Terms of Reference for the River Herring Stock Assessment Peer Review

1. Evaluate the justification for inclusion or elimination of available data sources.

The River Herring Stock Assessment Subcommittee (SASC) cast a wide net to collect and synthesize data from as broad a variety of sources as possible. The approach was inclusive rather than exclusive, and uncertainties and caveats were noted.

For inland/coastal information, 57 systems (mostly rivers) were included in the coast-wide assessment (Table 1). Nine categories of fisheries independent and dependent information were considered by the SASC. Most of the valid information was for northern systems; much information was lacking, particularly in southern states. It was noted that few state surveys actually target river herring per se. Some of the better count data were at fish passage facilities. For select data sets, a change in sampling methodology was a concern, as it limited utility of a data set for temporal trend analysis. Overall, however, there were sufficient data to undertake many of the analyses presented by the SASC.

Historical and modern catch data were obtained state by state and for the entire U.S. coast. NOAA Fisheries maintains data from 1950 onward, while pre-1950 data were from a combination of federal and state sources. Although the first reported catches dated from 1887, both the SASC and the Peer Review Panel noted that large data gaps occur prior to 1950 due to incomplete reporting by state. As an example, the U.S. Fisheries Commission reported river herring harvests in 1892 as coming solely from Massachusetts (3,651,000 lbs or 1,659.5 MT). On the other hand, the New York Times, which reported a great deal on fisheries in the 19th century, listed additional 1892 harvests of river herring from New York, Delaware, Maryland, and North Carolina totaling 19,932 MT – thus, the total harvest for that year was well over 20,000 MT or a factor of 12 larger than reported in U.S. statistics (NYT 1895). The Panel recognizes the difficulties in estimating catch from historical sources, but encourages the SASC to pursue these avenues in the future.

A problem with catch data is that these are generally reported only as 'river herring' or even as 'alewife'. Parsing out the species can be done by making reasonable assumptions about range distributions (cf. Limburg and Waldman 2009). However this was not done for the assessment.

Recreational catch data were not used because the only data source, NOAA's Marine Recreational Information Program, does not collect data in fresh water where most recreational fishing for river herring occurs. Additionally, there was concern about species misidentification in this dataset.

Trend analyses were conducted on most datasets, including catch-per-unit-effort data (loess smoothed, 11 rivers), run size estimates (23 rivers), young-of-year indices (13 rivers as well as lower Chesapeake Bay and Albemarle Sound), miscellaneous young-of-year, juvenile, and adult surveys (4 rivers), 19 trawl surveys, as well as the biological (mean length, maximum age) and population level (total mortality, Z, computed by age or by repeat spawning marks) information. The Panel noted that while the catch rate series were standardized for effort, analyses of these data would have benefited from use of Generalized Linear Modeling approaches which would have allowed more in-depth

exploration of the trends in the data as well as their uncertainties. Further comments on the uncertainties in the trend analyses are evaluated in ToRs 3 and 4.

Indices of run sizes based upon visual or electronic counters were available for six states for differing time periods preceding the 2010 surveys. Cluster analyses of three time intervals were conducted (1984-2010, 1999-2010, and 2003-2010) to explore temporal and spatial trends in run size. The first time period allowed for the longest time series to be analyzed but was restricted to 10 rivers (3 Maine, 4 New Hampshire, and 1 each in Massachusetts, Rhode Island, and Connecticut). A reduction in the time period (1999-2010) allowed more recent trends to be examined, increasing the analysis to 15 rivers (3 Maine, 6 New Hampshire, 3 Massachusetts, 2 Rhode Island and 1 Connecticut). The final time series (2003-2010) allowed the inclusion of 19 rivers (4 Maine, 6 New Hampshire, 3 Massachusetts, 3 Rhode Island and 3 Connecticut).

Although the run sizes in most rivers examined exhibited a decline, no geographic relationships could be detected by the cluster analysis. The data from 2003-2010 did show some promise as a geographic predictor of a latitudinal relationship and additional (future) analysis will be needed to bear this out. A problem with analysis of run counts is that the data are subject to both natural (i.e. spring rainfall) and anthropogenic modifications (i.e. river diversion or fishway modification) in upstream accessibility that can be acute or long term. Other confounding factors include the location of any obstruction or fishery component downstream of the census location and the absence of data on whether or not river herring use specific spawning locations within a river.

Length data were available from eight states (Maine, New Hampshire, Massachusetts, New York, Maryland, North Carolina, South Carolina, and Florida) along with the National Marine Fisheries Service Bottom Trawl survey. Sex-specific trends in length over time were examined for Maine through South Carolina; however large gaps in the Florida time series (1973-2001) prevented its inclusion. Although sampling methods were inconsistent between rivers, all trend analyses were based on within system sampling so gear selectivity should not have been a concern.

The trend analysis of the length data found a negative relationship in 4 of 10 rivers for alewife and 5 of 8 for blueback herring. The SASC noted significant trends were more common in times series that began in 1990 or earlier, and hence the length of the time series may be a confounding factor. The potential for a geographic bias may also be present for the two species because the number of rivers sampled was not even between regions. Of the six rivers where significant trends were found, only two were from New England while 8 of the 12 rivers examined were from this region. Evidence for this concern may also be seen in the results of the NEFSC Bottom Trawl survey where coast-wide trends were seen in alewife and "to a lesser extent in the blueback." It should be noted that Marcy (1969; cited as an ageing reference by SASC) notes a latitudinal trend in size that was apparent in the late 1960's. The panel realizes the SASC does not have the power to control data collection but encourages all attempts to obtain data from the under or non-represented states (regions).

River herring age data, determined by scales, were used for maximum age, length-at-age analyses, age at maturity, and associated mortality estimates. Potential problems with growth differences precluded use of length keys to develop age estimates.

All states cited the methods of Cating's (1953) study of ageing shad scales as part of the methodology for ageing their river herring. Several problems with the use of Cating's method have been discussed in recent years (McBride et al. 2005 and Duffy et al. 2011). Most recently Duffy et al. (2011) found that Cating's method does not reliably account for shad ages over large latitudinal ranges. Some of the discrepancy lies in the use of transverse grooves to establish the freshwater zone and ages one to three. They concluded transverse groove formation is more closely related to scale size (fish size) rather than a function of age. This would create a latitudinal interpretation problem that becomes more acute as the trends in decreasing length noted above develop.

The SASC clearly noted the weaknesses of using ages determined by scales:

> "These protocols have not been validated with known-age fish, and there have not been many efforts to standardize river herring ageing across states. As with any ageing method, there is the potential for bias both between labs and within labs over time as personnel change and methods are not consistently standardized."

Additionally, the Panel recognized that in the absence of validation (using known age fish) or alternate aging structures (i.e., otoliths) there were no alternatives. The Panel felt strongly that there is a need to develop a standardized, validated ageing process to reliably provide vital life history data.

Overall, the Panel concluded the SASC adequately justified the inclusion and exclusion of the available data in its analysis.

State	River	Time series	By species	Harvest	Age
ME	Damariscotta	1943-2010		●	
	St. George	1943-2010		●	
	Union	1975-2010		●	
	Orland	1943-2010		●	
	Androscoggin	1983-2010	●		●
	Sebasticook	2000-2010	●		●
	Merrymeeting Bay/Tribs	1979-2009	●		
	Gulf of Maine	2000-2010	●		
NH	Exeter/Squamscott	1991-2010	●	●	●
	Lamprey	1991-2010	●	●	●
	Winnicut	1991-2010	●	●	●
	Oyster	1991-2010	●	●	●
	Cocheco	1991-2010	●	●	●
	Taylor	1991-2010	●	●	●
	Great Bay Estuary	1997-2010	x		
MA	Mattapoisett	1988-2010	●	●	o
	Monument	1980-2010	●	●	o
	Nemasket	1996-2010	●	●	o
	Parker	1971-1978, 2000-2010	●	●	o
	Town	2000-2010		●	
	Agawam	2006-2010		●	o
	Back	2007-2010	●	●	●
	Charles	2008-2009		●	●
	Mystic	2004-2010	●		●
	Quashnet	2004	●		●
	Stony Brook	1978-2004	●		o
RI	Gilbert Stuart	1981-2010			●
	Nonquit	1999-2010			●
	Buckeye Brook	2003-2010			
	Pawcatuck	1988-2010			x
	Ocean waters	1979-2010			
	Naragansett Bay	1988-2010			
	Coastal ponds	1992-2010			
CT	Bride Brook	1966-1967, 2003-2011	●		
	Connecticut River	1975-2011	●		
	Farmington River	1976-2011	●		
	Thames River	1996-2011	●		
NY	Hudson	1975-2010	●	o	o
DE, NJ, PA	Delaware River	1980-2010	o	o	o
	Delaware Bay	1966-2010	o	o	o
MD	Nanticoke	1959-2010	o		o
	Susquehanna	1972-2010	o		
	Chesapeake Bay	1959-2010			o
MD, VA, DC	Potomac River	1959-2010		●	
VA	James	1966-2010	o	●	o
	Rappahannock	1966-2010	o	●	o
	York	1966-2010	o	●	o
NC	Albemarle Sound	1972-2009		o	
	Chowan River	1972-2009	●	●	●
SC	Wynah Bay				
	Santee-Cooper	1969-2010	o	●	o
	Savannah River				
	Ashley-Combahee-Edisto Basin				
GA	Altamaha River	2010			
	Ogeechee River	2010			
	Savannah River	2010			
FL	St. John's River	2001 -2010	●		

Length	Weight	Repeat Spawner	FI Adult	FI JAI	FD CPUE
●					
●					
●				●	
●			●		
●		o	●		
●		o	●		
●		o	●		
●		o	●		
●		o	●		
●		o	●		
x				x	
o	o		●		
o		o	●		
o	o				
o			●		
			●		
o					
●			●		
●	●	●	●		
●	●	●			
●	●	●			
o	o	o	o		
●	●	●	●	o	
●	●	●	●	o	
			●		
x	x	x	o	●	
●			●	●	
●			●	●	
●			●	●	
o			●		
o			●	o	
			●		
			●		
o	o	o	o	o	o
o			o	o	o
o			o	o	o
o		o		o	o
			x		
					o
o			o	o	o
o	o	o	o	o	o
o	o	o	o	o	o
o	o	o	o	o	o
		o	o	●	
●	●	o			●
				x	
o	o	o	o	x	●
				x	
				x	
				x	
				x	
				x	
●			●	o	

Table 1. Summary of available data and data quality by state, river, or other system (from SASC). Dark grey cells with filled circles indicate data sets available for the entire time series of interest; medium grey cells with open circles had partial data sets available; and light grey cells with "x" indicate data sets not reliable enough to use for the assessment. Blank cells indicate no data.

2. Evaluate the estimates of ocean bycatch of river herring and the methods used to develop estimates.

For many years, incidental bycatch in marine fisheries was a known but unquantified mortality source for river herring and shad, and was identified as a high priority in the most recent American shad stock assessment review (ASMFC 2007). For the current river herring assessment, incidental catch - defined as alosines brought aboard and either retained (landed) or discarded at sea - was quantified for the first time. The purpose was to compare the magnitudes of incidental catch from all sources to reported commercial catches.

Data were obtained from the Northeast Fishery Observer Program (NEFOP) and were quantified by fleet for 14 different gear types (see pg. 19 of the stock assessment report), by year, season, geographic area, gear group, and mesh size for each species. Bycatch was estimated by taking the ratio of bycatch weight to caught weight as reported on ships by a NEFOP observer, and then adjusting these by the weight of the sold catch as reported by dealers, which is considered a more accurate weight.

Bycatch was assessed from 1989–2010. However, methodologies changed in 2005 for subsampling bycatch in high-volume midwater trawls and became better estimations. Hence, midwater trawl incidental catches are only included for 2005–2010. Coefficients of variation (CVs) were calculated following Wigley et al. (2007).

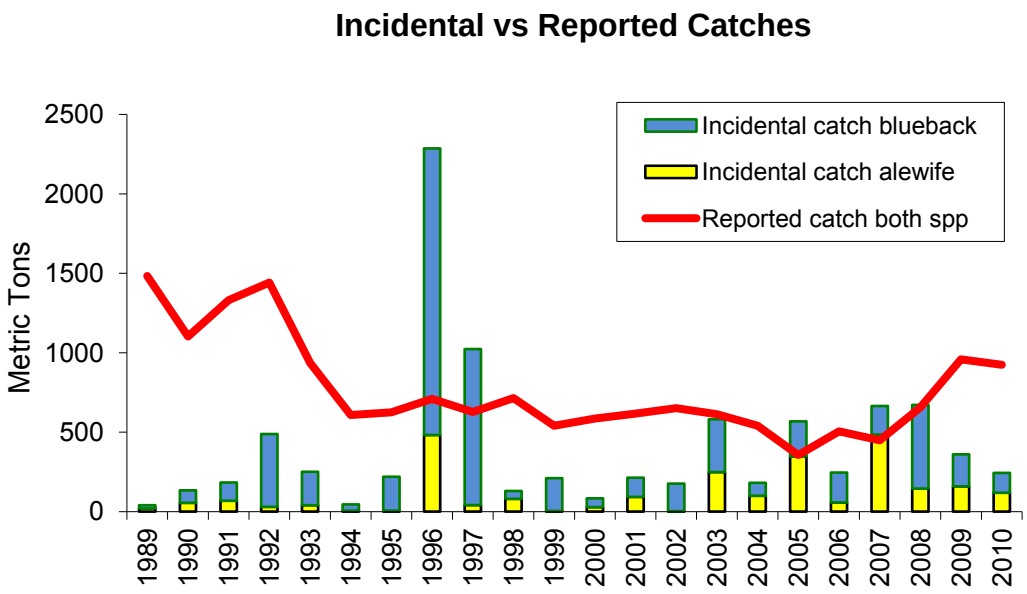

Figure. 1. Incidental catches of blueback herring and alewife, all gears and fleets reported by NEFOP observers, compared to total reported catches, 1989-2010. CVs not shown. Midwater trawl bycatch only included from 2005 onward.

Alewife bycatch ranged from a low of 2.72 MT in 2002 to 482 MT in 1996, with CVs ranging from 0.2–3.86 (20%–386%). Blueback herring bycatch ranged from 19.6 MT in 1989 to a high also in 1996 of 1803.4 MT, with CVs ranging from 0.2 to 2.1. Incidental marine catch estimates came close to or exceeded total reported commercial catches in 6

out of 22 years (Figure 1). Incidental catches occurred in all seasons, but tended to be highest during October – March. Midwater trawl catches were about equally proportioned between New England and Mid-Atlantic statistical areas, although New England small-mesh trawls took more incidental catch than Mid-Atlantic ones. Overall, New England incidental catches formed the larger part of the total (56%).

An unknown fraction of incidental catch is reported as 'landed catch' and thus the actual incidental bycatch reported as alewife and blueback herring is likely a bit lower than shown in Figure 1. However, an additional category of bycatch, called 'Herring – Unknown' (2.1 – 328 MT during this period) likely also includes river herring.

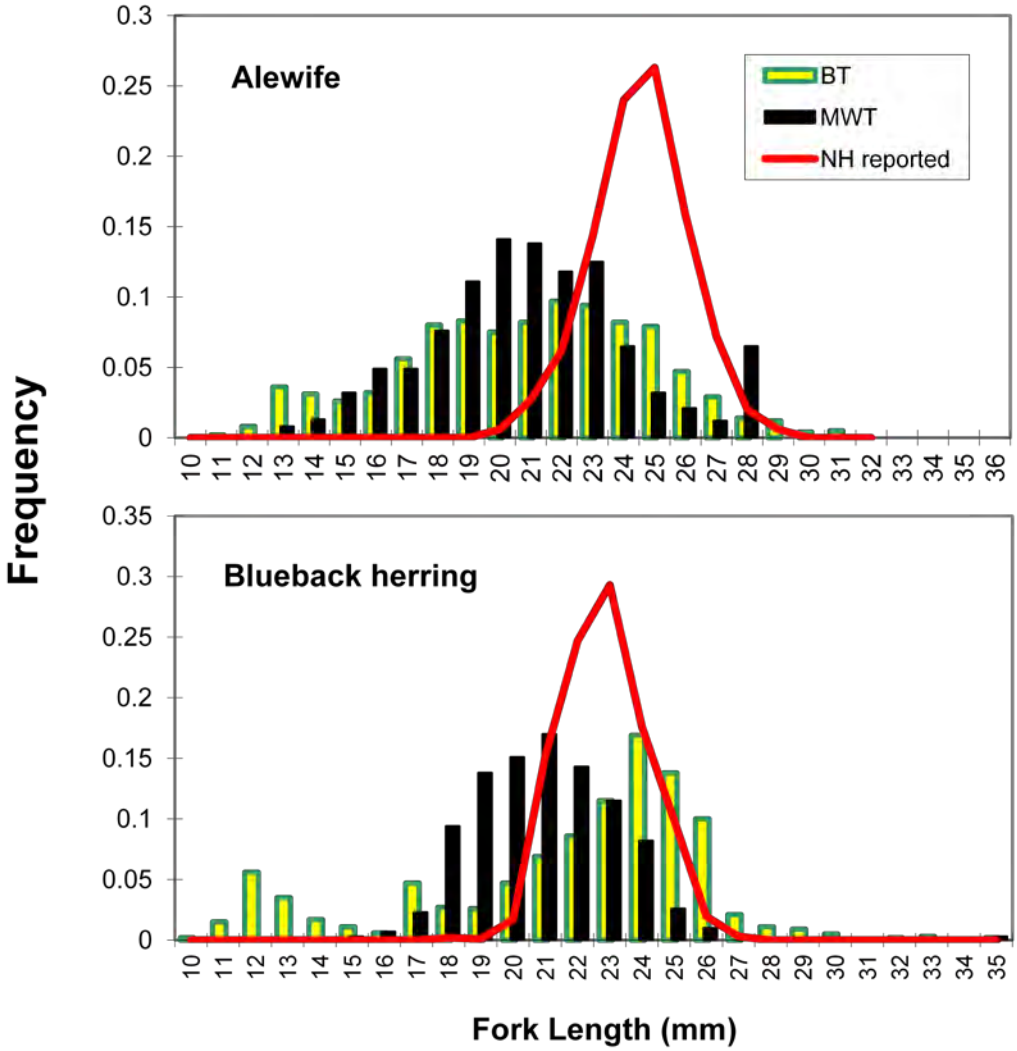

Figure 2. Length frequency distributions of alewife and blueback herring captured in bottom trawls (BT), midwater trawls (MWT), and compared to the spawner length frequency in New Hampshire. Data are from 2005-2010 added together.

Observers also record the sizes of incidentally caught river herring. It is noteworthy, even if expected, that a far broader range of sizes of both species were caught at sea than is the case in inland fisheries (Figure 2, using New Hampshire inland catches as a typical example of spawner size frequencies). For both species, large proportions of immature individuals were captured at sea. This is cause for concern.

Overall, the Panel considered the approach used by the SASC to assess incidental catches of river herring as reasonable and followed established protocols. Uncertainties were acknowledged. The Peer Review Panel encourages the assessment team to work to reduce uncertainties going forward, noting that CVs were lower in later years of the data presented. This likely is due to improvements in midwater trawl subsampling, among other things.

3. Evaluate the methods and models used to estimate population parameters (e.g., Z, biomass, relative abundance) and biological reference points, including but not limited to:

- **a. Evaluate the choice and justification of the preferred model(s) or method(s) of calculation. Was the most appropriate model or method chosen given available data and life history of the species?**
- **b. If multiple models were considered, evaluate the analysts' explanation of any differences in results.**
- **c. If appropriate, evaluate model parameterization and specification (e.g. choice of CVs, effective sample sizes, likelihood weighting schemes, calculation/specification of M, stock-recruitment relationship, choice of time-varying parameters, plus group treatment).**
- **d. Evaluate the diagnostic analyses performed, including but not limited to:**
  1. **Sensitivity analyses to determine stability of estimates and potential consequences of major model assumptions**
  2. **Retrospective analysis**

Besides examining trends in fishery-dependent and -independent indices of abundance, the SASC pursued three main categories of analyses to estimate population parameters. The first consisted of the estimation of river-specific total mortality (Z). Associated with this was derivation of Z reference points based upon a Spawner per Recruit (SPR) analysis. The second category consisted of the estimation of both river-specific and coast-wide exploitation rates (u). The third category consisted of two sets of population models, one set for specific rivers (Monument, Chowan and Nanticoke) and a second set for the coast-wide stock.

Total Mortality (Z)

Age frequency information was available for many of the coast's rivers from a variety of fishery-dependent and -independent sources (see ToR 1). The Chapman-Robson (1960) survival estimator, which is comparable to catch curve analysis but less biased, was applied to the annual age frequency data to provide a total mortality estimate by river,