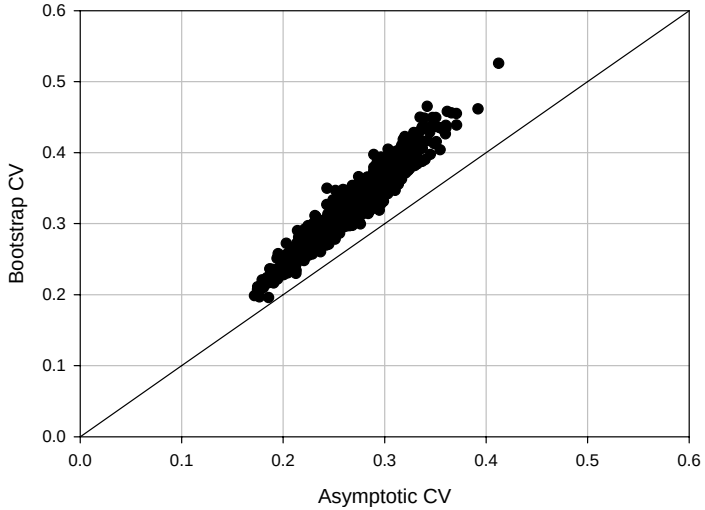
WP#3, Table 9

Estimates of variance: asymptotic vs. bootstrap

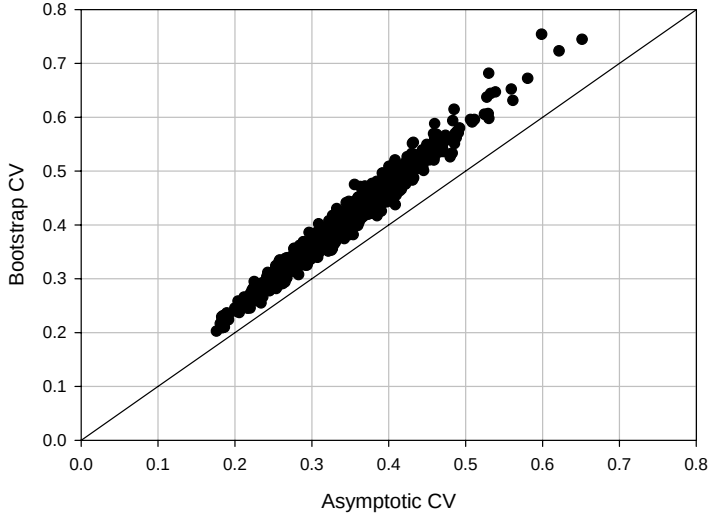
- Bootstrap variance calculated by sampling with replacement (500 times) from the observed set for each simulation iteration.



Gulf of Maine cod longline fishery (08jan2010075410)



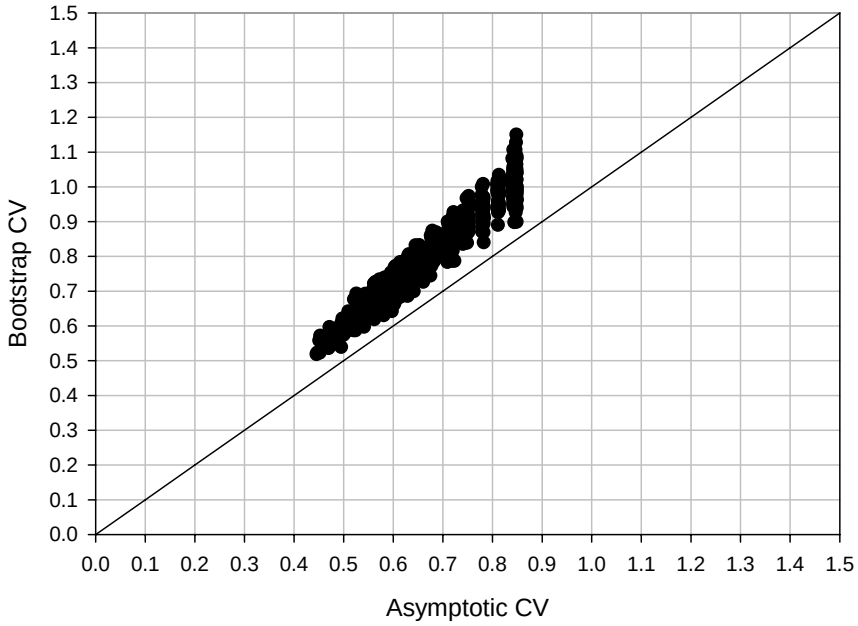
Georges Bank yellowtail trawl flounder fishery (08jan2010083900)

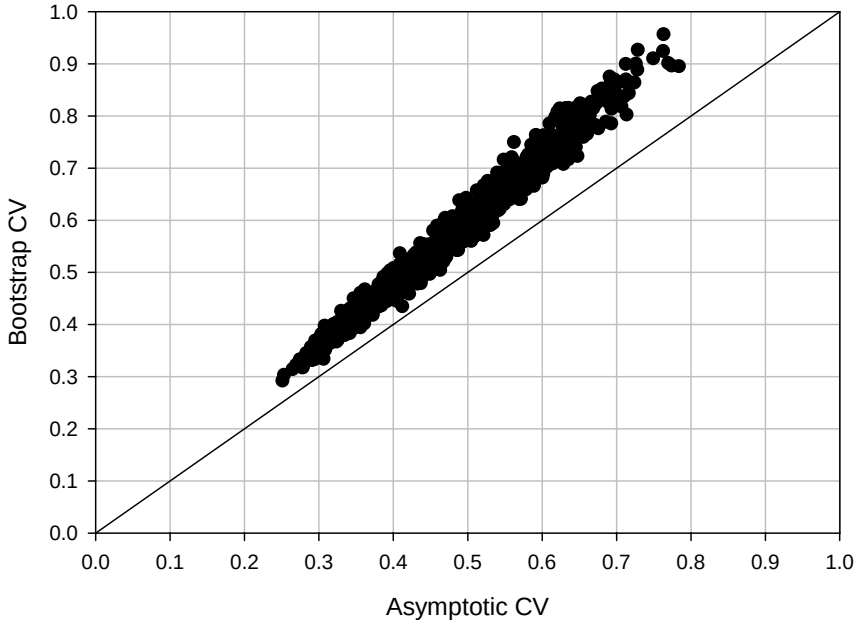


WP#3, from Figures 8 & 9

Estimates of variance: asymptotic vs. bootstrap (back pocket slide)

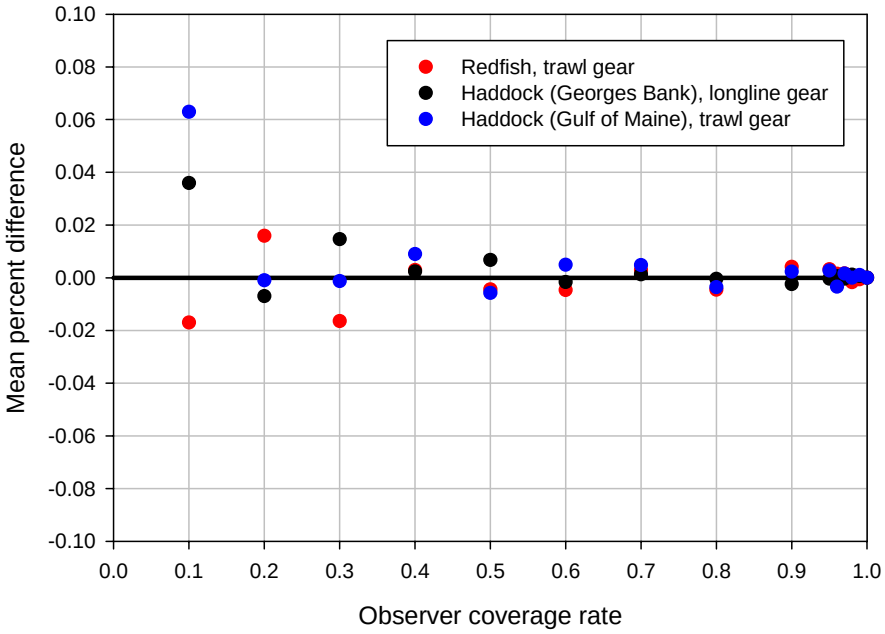
- Effects of low sample size on asymptotic variance estimates

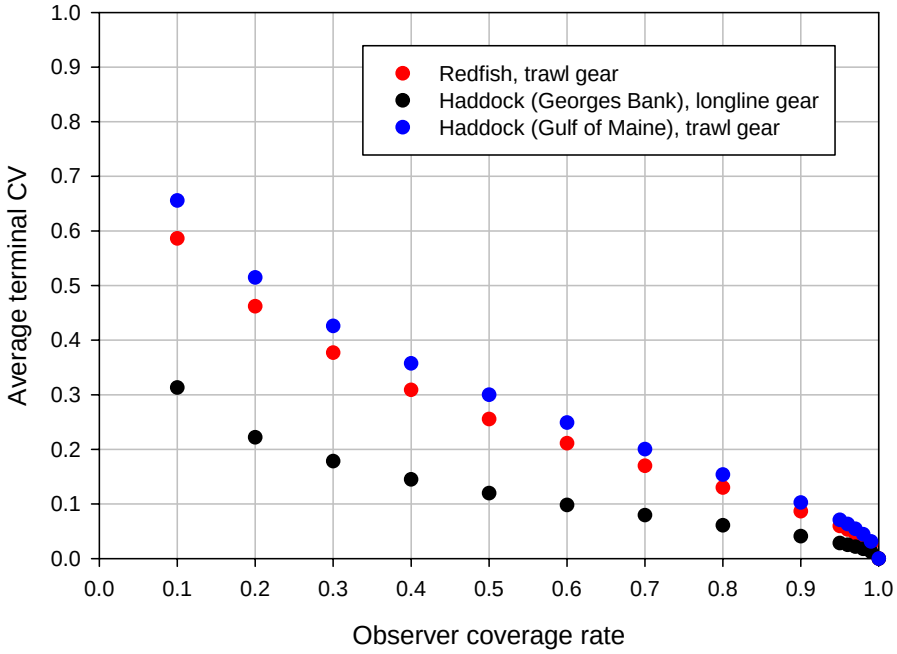




WP#3, Figures 9c & d

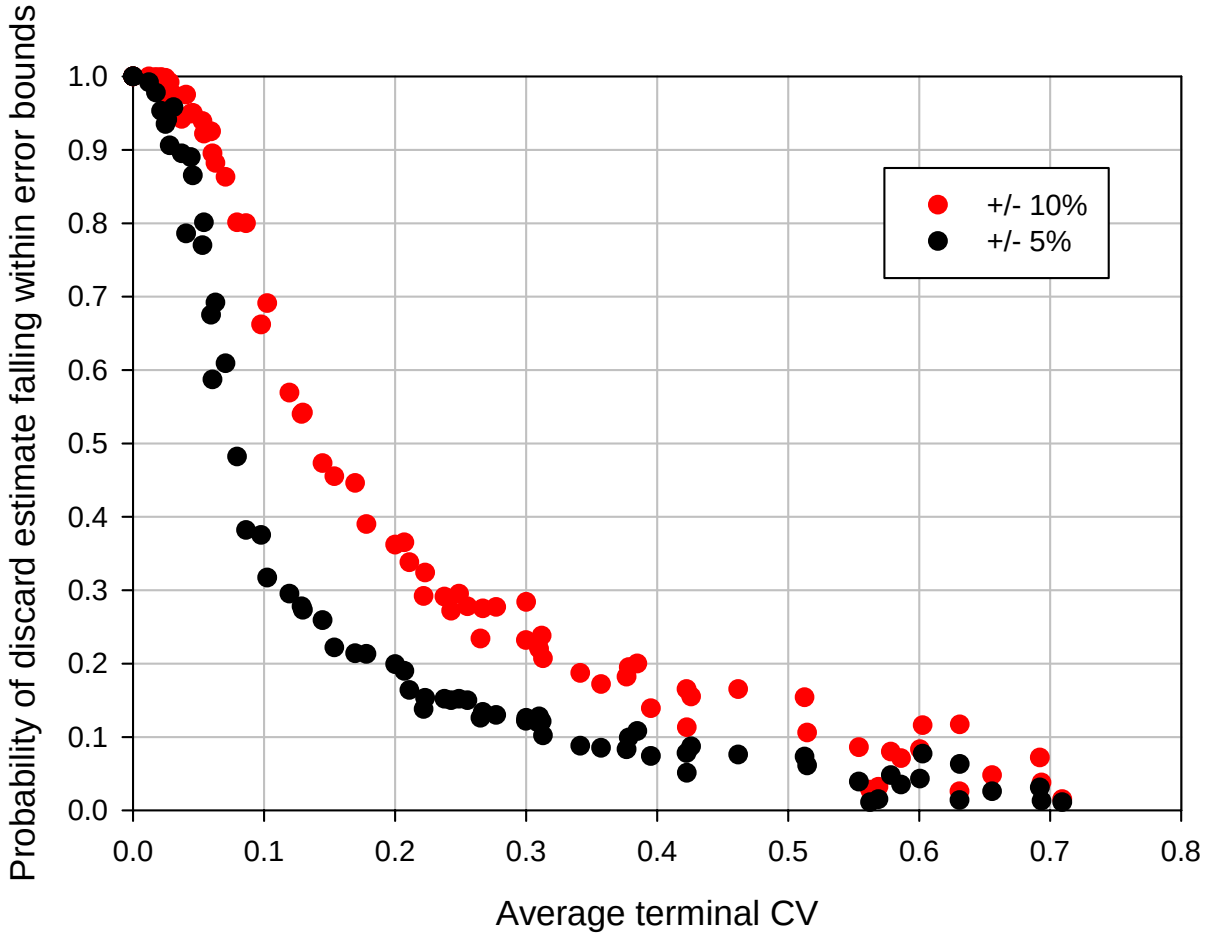
How do observer coverage rates affect the accuracy and precision of discard estimates?





WP#3, Figures 10 & 11

Can the CV of the terminal discard estimate indicate anything about the accuracy of an unbiased discard estimate?



WP#3, Figure 12

Summary

- The temporally stratified method is susceptible to estimation bias when applied to small strata and/or strata with high variability in the dk ratios.
    - The temporally stratified method provides more precise estimates, however when applied to small strata the estimates of precision may be artificially low.

- The yearly separate ratio method (a.k.a., cumulative) appears to be the most robust of the discard estimators examined.
    - Temporal variability in the dk ratios does not necessarily translate to large variability in the discard patterns.

- Analytic (asymptotic) methods of estimating variance are slightly biased (low), but the bias problems are small relative to the scale of the precision. The analytical method may be sufficient for providing uncertainty information needed to inform management decisions.

- Uncertainty in estimating discards may complicate ACE monitoring; however, the extent will depend not only on the uncertainty in the discard estimate, but also the contribution of discards to the overall ACE accounting (variable by sector and stock).