A Description of the Allocation Procedure

Presented by

Susan Wigley
Operations Research Analyst
Susan.Wigley@noaa.gov

NEFSC Data Review
August 5-9, 2013

Outline

I. Overview of Area Allocation Procedure

II. Measures of Uncertainty in Trip-based Allocated Landings

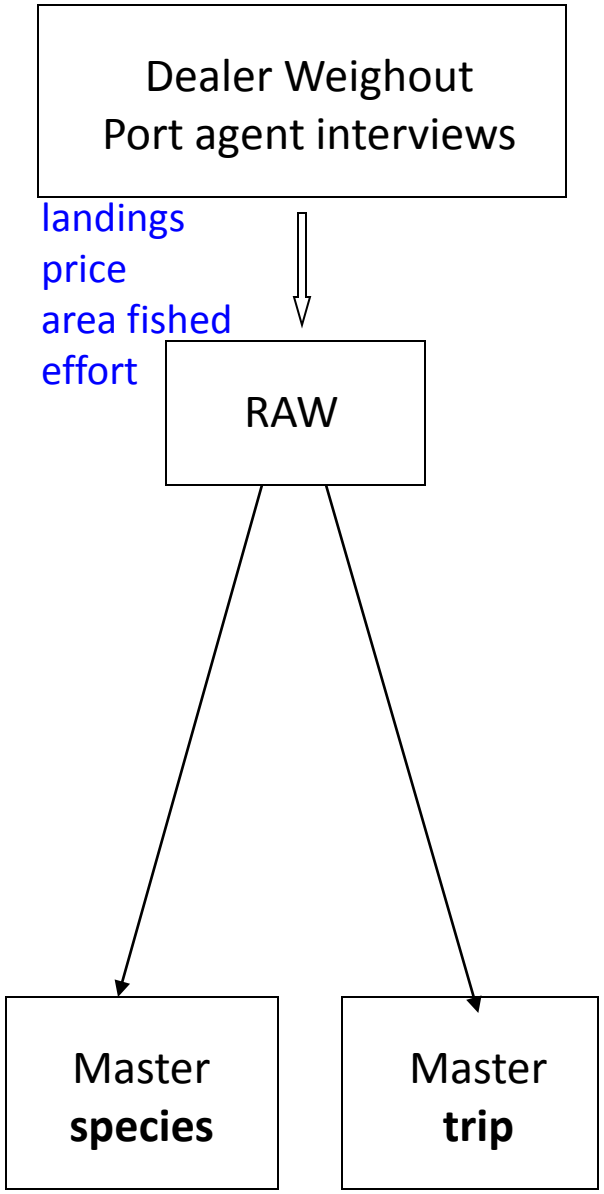
III. Investigative Analyses

IV. Proposed Solutions

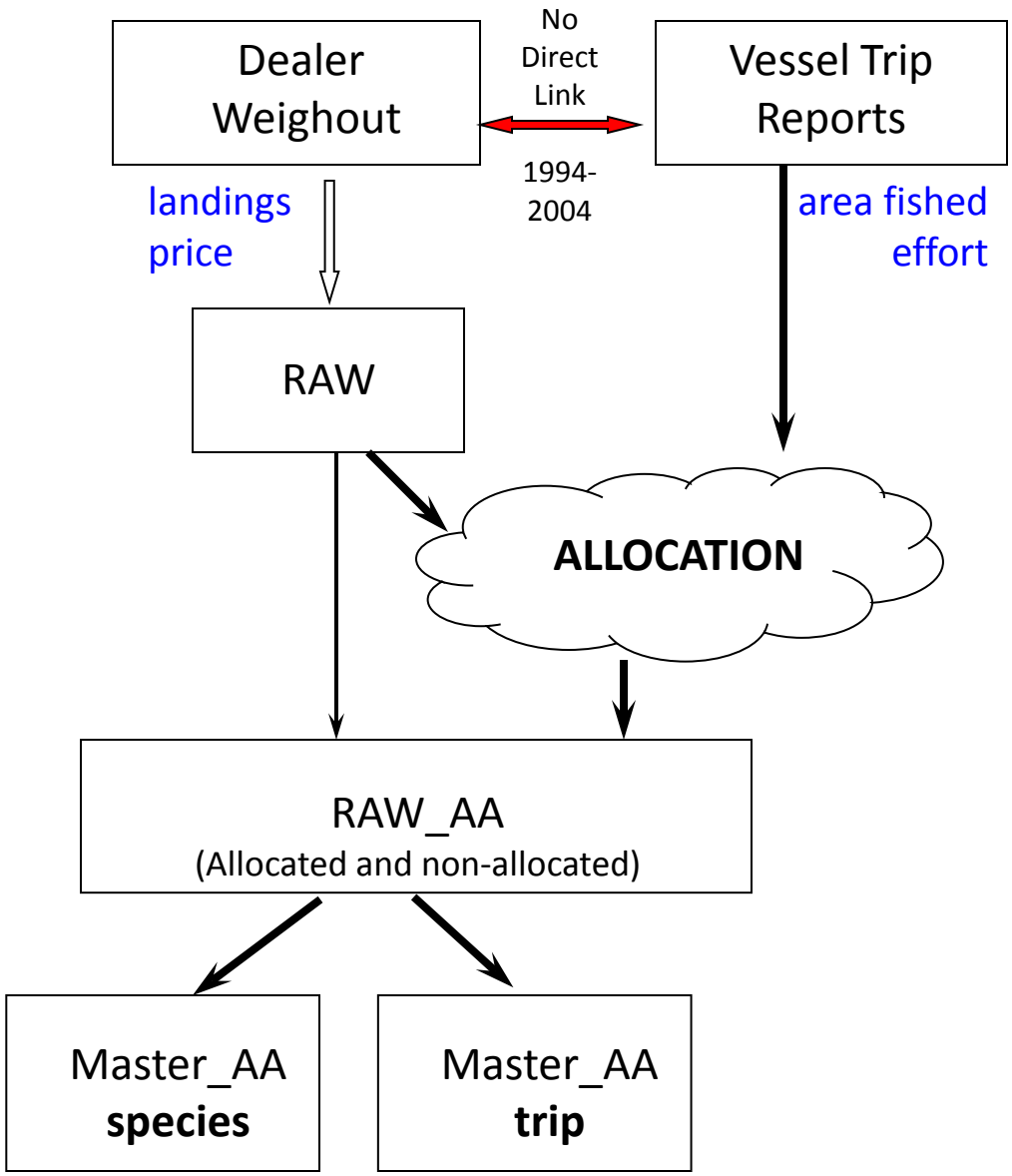
I. Area Allocation--Background

- Landings by statistical areas are used to define stocks
- Historically, this was done by port agents who interviewed fishermen
- In 1994 mandatory self-reported log books were introduced.
- Two reporting systems
  - Logbooks (vessel trip reports or VTR)
  - Dealer Reports
- Only an indirect link between these databases
- Methods were needed to link records for use in stock assessments

Pre-1994 Data
"Weigh Out Databases"



1994-present Data
"AA=Area Allocation"



Allocation

Purpose:
To supplement the mandatory commercial landings data with area fished and effort information using the Vessel Trip Report (VTR) data.

Goal:
To eliminate the need for single species prorations for each analysis conducted; and

To maintain a consistent, comprehensive, commercial landings database from 1963 – present containing the information (landings, value, area and effort) needed to address management questions, conduct stock assessments, meet NAFO reporting requirements, and support economic analyses and ecosystem research.

Allocation Design

Trip-based A Dealer trip (and all associated species landings) is assigned area fished and effort based upon VTR data.

Multi-tier A Dealer trip will be matched to VTR data at one of 4 levels:

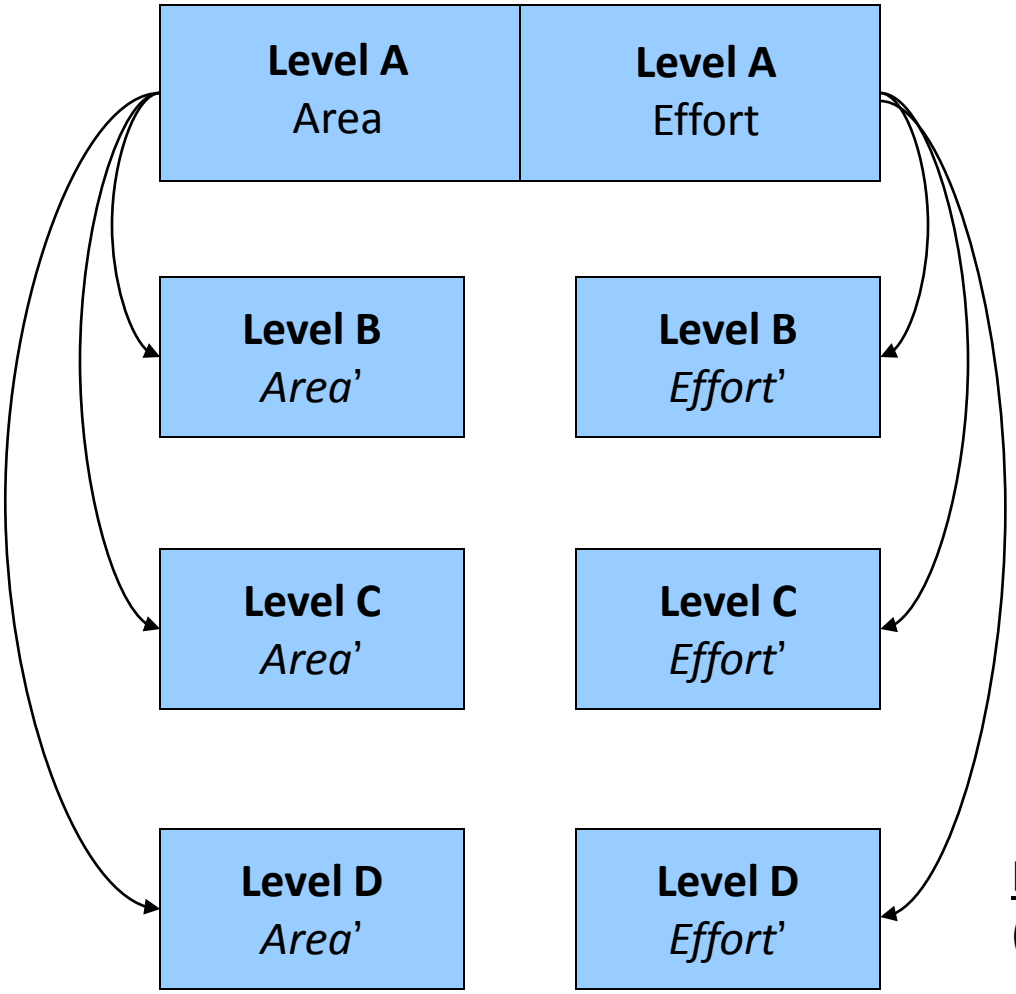
LEVEL A: 1:1 match
Dealer trip matches a VTR trip based upon permit-month-day

LEVEL B: 1: vessel match
Dealer trip matches a group of VTR trips for the same vessel within a month, species group, and gear type

LEVEL C: 1: Fleet match
Dealer trip matches a group of VTR trips for the same fleet: ton class, port group, gear group, species group, and quarter

LEVEL D: 1: Fleet match
Dealer trip matches a group of VTR trips for the general fleet: port group

VESSEL TRIP REPORT DATA



Individual trips/subtrips:
permit, month, day

Vessel: permit, gear group, species group, month
(*plus* area *for Effort matching*)

Fleet: ton class, port group, gear group, species group, quarter
(*plus* area *for Effort matching*)

Fleet: port group
(*plus* area *for Effort matching*)

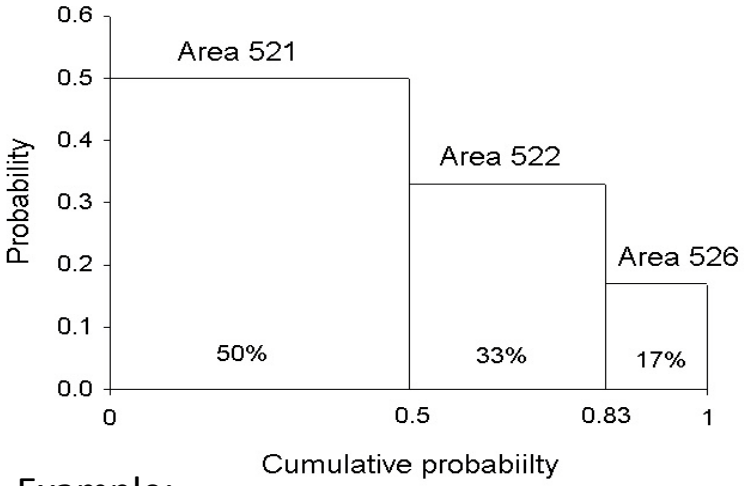
Estimation of Area' at Levels B, C & D

A statistical area is assigned to a Dealer trip on a probabilistic basis by sampling (with replacement) the distribution of unique areas within the stratification cell of a Level.

For each group of VTR trips (including subtrips), a discrete cumulative probability density function is formed based on the number of trips and unique areas in the cell.

Each Dealer trip has been randomly assigned a value between 0 and 1. This value is compared with the cumulative probabilities within the cell, when RAN <= CUMPROB, the area associated with that cumulative probability is assigned to the Dealer trip.

No split trips at Level B, C, and D



Example:
6 trips fished in 3 unique areas
(3 trips in 521, 2 trips in 522 and 1 trip in 526)

	prob.	cum prob.
Area 521 (3 trips / 6 trips)	0.50	0.50
Area 522 (2 trips / 6 trips)	0.33	0.83
Area 526 (1 trip / 6 trips)	0.17	1.00

If a Dealer trip had a random number = 0.75, then this Dealer trip would be assign Area 522.

On average, 50% of the Dealer trips that match this cell would be assigned Area 521, 33% would be assigned Area 522 and 17% would be assigned Area 526

Estimation of Effort' at Levels B, C and D

Total effort is not known; effort is estimated.

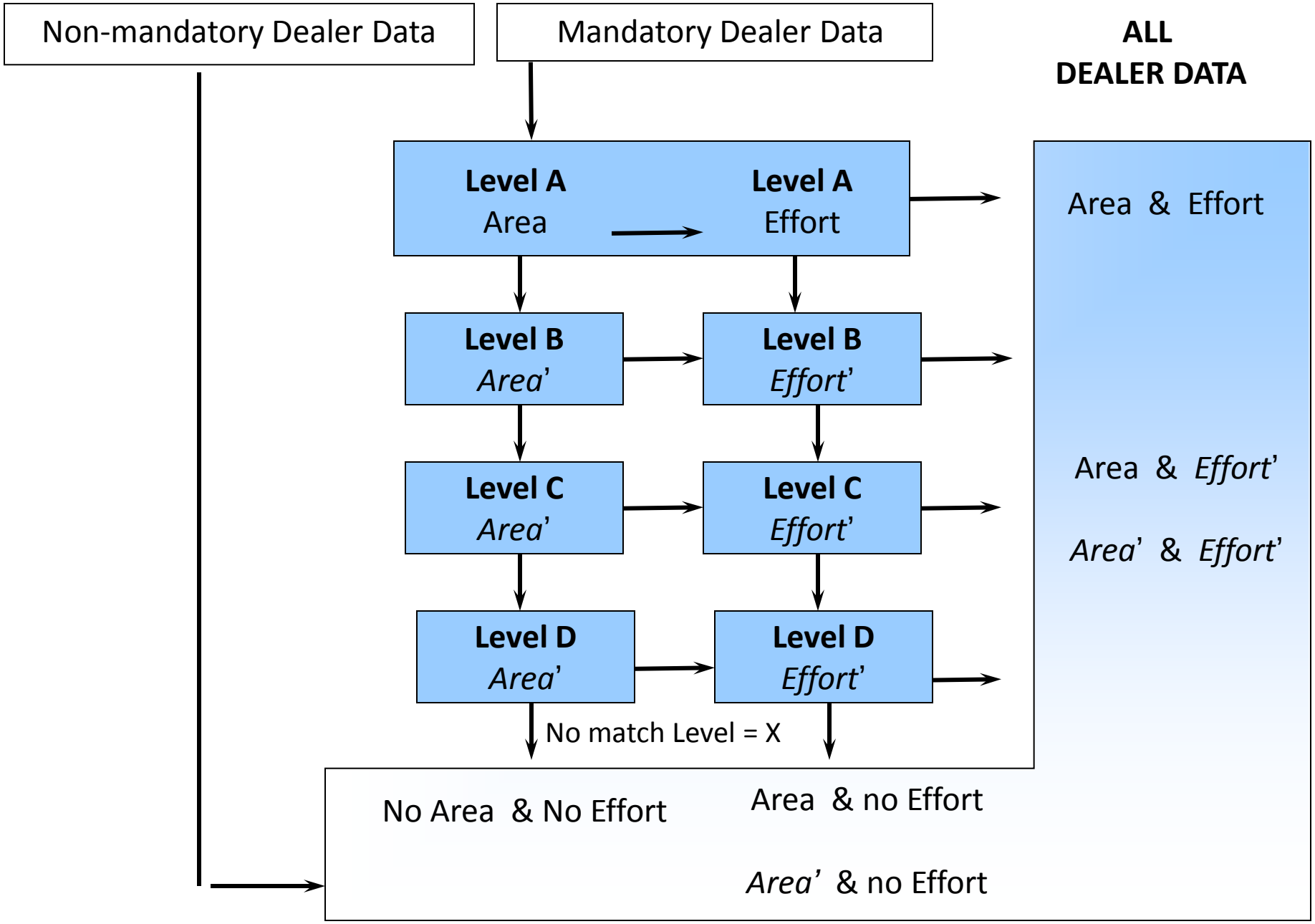
Days fished (DF) and days absent (DA), although correlated, are independent of each other and each can be related to area fished.

Within a cell, the median days fished and median days absent are derived.

The median was selected as the simplest statistic of central tendency for distributions of various shapes.

1st and 3rd quartiles are also calculated (for cells with 4 or more observations),
the quartiles provides a measure of dispersion:
semi-interquartile range or quartile deviation: (Q3 – Q1)/2

At LEVEL B, C, & D, effort is assigned to one Dealer trip
(no split trips at Level B,C, & D)


Non-mandatory Dealer Data
Mandatory Dealer Data
ALL DEALER DATA
Level A Area
Level A Effort
Area & Effort
Level B Area'
Level B Effort'
Level C Area'
Level C Effort'
Area & Effort'
Area' & Effort'
Level D Area'
Level D Effort'
No match Level = X
No Area & No Effort
Area & no Effort
Area' & no Effort

Allocation meta fields

- *Alevel* = stores the level at which area was obtained
  (A, B, C or D)

- *Elevel* = stores the level at which effort was obtained
  (A, B, C, D, or X)

Allocation checks

- Counters are in place at Level A, B, C, & D to track which cells are used and how often each cell is utilized.
  At Level A, no cell is used twice!

Matching statistics of the Allocation 1994 – 2006

…and updated through 2012

Example: 1995

Dealer Data	mt	% mt	% allocated mt
before Allocation	757964.9		
after Allocation	757965.3		
Difference	0.4		
Non-Allocated data	572974.6	75.6%	
Allocated data	184990.7	24.4%	
AREA match			
Level A	106473.0	14.0%	57.6%
Level B	39078.9	5.2%	21.1%
Level C	31730.3	4.2%	17.2%
Level D	7708.5	1.0%	4.2%
Allocated data	184990.7	24.4%	100.0%
subject to random component		10.4%	42.4%
EFFORT match			
Level A	91930.4		49.7%
Level B	35391.5		19.1%
Level C	21887.3		11.8%
Level D	28721.8		15.5%
Level X	7059.7		3.8%
Allocated data	184990.7		100.0%

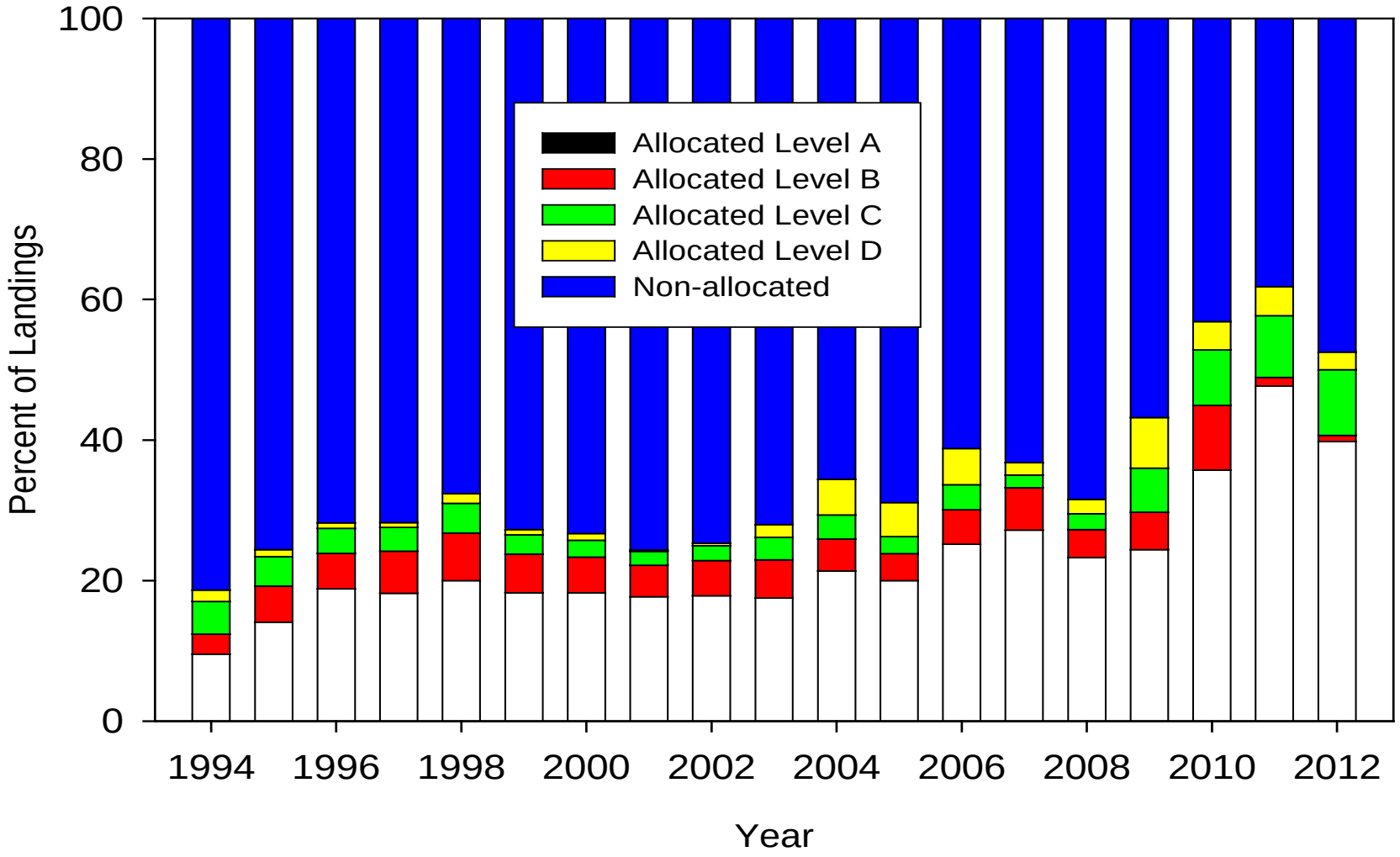
some rounding occurs; < 1 mt

< 39% of annual landings enter allocation (varies by species); recent years higher

most landings match at the vessel level (A&B);
a funnel pattern is present for Area (most mt at Level A, fewest mt at Level D)

Average 11% of annual landings are subject to the random component

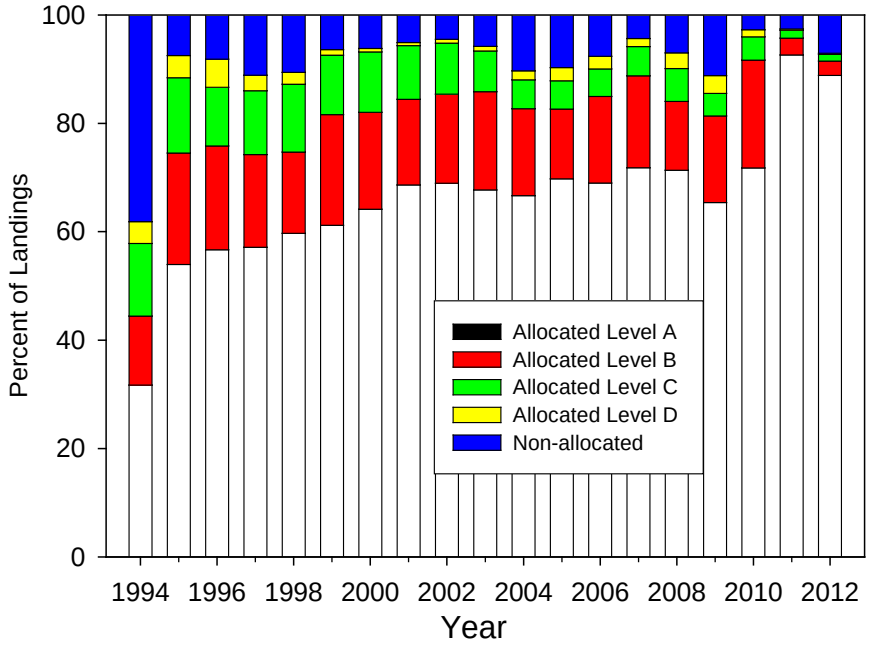
Matching statistics of the Allocation 1994 – 2006 and updated through 2012



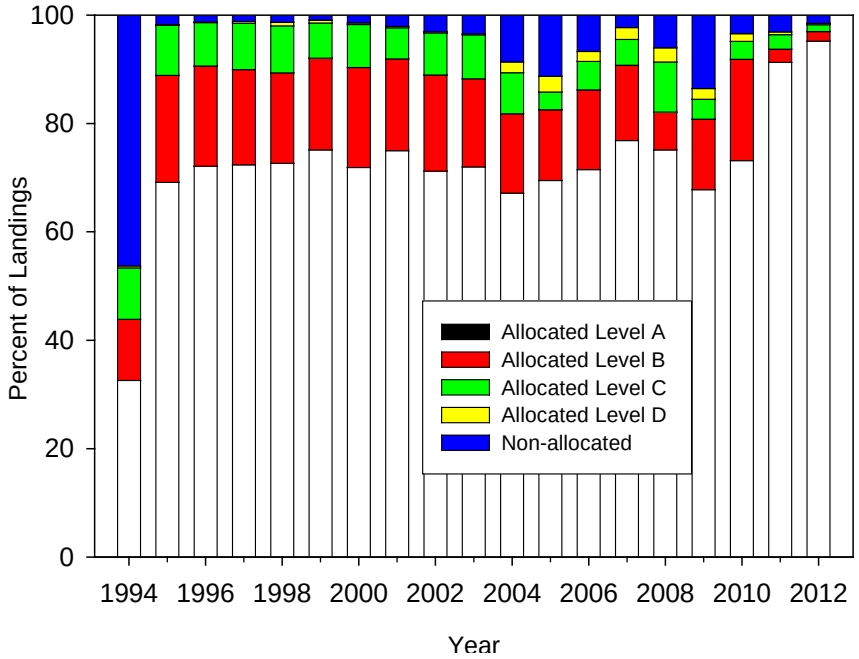
< 39% of the total Dealer metric tons enters the allocation; varies by species

Update: increase percentage in recent years

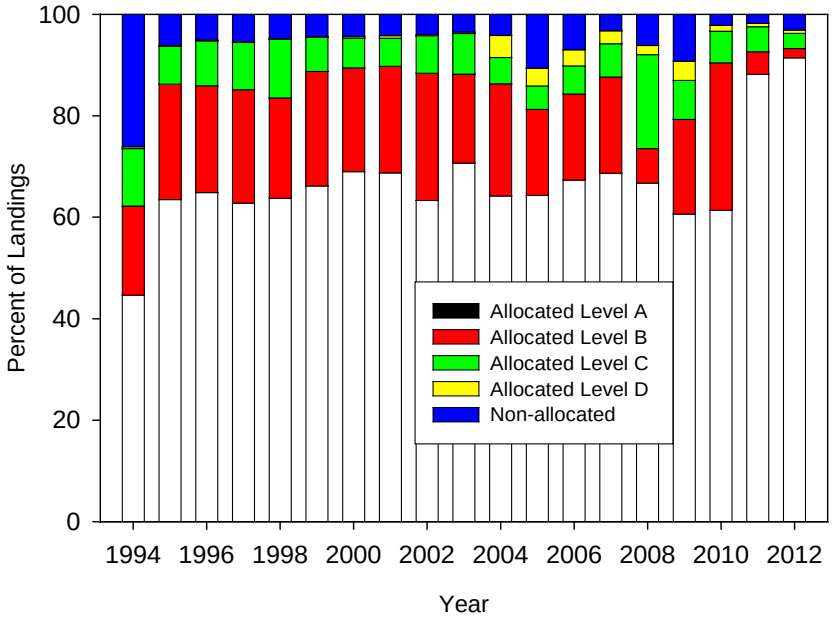
MONKFISH



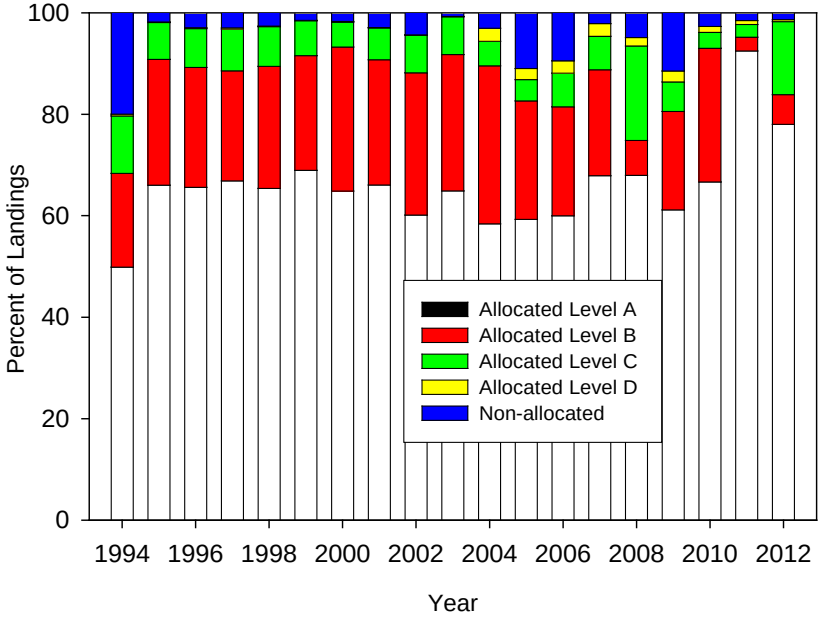
COD



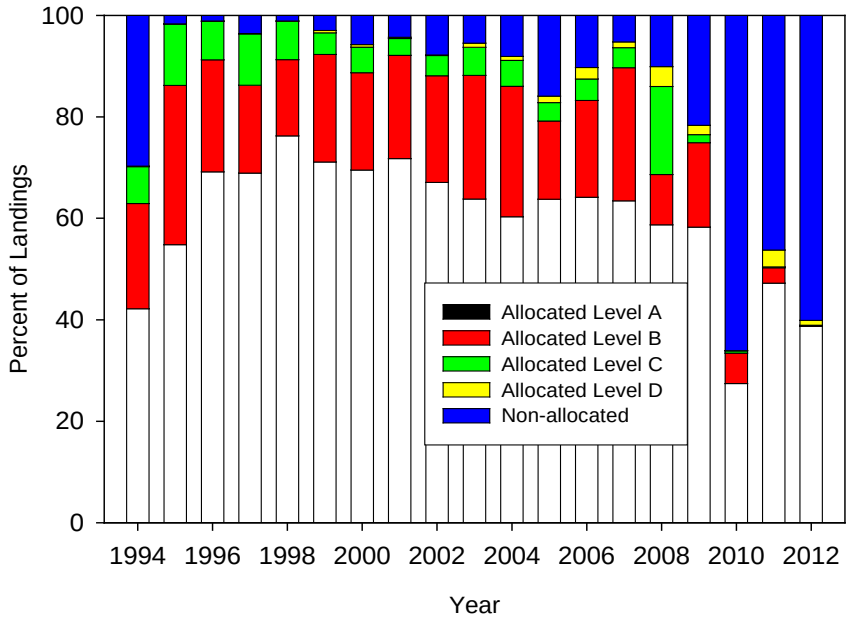
WINTER FLD



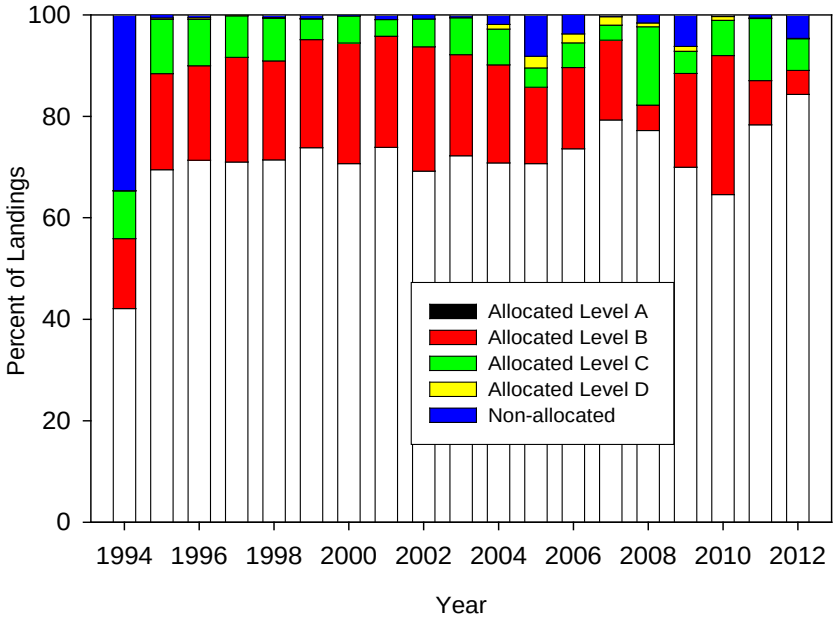
YELLOWTAIL FLD



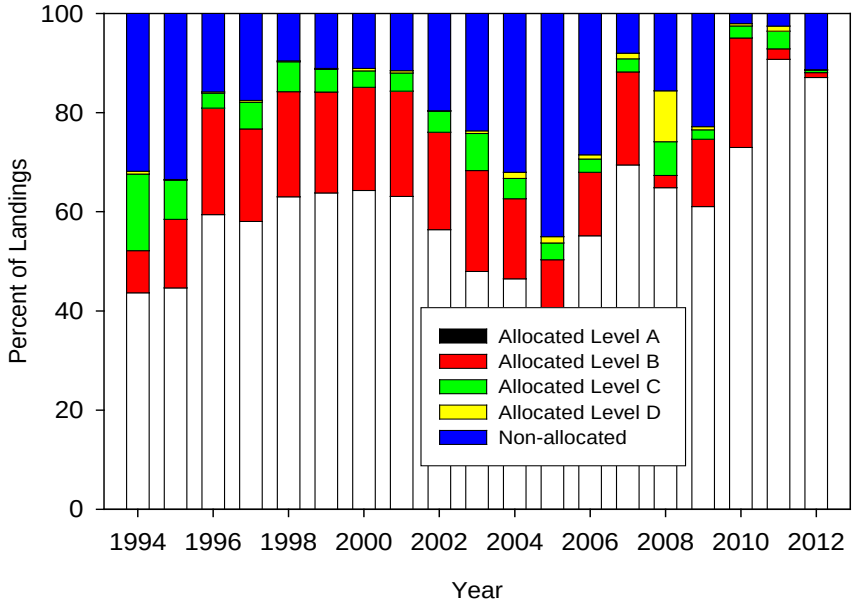
WINDOWPANE



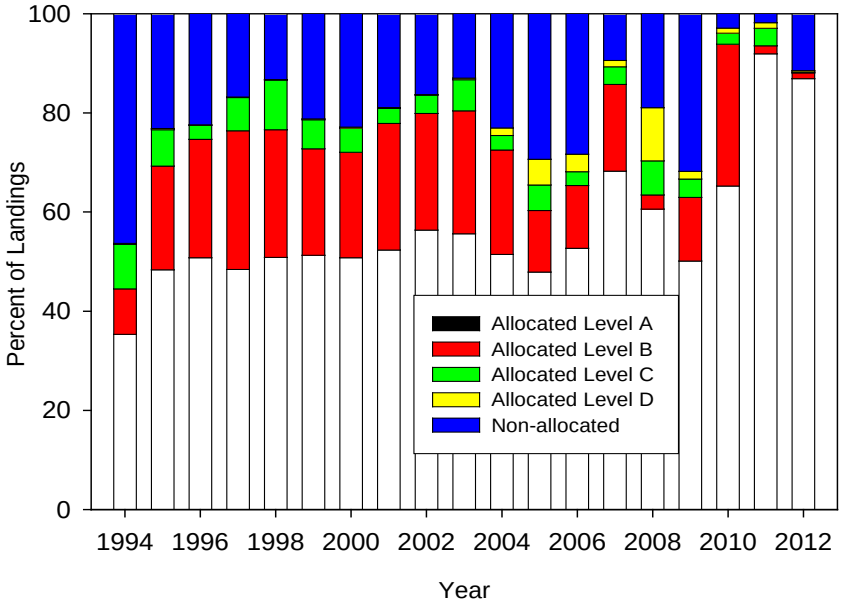
HADDOCK



RED HAKE



SILVER HAKE



Area Allocation Procedure Summary

Trip-based multi-tier allocation procedure is a comprehensive method

- all species are addressed, landings at statistical area, estimates effort
- an improvement over the single-species proration

Majority of allocated landings matched at the Vessel level (Level A & B)

VTR data are a representative subset of Dealer data

Random component of the allocation is not creating high variability in estimates of 1994 stock landings

Majority of the biological samples do not change stock area

Strengths and Challenges

Strengths:

- Trip-based allocation provides landings at a finer temporal and spatial resolution than the single species prorations
- Comprehensive approach to create a master dataset to support *all* analyses
- Uses same data sources as single species proration
- Multi-tier approach is similar to port agent's interview; VTR provides higher coverage rate

Challenges:

- VTRs are self-reported data; but no other data source is available.
- VTR compliance and data auditing could be improved.
- Trip-based approach to maintain link to biological samples.
- Examination of effort over entire time series is needed.

II. Measures of Uncertainty in Trip-based Allocated Landings

by
S Wigley, J Blaylock, and M Palmer

WP# 11 SARC 52 Winter Flounder
in support of TOR

Perform a sensitivity analysis which examines the impact of allocation of catch to stock areas on model performance.

Uncertainty in matching method

For Dealer AA, uncertainty at Alevel = B, C & D characterized using multinomial probability

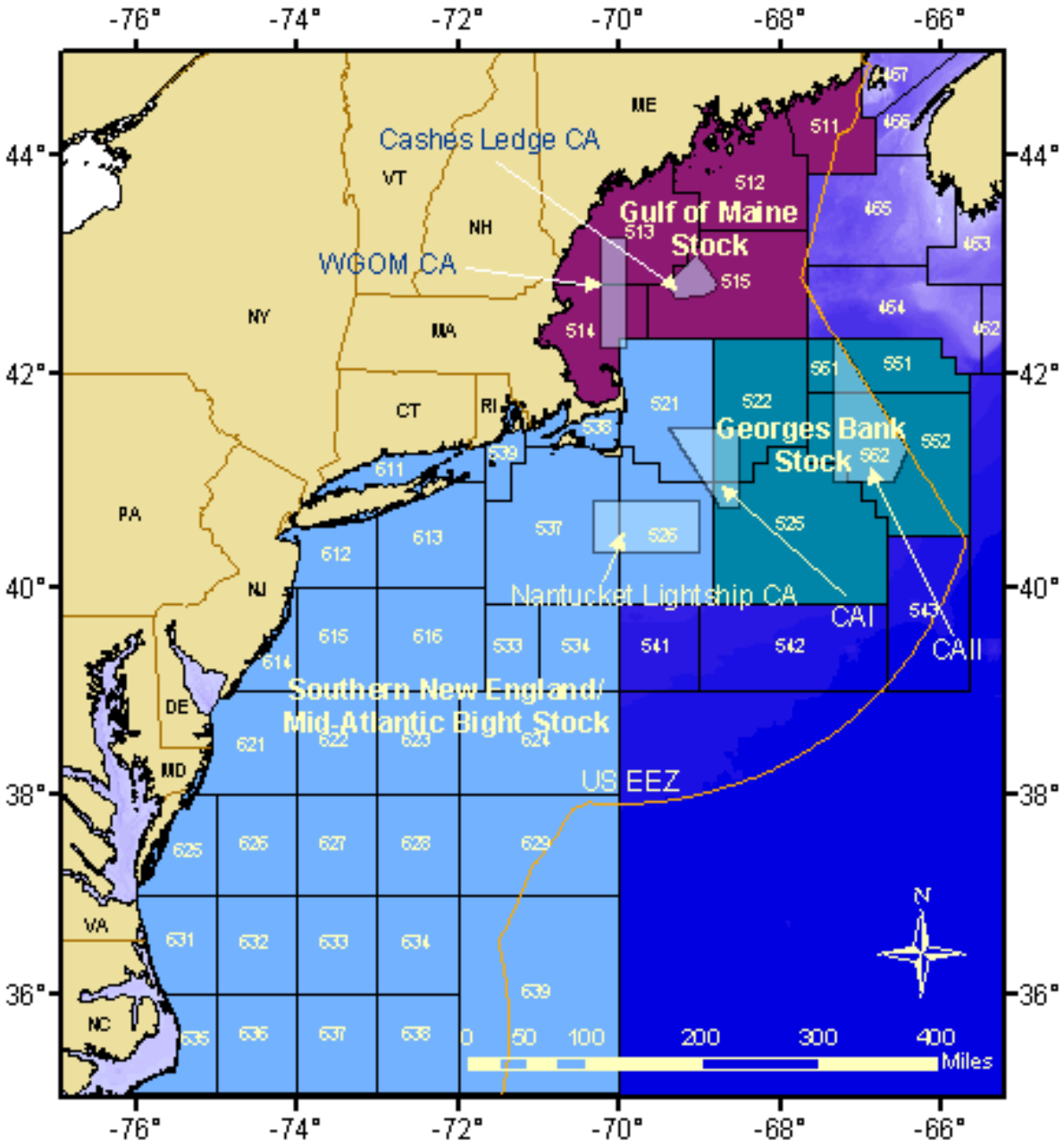
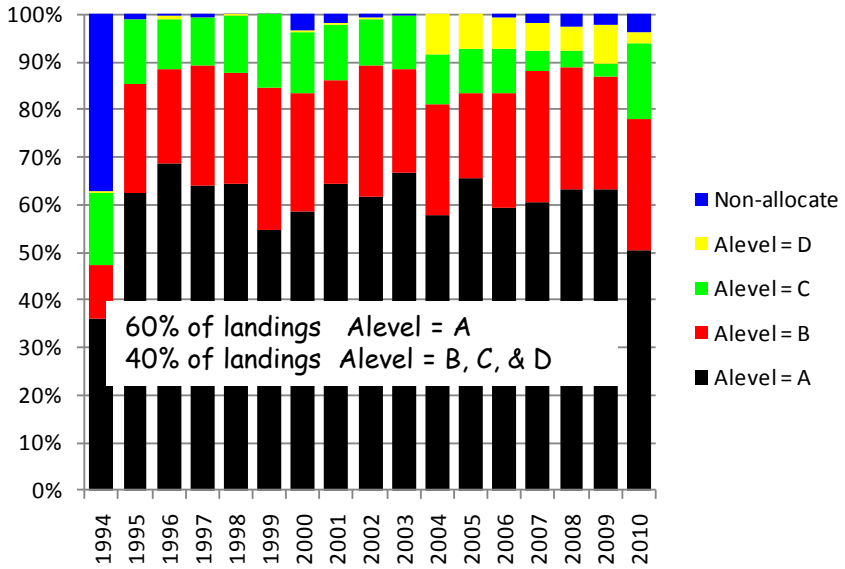


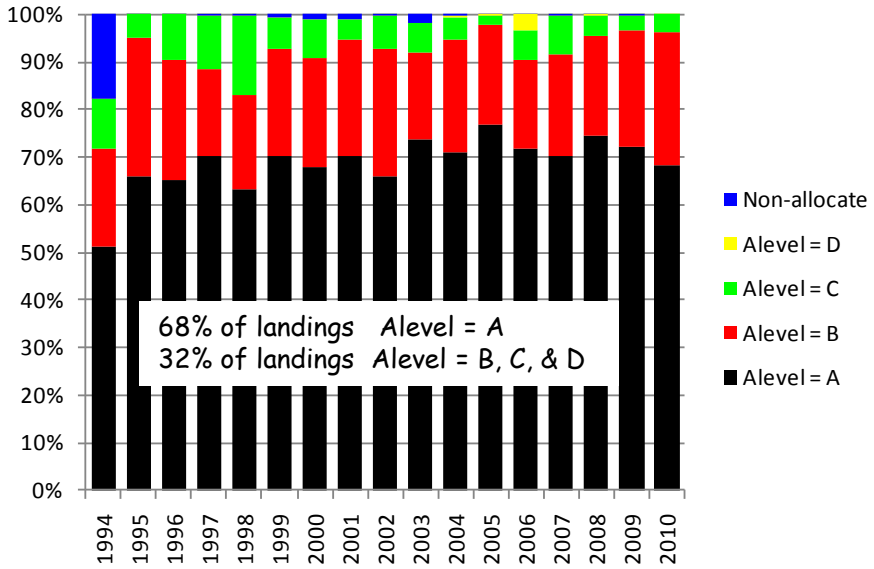
Figure 11.1. Statistical areas used to define the Gulf of Maine, Georges Bank, and Southern New England/Mid-Atlantic Bight winter flounder stocks.

Percentage of stock landings, by year and Alevel

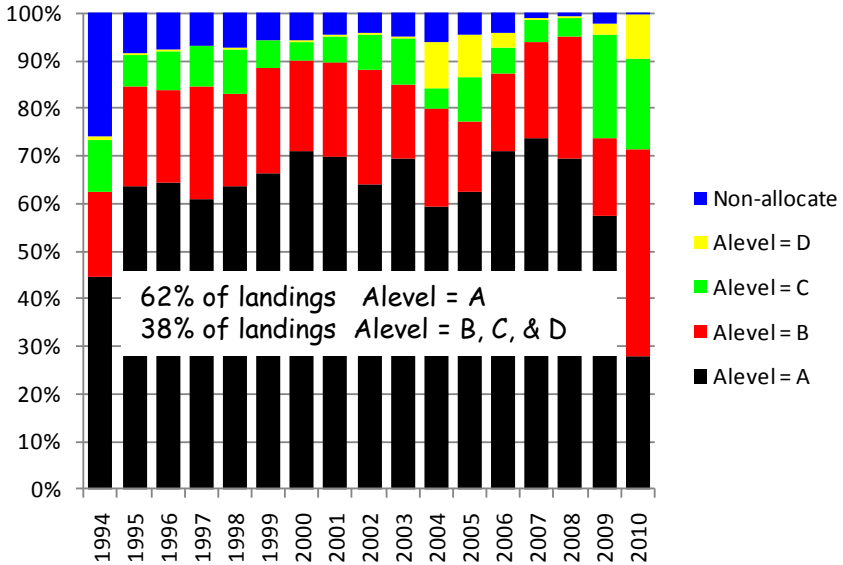
GOM Winter Flounder



GB Winter Flounder



SNE Winter Flounder



Variance associated with allocated landings at Alevel B, C & D

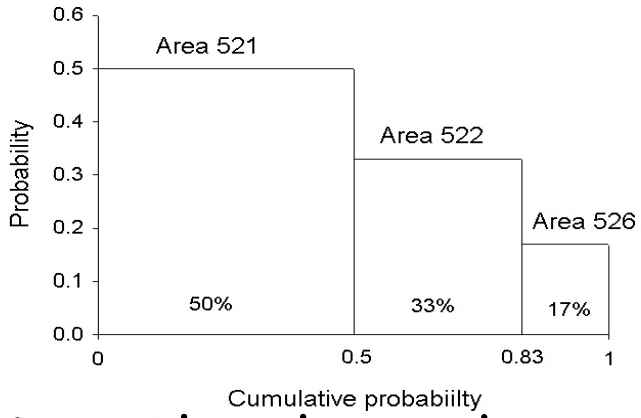
Variance of a multi-nominal probability distribution

$V(T) = pq = p * (1-p)$
where p = probability associated with the Dealer trip T

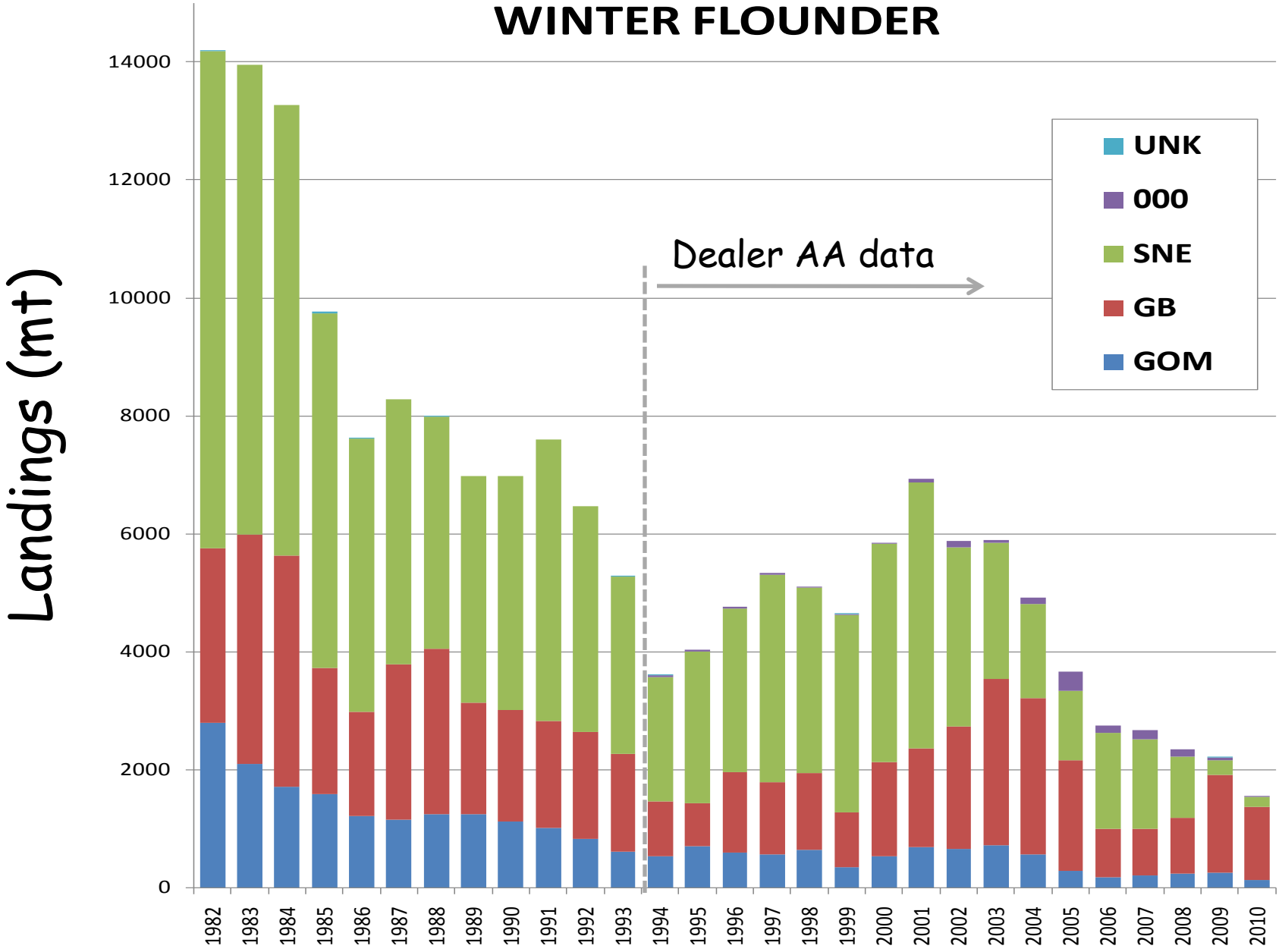
$CV(T) = sqrt(pq)$

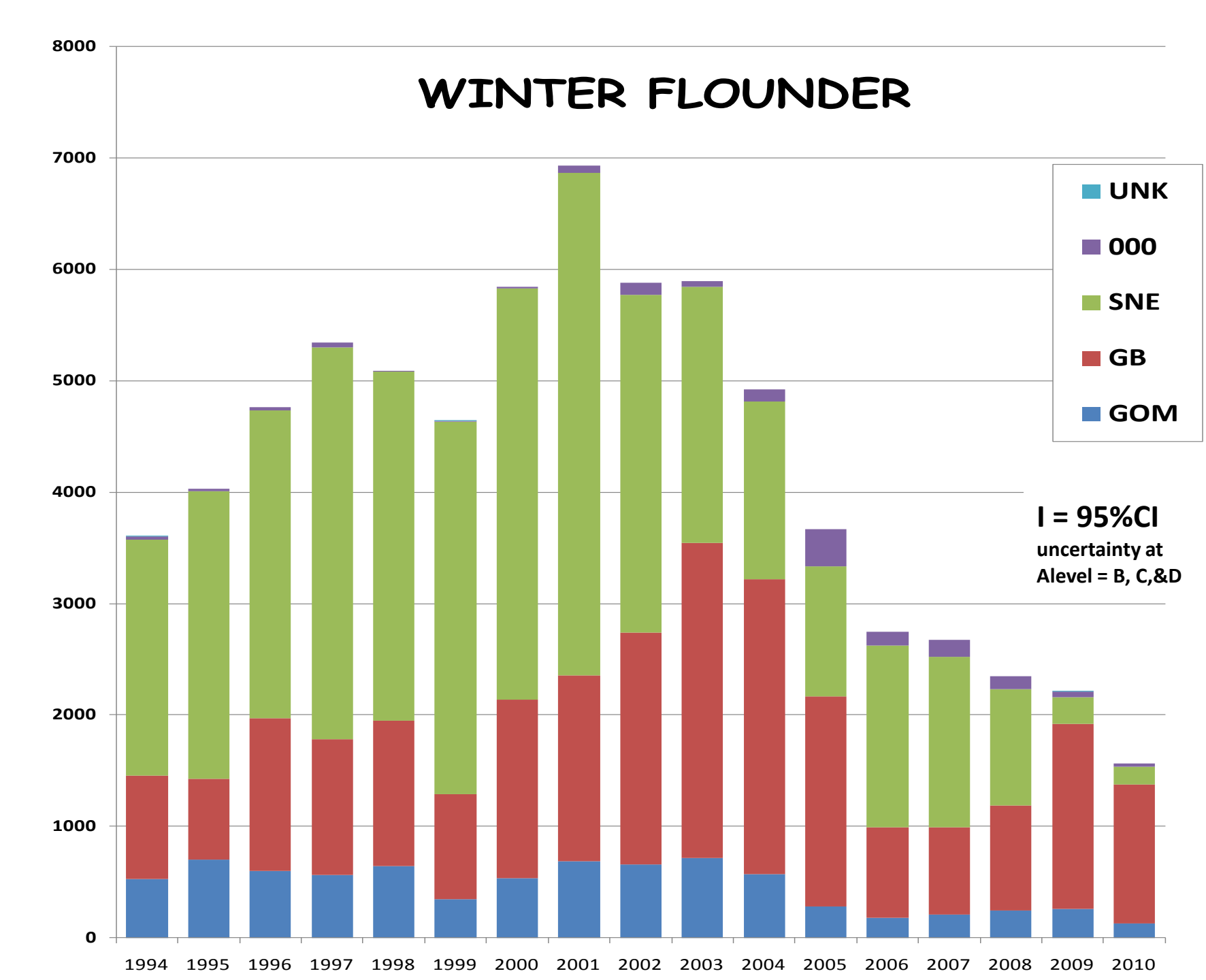
To translate between Dealer trip T and trip pounds (L) we assume the $CV(T) \sim CV(L)$

Then back out the variance
$V(L) = (CV(T) * L)^2$
$var\_mt = prob * (1-prob) * mt^2$



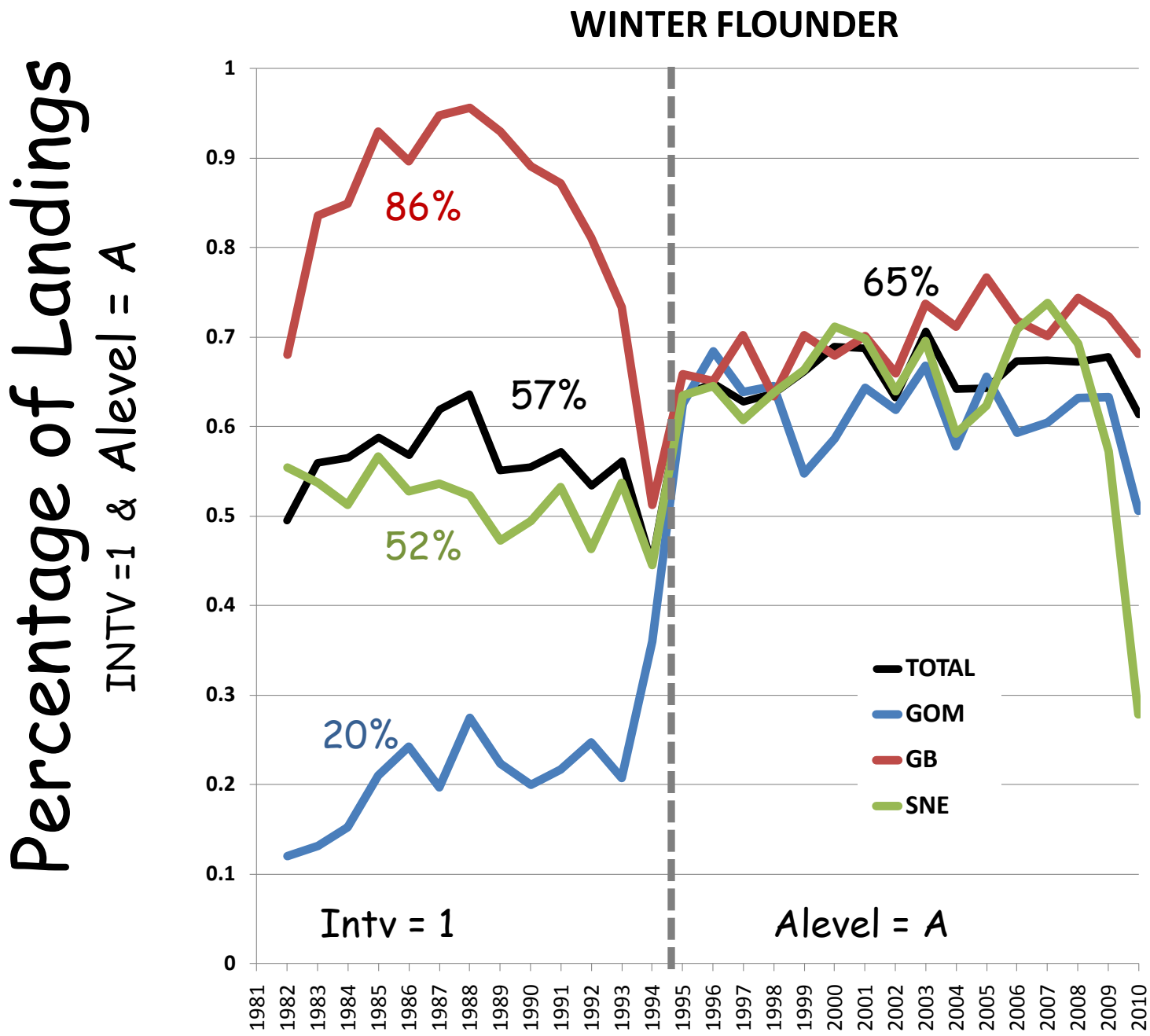
Calculate confidence intervals by species, Alevel, stock


WINTER FLOUNDER
Landings (mt)
14000
12000
10000
8000
6000
4000
2000
0
Dealer AA data
UNK
000
SNE
GB
GOM
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010


WINTER FLOUNDER
8000
7000
6000
5000
4000
3000
2000
1000
0
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
UNK
000
SNE
GB
GOM
I = 95%CI
uncertainty at
Alevel = B, C,&D

Winter flounder landings (mt), 95%CI(mt) and percentage of uncertain mt, by stock and year. NOTE—Generally less than 2%

Year	GOM	95 CI	%	GB	95 CI	%	SNE	95 CI	%	Area 000	UNID stock	
1982	2798.7			2958.6			8420.1				0.2	
1983	2099.1			3893.8			7963.7					
1984	1706.0			3926.6			7635.2					
1985	1583.4			2151.0			6005.4				23.2	Grand Banks
1986	1216.0			1761.3			4639.4				7.1	Grand Banks
1987	1159.9			2636.6			4482.7					
1988	1250.6			2803.9			3932.1				0.0	
1989	1252.9			1880.1			3846.9					
1990	1117.0			1898.0			3963.5					
1991	1008.3			1814.3			4782.8					
1992	824.6			1821.5			3815.5					
1993	611.5			1659.6			3010.4				2.3	
1994	528.5	4.6	0.9%	929.1	16.3	1.8%	2113.7	18.3	0.9%	30.5	1.2	stat area 460s
1995	699.9	11.3	1.6%	728.3	16.0	2.2%	2582.9	18.1	0.7%	18.1		
1996	602.2	11.5	1.9%	1366.3	24.0	1.8%	2767.7	29.3	1.1%	23.9		
1997	566.3	16.4	2.9%	1219.0	24.4	2.0%	3515.5	47.8	1.4%	42.6		
1998	640.7	7.8	1.2%	1308.0	32.1	2.5%	3134.8	42.1	1.3%	5.4		
1999	348.5	4.7	1.3%	937.5	21.5	2.3%	3342.8	32.5	1.0%	8.3	0.1	
2000	533.1	5.6	1.0%	1603.1	31.0	1.9%	3692.8	28.1	0.8%	13.7		
2001	691.0	11.3	1.6%	1667.4	32.6	2.0%	4509.0	32.4	0.7%	63.0		
2002	658.2	14.3	2.2%	2079.7	34.0	1.6%	3033.2	33.2	1.1%	106.4		
2003	716.0	4.9	0.7%	2828.2	38.9	1.4%	2301.8	25.8	1.1%	46.0		
2004	573.0	6.2	1.1%	2647.2	39.8	1.5%	1593.3	39.0	2.4%	106.0		
2005	282.5	4.4	1.5%	1882.0	24.0	1.3%	1168.0	26.8	2.3%	334.5		
2006	180.7	2.4	1.3%	814.1	13.0	1.6%	1632.0	14.5	0.9%	119.4		
2007	209.8	1.8	0.9%	785.9	15.0	1.9%	1525.5	17.4	1.1%	155.1		
2008	242.4	2.9	1.2%	944.5	14.7	1.6%	1043.0	12.9	1.2%	117.2		
2009	261.3	1.7	0.7%	1656.4	30.8	1.9%	242.1	10.9	4.5%	52.6	2.2	stat area 468
2010	129.4	1.6	1.3%	1249.6	32.4	2.6%	157.8	13.9	8.8%	28.5		
avearge			1.4%			1.9%			1.8%			


WINTER FLOUNDER
Percentage of Landings
INTV =1 & Alevel = A
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
86%
57%
52%
20%
65%
TOTAL
GOM
GB
SNE
Intv = 1
Alevel = A
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010

SARC 52 WP 11 Summary

- High percentage of landings match a the vessel level (Levels A and B)
  - Level A percentages are greater than interview percentages for SNE and GOM winter flounder stocks
- Uncertainty can be characterized for stock landings at Alevel = B, C, & D for 1994 onward

- No measure of uncertainty for landings prior to 1994

III. Investigative Analyses

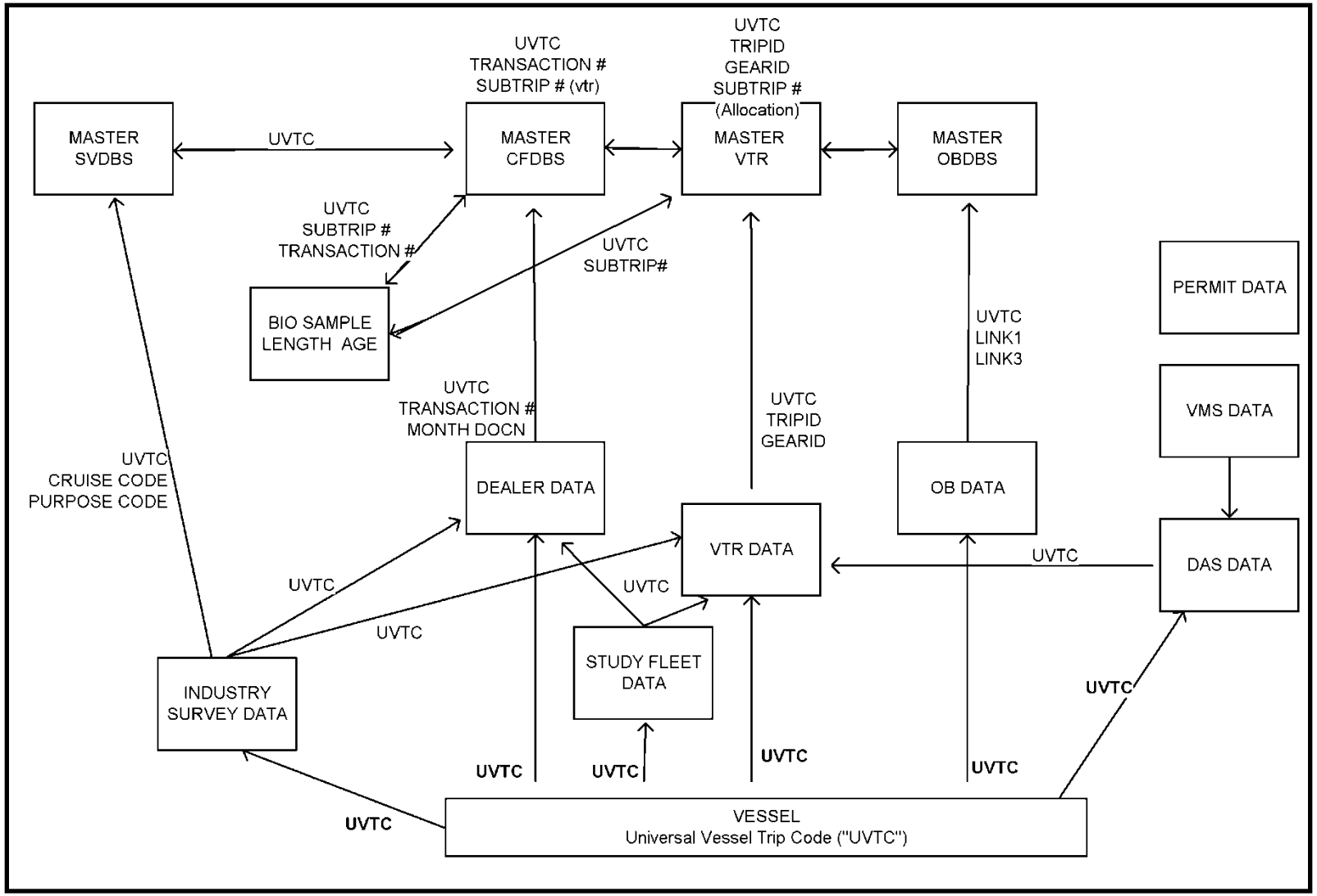
by Palmer and Wigley

A short list of examples (not exhaustive)

- only intended to highlight a few key issues

Example of Need: Trip Identification

The current trip identifier does not allow seamless linking/matching of fishing trips between systems



Unique Trip Identifier

Example of Need: Trip Identification (con't)

- The QA/QC on the current trip identifier (VTR serial number) is not sufficient to match trips across databases with the accuracy needed to support current and future scientific and management needs.
- To match trips across databases other methods must be used
  - Can use match on vessel permit number and landing/transaction date
    - Assumes a single transaction date per trip and the transaction occurs on the date of landing
    - This is the only method that can be used prior to 2004
  - A combination of methods can improve the match rate compared to use of only a single method…but higher matching rates are needed.

…

Example of Need: Trip Identification (con't)

Table 1. Trip matching rates between VTR and dealer trips using various methods, 2004 to 2012. Only trips where fishing occurred and fish other than lobster were retained were used in the analysis.

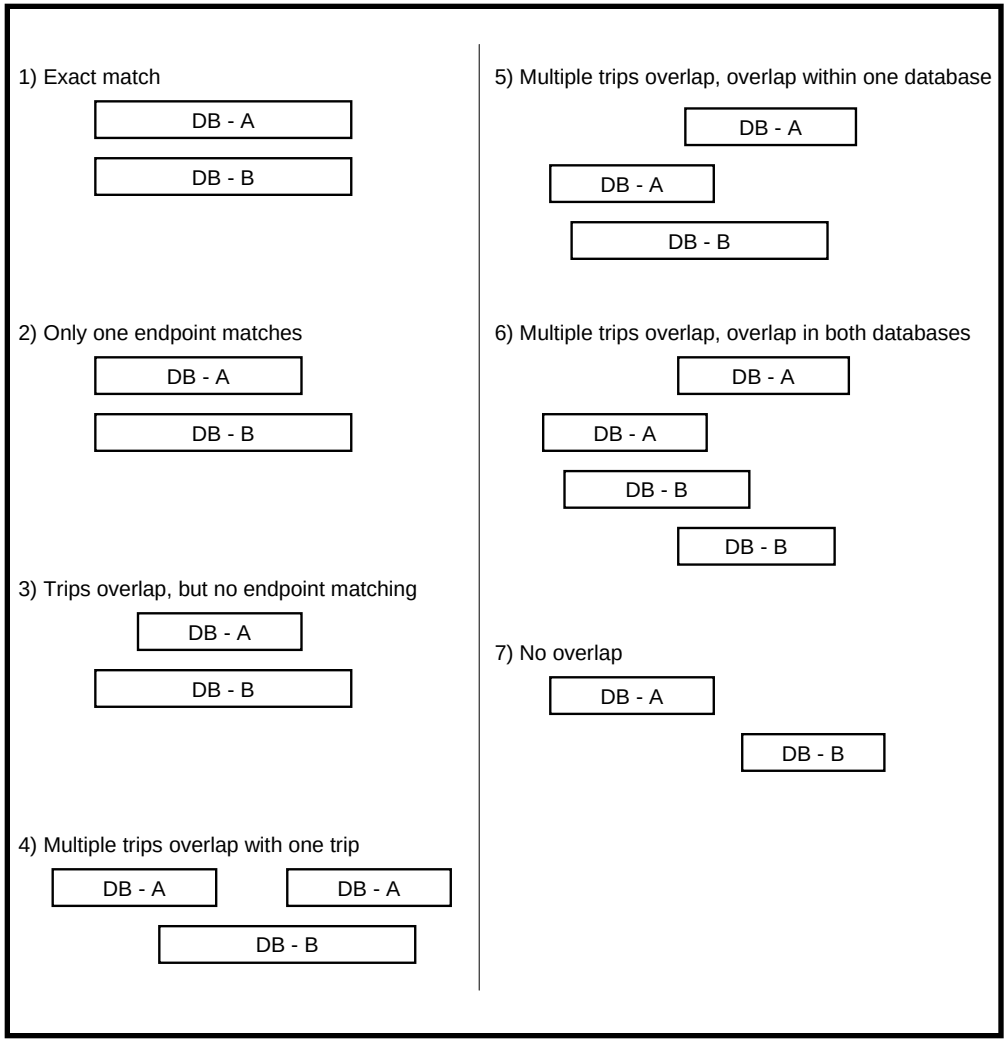
Year	Total commercial VTR trips	Match on only VTR serial number		Match on VTR serial number, vessel permit and month landed/sold		Match on vessel permit, month and day landed/sold		Combined methods		Annual improvement (%)
		Number of trips	Match ratio (%)	Number of trips	Match ratio (%)	Number of trips	Match ratio (%)	Number of trips	Match ratio (%)	
2004	86,287	54,527	63.2	51,672	59.9	50,713	58.8	59,435	68.9	
2005	93,213	65,890	70.7	62,203	66.7	58,671	62.9	71,318	76.5	7.6
2006	92,578	65,991	71.3	62,235	67.2	57,933	62.6	72,147	77.9	1.4
2007	86,570	63,703	73.6	60,030	69.3	56,301	65.0	69,128	79.9	1.9
2008	82,480	64,951	78.7	61,309	74.3	57,381	69.6	69,116	83.8	3.9
2009	80,711	64,413	79.8	60,607	75.1	55,868	69.2	67,336	83.4	-0.4
2010	75,286	61,832	82.1	58,087	77.2	52,230	69.4	62,765	83.4	-0.1
2011	73,464	63,663	86.7	61,096	83.2	53,486	72.8	63,972	87.1	3.7
2012	70,608	62,450	88.4	59,292	84.0	51,235	72.6	61,354	86.9	-0.2
2004 -2012 summary	82,355	63,047	77.2	59,615	73.0	54,869	67.0	66,286	80.9	2.3

Taken from Palmer 2007; updated through 2012

Example of Need: Trip Identification (con't)

Surrogate methods are inferior to a direct link using a unique trip code/identifier with sufficient quality controls.

Between database checking



Within database checking

Fig. 2. Agreement scenarios for matched trips between the Vessel trip report database (A) and other fisheries-dependent databases (B) (DB = database). Taken from Wigley et al. 2008 (NEFSC CRD 08-02).

Example of Need: VTR Compliance regarding Area Fished

- Commercial landings are assigned to stock areas using the statistical areas/positions reported on VTRs

  Trip-based allocation (Wigley et al. 2008)

- Misreporting of area fished on VTRs is a known problem (Palmer et al. 2007; Nies and Applegate 2007)

  Primarily, fishers under-report the number of statistical areas fished

- VTR compliance needs to be improved, and scope of the problem is manageable

  Of the approx. 2,500 vessels submitting VTRs annually, there are ~270 vessels which frequently under-report statistical areas on their VTRs

  These vessels have been, and can be, easily identified when VTRs are cross-validated with VMS data

Reporting Compliance: Underreporting of Stat Areas in VTR

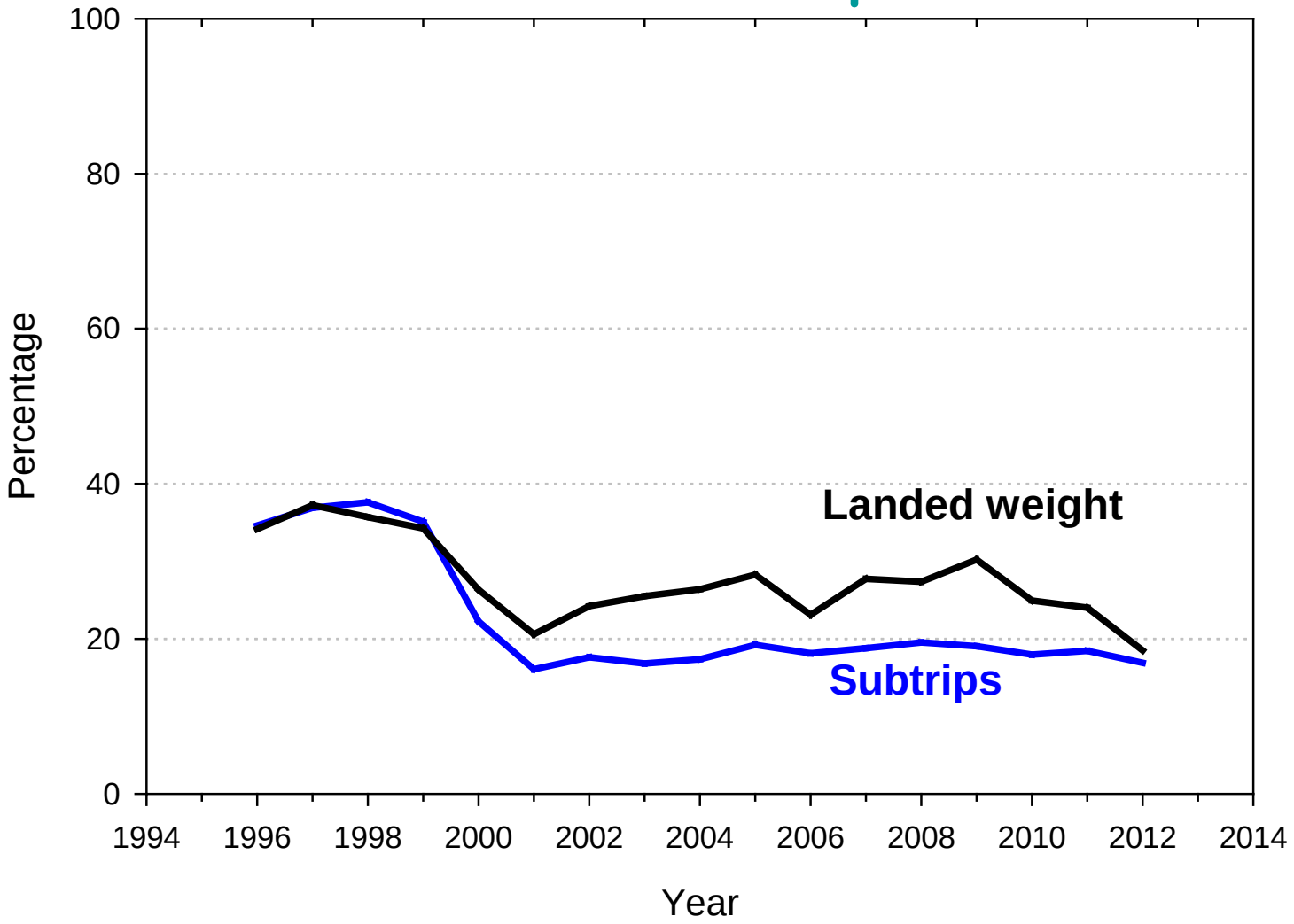
Year	Percentage of multi-area trips with correctly filed VTRs
2006	4.1
2007	4.7
2008	4.3
2009	5.0
2010	6.6
2011	8.6

Vessels with 5+ VTRs incorrectly reported.

Year	# Vessels	# Trips
2006	269	3849
2007	307	4485
2008	199	2747
2009	280	4334
2010	255	3803
2011	239	3961

Taken from Palmer CRD 07-22 and update

Vessel Trip Report Internal Inconsistency between Statistical Area and point location



2000-2012 mean : 18% VTR subtrips and 25% of VTR landed weight have a point location that is not within the reported statistical area.

IV. Proposed Solutions

Area Allocation procedure is a peer-reviewed method to link critical information, BUT… it is not the solution!

- Remove need for allocation procedure

- Robust trip identifier to link trips cross all databases

- All VTR reporting submission times must be the same

- Core data collection programs need fine-tuning to meet current and future management needs (fix/expand existing data collection systems)

- Strongly encourage electronic reporting for all systems

- Improve reporting compliance & Tighten QA/QC using within and between cross-checking; Utilizing data leveraging between VMS and VTR is best way to improve reporting compliance

- Expand data elements to identify fleets (SAPs, gear & gear modifications)

NOAA FISHERIES

Questions?