Abundance Indices Workshop:
Developing protocols for submission of abundance indices to the SEDAR process

SEDAR Procedures Workshop 1

Oct. 14-17, 2008

Miami, Florida

Version I October 2009

Executive Summary

Given the complexity of developing indices of abundance and determining which indices may be useful for use in stock assessments, and that SEDAR independent peer review panels have commented that objective and consistent criteria for evaluating indices are currently lacking, a 4 day workshop was convened to gather key members of the SEDAR assessment community to prepare a guide to index development and evaluation. Over the course of developing stock assessments through the SEDAR process it was recognized that certain topics arose repeatedly in discussions. Such topics typically involve well known areas of uncertainty in assessment methods and input data sources, and discussions related to these topics often occupy considerable time in data and assessment workshops as participants strive to come up with reasonable solutions. One area where this became apparent was with regards to indices of relative abundance. SEDAR independent peer review panels have commented that objective and consistent criteria for evaluating indices are currently lacking. The criteria used to select particular indices for inclusion in a particular assessment have not been clearly documented, and those indices selected, along with the weighting scheme applied to them, for inclusion by the Data Workshop have been questioned. The intent of this workshop was to fulfill the recommendations of several SEDAR independent peer review panels by developing and documenting selection criteria, and improve future SEDARs through increasing efficiency by providing code for developing indices and diagnostics as well as guidance regarding what information is expected to be available for review, and the quality of assessment products.

TABLE OF CONTENTS

1. INTRODUCTION

1.1. Workshop Time and Place

The SEDAR Procedural Workshop I entitled "Abundance Indices Workshop: Developing protocols for submission of abundance indices to the SEDAR process" was held October 14-17, 2008, in Miami, Florida.

1.2. Statement of Work

Overview:

Measures of population abundance are critical to developing adequate and reliable quantitative stock assessments. Given the complexity of developing indices of abundance and determining which indices may be useful for use in stock assessments, and that SEDAR independent peer review panels have commented that objective and consistent criteria for evaluating indices are currently lacking, a 4 day workshop was convened to gather key members of the SEDAR assessment community to prepare a guide to index development and evaluation. Over the course of developing stock assessments through the SEDAR process it was recognized that certain topics arose repeatedly in discussions. Such topics typically involve well known areas of uncertainty in assessment methods and input data sources, and discussions related to these topics often occupy considerable time in data and assessment workshops as participants strive to come up with reasonable solutions. Because each workshop is composed of a slightly different mix of individuals, differing solutions are at times proposed for these common uncertainties. While it is within the charge to each individual workshop to make the best decisions possible based on the information before them, the lack of a consistent approach for addressing common problems across stocks and assessments can contribute to unnecessary criticisms of the assessments and the assessment process itself.

One area where this became apparent was with regards to indices of relative abundance. SEDAR independent peer review panels have commented that objective and consistent criteria for evaluating indices are currently lacking. The criteria used to select particular indices for inclusion in a particular assessment have not been clearly documented, and those indices selected, along with the weighting scheme applied to them, for inclusion by the Data Workshop have been questioned. The Panels have questioned why indices of abundance displaying conflicting trends have been included in the base assessment and it has appeared to some Reviewers that little or no evaluation of the indices with respect to their relationship and/or ability to track overall abundance has been considered in the selection process. Lack of clear documentation explaining why such choices had been made has left it to the lead analyst to try and explain the rationale, despite the fact that the analyst in question may not have participated in the Indices Working Group.

The intent of this workshop was to fulfill the recommendations of several SEDAR independent peer review panels by developing and documenting selection criteria, and improve future SEDARs through increasing efficiency by providing code for developing indices and diagnostics as well as guidance regarding what information is expected to be available for review, and the quality of assessment products.

Scope of Work:

Participants produced summaries or documents of various issues, which were discussed during the workshop.

Objectives of the workshop:

- Discuss and document factors that need to be addressed when developing indices of abundance, relative to their information source:
  - Fishery independent factors may include: survey objectives, sampling design, species targeted, life stage of target species at time of sampling, time of year of survey, areas where surveys are conducted, length of time series
  - Fishery dependent factors may include: species targeted (including data reduction techniques), life stage of target species at time of fishery prosecution, effects of data reporting errors, time of year of fishery prosecution, areas where fishery is prosecuted, length of time series, regulation and closure effects
- Determine criteria for selecting indices for inclusion in the assessment model given that they meet the minimum requirements set forth by the group
- Provide minimum requirements and an expected format for submission of SEDAR data workshop documents concerning development of abundance indices
- Provide appendices containing generalized computer code for developing abundance indices.

1.3. List of Participants

Workshop Panel

Walter Ingram... NMFS, SEFSC Pascagoula, MS
Julie Neer... SEDAR Coordinator, SEDAR, Charleston, SC

Craig Brown.. NMFS, SEFSC Miami, FL
John Carlson..NMFS, SEFSC, Panama City, FL
Shannon Cass-Calay.. NMFS, SEFSC Miami, FL
Paul Conn... NMFS, SEFSC Beaufort, SC
Nancie Cummings.. NMFS, SEFSC Miami, FL
Todd Gedamke.. NMFS, SEFSC Miami, FL
Brian Linton.. NMFS, SEFSC Miami, FL
Camilla McCandless...NMFS, NEFSC, Narragansett, RI
Kevin McCarthy.. NMFS, SEFSC Miami, FL
Joseph Munyandorero... Florida FWCC, St. Petersburg, FL
Joe O'Hop.. Florida FWCC, St. Petersburg, FL
Mauricio Ortiz.. NMFS, SEFSC Miami, FL
Clay Porch... NMFS, SEFSC Miami, FL
Marcel Reichert..SCDNR, Charleston, SC
Andi Stephens..SAFMC, Charleston, SC
Steve Turner.. NMFS, SEFSC Miami, FL
John Walter.. NMFS, SEFSC Miami, FL

2. Fishery Independent Indices

2.1. A Brief Overview of Large-Scale, Fishery-Independent Surveys of NOAA Fisheries, MS Labs – Walter Ingram

The following is a brief overview of the large-scale, fishery-independent surveys of NOAA Fisheries, Mississippi Laboratories. The surveys mentioned below generally contain useful information about time-series resource relative abundance levels, although of different species and ages/stages and over different lengths of time, depending on survey type. The following is a description of the survey-by-survey type of time-series information available for use in the SEDAR process.

Plankton Surveys
The plankton surveys in their present form started data collection in 1982. There are five general time frames where data were collected in the present format. These include spring plankton, summer ground fish, fall plankton, fall ground fish, and winter plankton. Plankton surveys are 'piggybacked' on the summer and fall groundfish survey designs and are hence categorized as 'piggyback plankton surveys.' The spring plankton and fall ground fish cruises have occurred each year since 1982. Some early plankton collection occurred in 1950-1976, but mainly for bluefin tuna. There were data in the spring survey time from 1977 through 1981, but the sampling design was different than the current version. Fall plankton started in 1986 and has occurred most every year since that time. The winter surveys have been conducted sporadically since 1992. Bongo and neuston nets are used during each of the surveys and use of MOCNESS (Multiple Opening/ Closing Net and Environmental Sampling System) and CUFES (Continuous Underway Fish Egg Sampler) applications are under consideration.
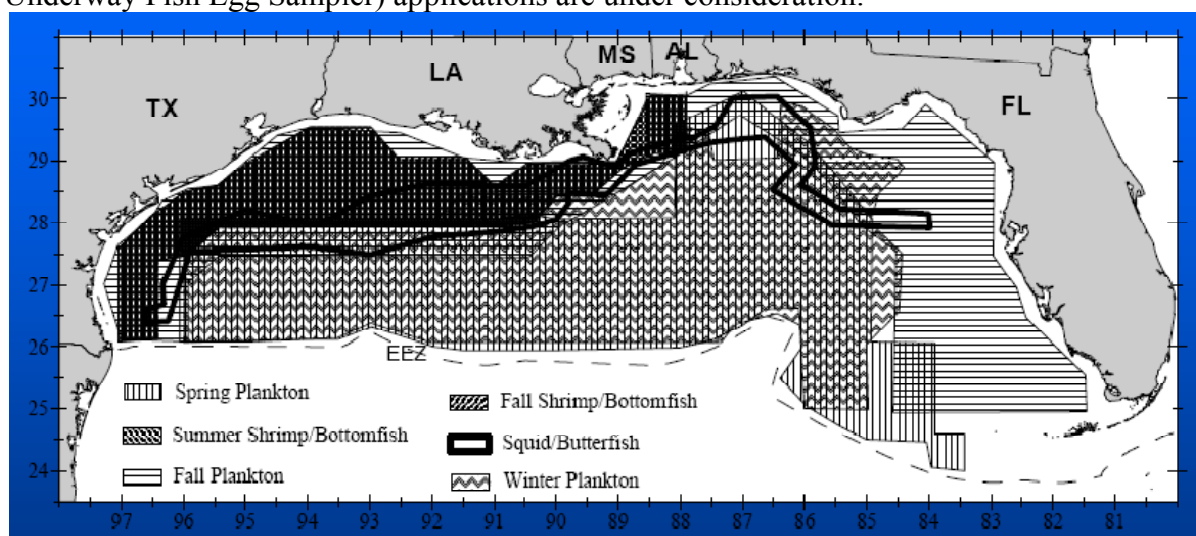

Figure 1. Aerial coverage of plankton surveys.

Several issues concerning plankton surveys were discussed. Almost all indices of abundance developed from these data have large CVs, which could result from insufficient sample sizes. The panel suggested that quadrupling current sample sizes would aid in lowering this variability. Next, since these data are primarily used to index spawning-stock biomass, the panel suggested that sampling eggs would be more beneficial, due to large changes in larval mortality both

temporally and spatially. Also, possible vessel effects were discussed, and proper recording of gear changes and changes in the spatial coverage of sampling was discussed. Finally, the sorting of the plankton samples was discussed, and it was suggested by the panel that a more molecular-based approach should be pursued.

Reef Fish Surveys
After more than a decade of development, the reef fish survey design currently in place was instituted in 1992 using video/trap technology. The survey has been conducted most years since the initial cruise. Video gear has continued to be refined over the time frame, with stereoscopic imaging (for length/mass measures) the next step. The survey could be considerably improved by increasing the sampling frame, especially along the Florida shelf. This would require more detailed bathymetric information on reef structure than is currently available. This video technology is also being used at the Madison – Swanson marine management reserve. After several years of development testing appropriate methods, a randomized bandit reel survey at oil platforms was started in 2007.
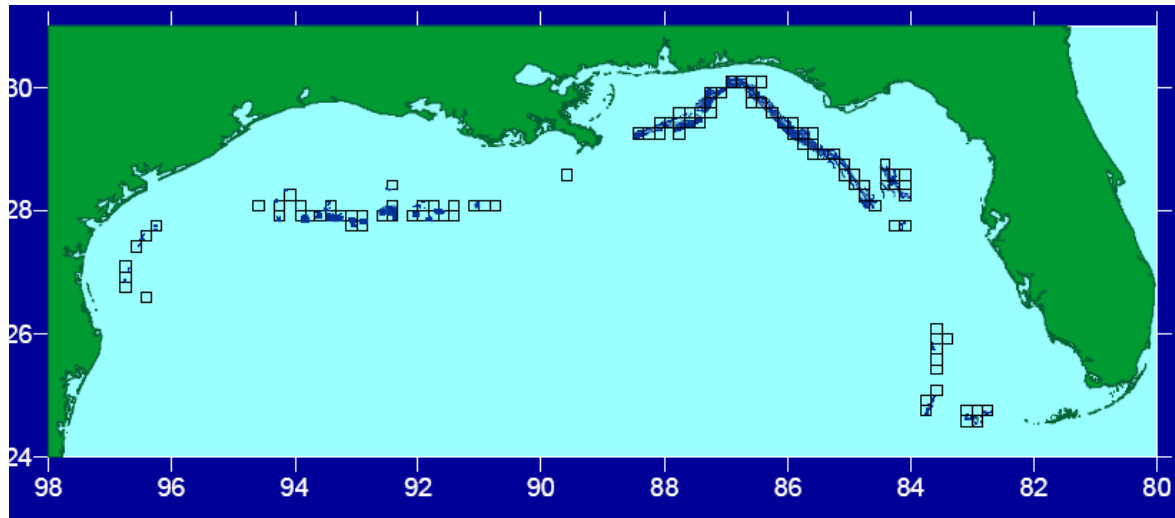

Figure 2. Areas sampled during the reef fish video survey.

Several issues concerning reef fish video surveys were discussed. Almost all indices of abundance developed from these data have large CVs, which could result from insufficient sample sizes. The panel suggested that quadrupling current sample sizes would aid in lowering this variability. However, to accomplish this, the panel suggested that more habitat mapping would be needed, especially in those areas currently not mapped (Figure). Next, for many years the tail ends of the survey were missing due to time constraints or mechanical failures. Also, there are years where there were no data collected due to lack of funding. Finally, this survey is still has a relatively short times series and covers only shelf-edge banks. However, it is very valuable in that this survey constitutes one of the most efficient and non-destructive ways to sample this environment.

Trawl Surveys

Ground Fish Trawl Surveys

After several decades of development, the fall ground fish (*aka* "bottom fish") survey using started in 1972 and the summer component of the ground fish survey started in 1981, with the same format as developed in the fall. A 40' shrimp otter trawl is the gear used for these surveys. It should be noted that many fishery exploratory cruises were conducted on a variety of vessels and a multitude of seasons during the early years (1950 – 1971). Apparent change in resource abundance during this period was one driving force for instituting the statistically based sampling design for resource abundance monitoring. However, information from these early period surveys can prove to be highly valuable in stock assessments and data from these surveys need to be incorporated into the database, to the degree possible.
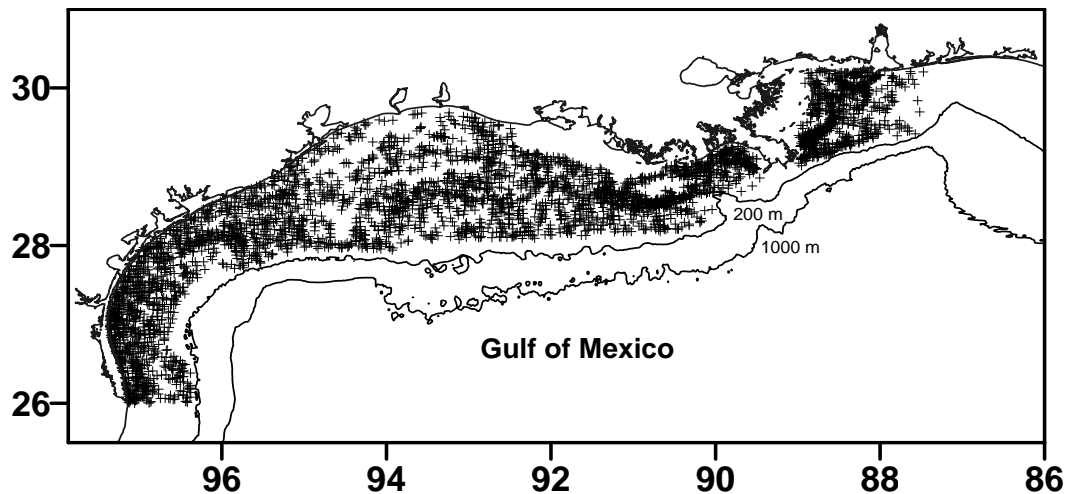

Figure 3. Fall groundfish effort through 2007. Each cross represents a trawl station.
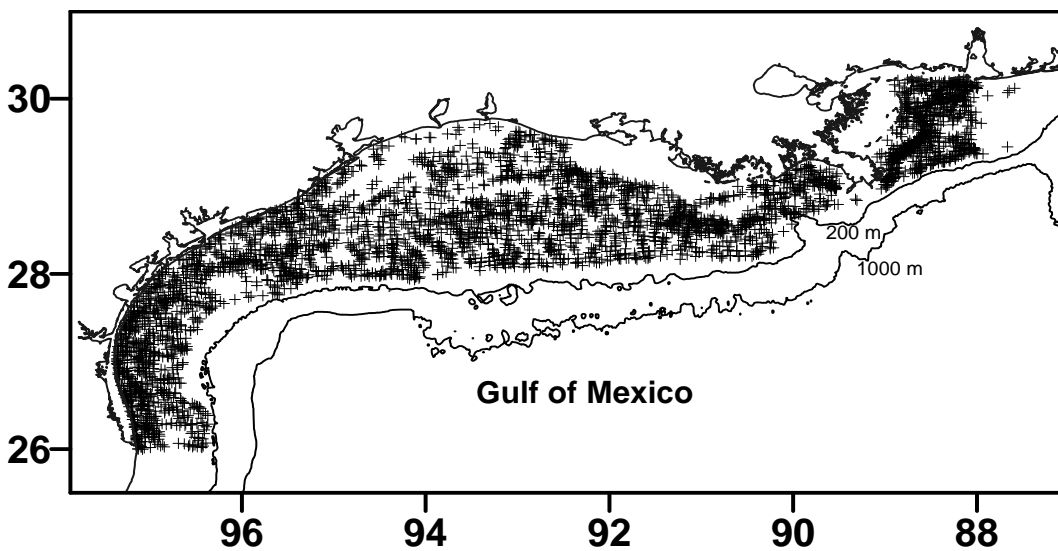

Figure 4. Summer groundfish effort through 2007. Each cross represents a trawl station.

10

Small Pelagics Survey
After a decade of development, the small pelagics survey began in 1988, and has been conducted intermittently since then due to vessel access limitations. The survey uses a 90' high opening bottom trawl (HOBT) for data collection. While the time-series is relatively short (compared to groundfish), the species composition and age structure sampled provides valuable information for stock assessment of Federally managed species.
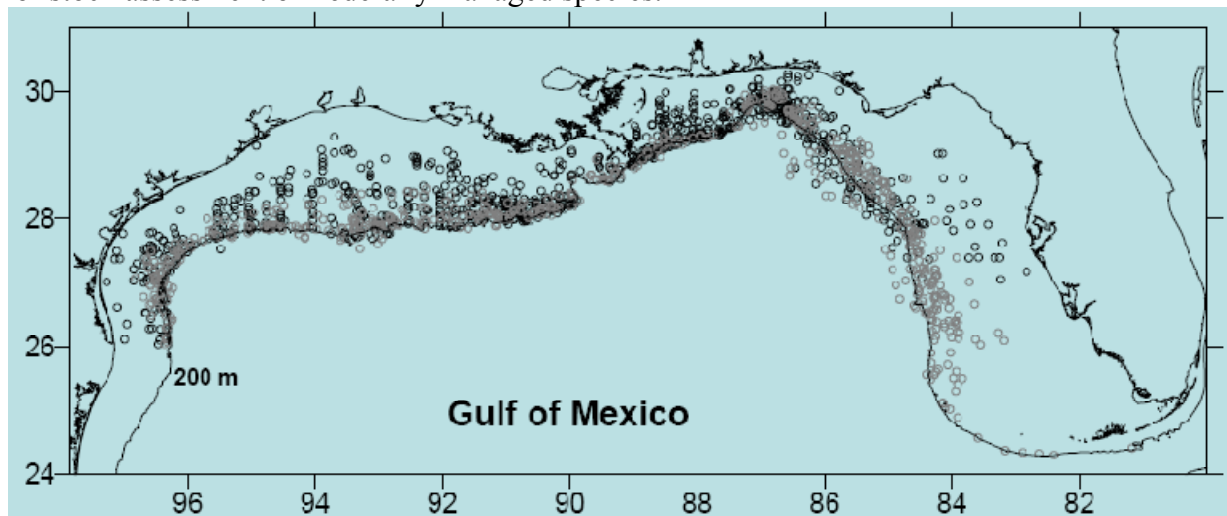

Figure 5. Small pelagics effort through 2007. Each circle represents a trawl station. Black circles represent early surveys (i.e. before 1997), and the gray circle represent recent surveys (after 2001).

Several issues concerning the groundfish and small pelagics surveys were discussed. Almost all indices of abundance developed from these data have large CVs, which could result from insufficient sample sizes. The panel suggested that quadrupling current sample sizes would aid in lowering this variability. Recently (i.e. FY2009), changes were made to the groundfish sampling protocol that allowed in increase in the number of trawling stations sampled per survey without changing the gear or the focus of the survey. Also, the panel discussed the changes in the small pelagics survey design that occurred during the hiatus in the survey during the late 1990s, and how this could affect development of abundance indices. Finally, the panel discussed how trawls collect fish during the surveys. This included discussions of how small pelagics may be primarily caught on ascent, of how the amount of biomass caught on long tows may change the way the gear fishes, and of how trawl cables may stretch and need to be monitored and documented.

Longline Surveys
Longline surveys were begun in 1995 with the bottom longline survey. A standard bottom longline gear is used during the data collection: one-mile main line with 100 hooks, soaked for one hour. There was a change from J-type hooks to circle-type hooks, which took place between 1998 and 2001. These surveys have been used for HMS (particularly shark) stock assessments and have proved invaluable to these assessments. A pelagic longline survey was conducted in 2004 -2006, and a regional inshore bottom longline program was conducted in the last two years.
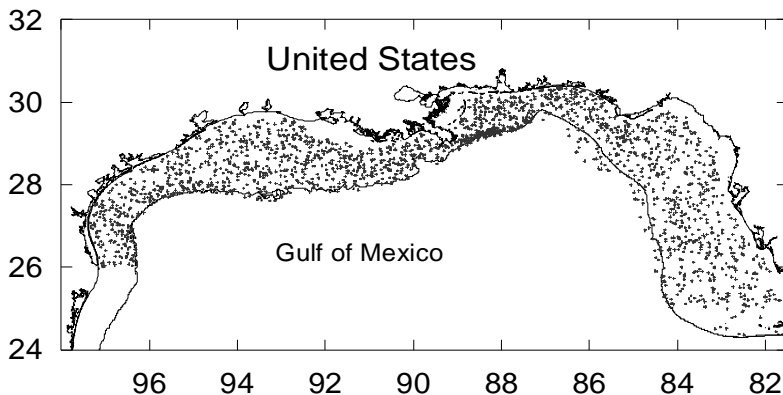
Figure 6. Bottom longline effort through 2007.
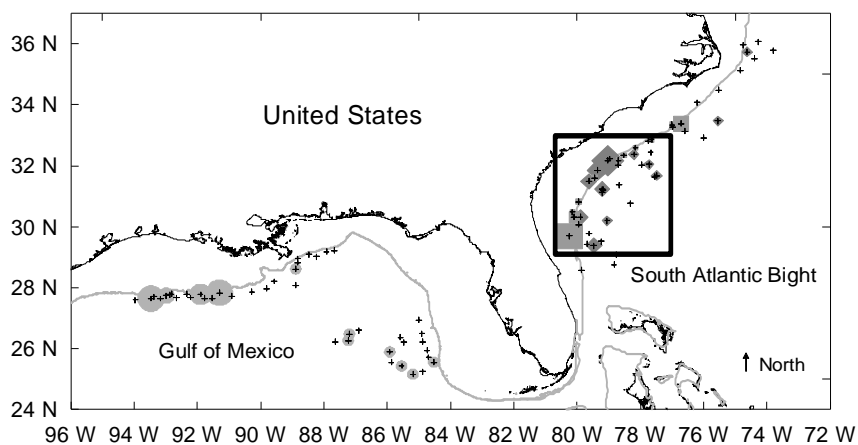
Figure 7. Pelagic longline effort (black crosses) from 2004 – 2006.

The bottom longline surveys are anticipated to be critical for reef fish assessments, but as the time-series remained short as of the most recent assessments for these species, the information has not yet been used in the reef fish stock assessments. The Panel recommends increasing sampling intensity and expansion (further inshore and further into Atlantic waters) of the bottom longline survey. Collaboration with the NEFSC Apex Predator Program which conducts shark longline surveys may be warranted. Almost all indices of abundance developed from these data have large CVs, which could result from insufficient sample sizes. The panel suggested that quadrupling current sample sizes would aid in lowering this variability. The panel recommends a model assisted survey and the use of satellite imagery to aid during future pelagic longline surveys.

2.2. Overview of small scale fishery independent surveys – Camilla McCandless and John K. Carlson

Fishery independent surveys are primarily developed to monitor temporal and spatial changes in the relative or absolute abundance of a target fish population in a manner that is not subject to the biases inherent in data from commercial or recreational fisheries (NRC 1998). Small scale fishery independent surveys allow for focus on a particular life stage and/or area of a target species overall distribution. These surveys often use smaller vessels and reduced gear volume in comparison to large scale surveys. This allows for sampling in shallow areas not accessible to large scale survey vessels. In addition, visual surveys can be conducted in areas that are protected from survey gear such as reef environments.

Access to shallower and protected environments is key for monitoring many juvenile populations of fish and therefore necessary to calculate recruitment indices for these species. For all life stages of a fish population being targeted, the ideal survey should maintain the same gear, area of coverage, and time period throughout the time series of the survey. It is generally assumed that keeping these factors constant over time allows for year-to-year changes in measured abundance to be directly related to true or relative changes in population size, given consistency in the area of coverage and time period that the targeted component of the population is available to the survey (NRC 1998).

Fishery independent surveys are usually standardized to the fullest extent possible. Stratified random designs are the most common designs used for fishery independent surveys (Gunderson 1993). These surveys often include stratification by depth and area but may also reflect management boundaries. Another type of survey design used in fishery independent surveys is the fixed station or fixed transect design, which maintains the same stations or transects each year.

Gear commonly used in small scale fishery independent surveys include trawls, longlines, gillnets, seines, traps, and hook and line as well as the use of visual surveys. Target species for some recent and/or ongoing small scale surveys include sharks (Carlson and Brusher 1999, McCandless 2005), gag grouper (Koenig and Coleman 1998), red porgy (Devries 2006) and multiple reef fish (Gomez 2000, Tobias et al 2002 and Mateo 2002). These surveys vary in geographic coverage from only one or a few small embayments to the majority of coastal waters of a single state.

One common use of fishery independent survey data is to monitor fish populations not targeted by the survey because of lack of data on a particular species or a species life stage. In addition, some surveys monitor multiple species and the survey design cannot be optimal for all species surveyed. Lack of consistent funding for surveys can also cause problems leading to variable effort and discontinuous or short time series. Other issues encountered in using small scale fishery independent data for both targeted and non targeted species monitoring is accounting for gear changes, inadequate spatial or temporal coverage, high proportions of zero catch sets, gear selectivity, and uncontrollable environmental factors. The majority of these issues can be dealt with by using a general linear model (GLM) to reduce the variance in catch per unit effort (CPUE) data caused by these factors, producing indices much more suitable for use in stock

assessments. Factors generally found to influence catch rates are area, season, depth and temperature (correlation).

One of the major difficulties in modeling catch rates for fishery independent surveys is the high incidence of zeroes in the data. Excess zeroes can make traditional approaches to modeling inappropriate, particularly because such models fail to deal with the high overdispersion that can be present. A more direct way to deal with large numbers of zeroes is to model them explicitly, where the presence or absence of a given phenomenon is specified through a Bernoulli distribution and the number of occurrences - given its presence - is specified through some other distributional form – often as a Poisson (Hall 2000), binomial (MacKenzie et al. 2006), or negative binomial (Minami et al. 2007) model. Such models are often called "zero-inflated" or "mixture" models, highlighting the amalgamation of two distinct distributions to distinguish between the presence/absence and count processes that may generate a set of observed data.

One of the more common approaches utilized in the SEDAR process for highly skewed survey data containing a large portion of zeros is based on a delta-lognormal model and is a two-step approach that models the zero catch separately from the positive catch. This method was originally proposed by Lo et al. (1992) for use in analyzing fish spotter data for northern anchovy, *Engraulis mordax*, from the southern California purse-seine fishery. The Lo et al. method for standardizing data can correct the bias that may be introduced into log-normal error models when a significant number of zero catches in the data may cause zero catches with low effort to appear higher. An example of small scale fishery independent data modeled using the Lo et al. method is included for the NMFS survey of juvenile sandbar sharks in Delaware Bay (Table 1, taken from McCandless 2005). By using the Lo et al. method to model the data from this survey it provides a reduction in variance associated with region, month and depth in the CPUE data in addition to limiting the effects of the abundant zero catches in this dataset. In this example, standardization of the data explains between 10 and 20% of the deviance between the null and the final models with the spatial factor, region, explaining the majority of the deviance (Table 1). All diagnostic plots met assumptions in this example, and supported an acceptable fit to the selected model. However, catch rate indices developed for juvenile blacknose shark using the delta-lognormal approach resulted in a poor fit and the resulting model was not utilized (Carlson and Bethea 2007). Thus, the necessity of alternate distributions needs to be explored more frequently when modeling fishery independent data.

References

Carlson, J.K. and D.M. Bethea. 2007. Standardized catch rates of small coastal sharks from a fishery-independent gillnet survey in northwest Florida. SEDAR 13-DW-06.

Carlson, J.K. and J.H. Brusher. 1999. An index of abundance for coastal species of juvenile sharks from the northeast Gulf of Mexico. Marine Fisheries Review 61(3):37-45.

Devries, D.A. 2006. The life history, reproductive ecology, and demography of the red porgy, *Pagrus pagrus*, in the northeastern Gulf of Mexico. Dissertation. Florida State University.

Gomez, R.M. 2000. Fisheries independent sampling of the shallow water reef resources in the U.S. Virgin Islands. Caribbean/NMFS Cooperative SEAMAP Program. Bureau of Fisheries, Division of Fish & Wildlife, Department of Planning and Natural Resources, U.S. Virgin Islands, 7 pp.

Hall, D. 2000. Zero-inflated Poisson and binomial regression with random effects: a case study. Biometrics 56:1030–1039

Gunderson, D.R. 1993. Surveys of fisheries resources. John Wiley & Sons, Inc., New York. 248 pp.

Koenig C.C., Coleman F.C. 1998. Absolute abundance and survival of juvenile gags in sea grass beds of the northeastern Gulf of Mexico. Transactions of the American Fisheries Society 127:44–55.

Lo, N.C., L.D. Jacobson, and J.L. Squire. 1992. Indices of relative abundance from fish spotter data based on delta-lognormal models. Canadian Journal of Fisheries and Aquatic Sciences 49:2515-2526.

Mateo, I. 2002. Nearshore habitats as nursery grounds for recreationally important fishes. U.S. Fish and Wildlife Service, Sport Fish Restoration Program, U.S. Virgin Islands, 72 pp.

MacKenzie, D., Nichols, J., Royle, J., Pollock, K., Hines, J., and Bailey, L. 2006. Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence. Elsevier, San Diego, USA. 324 pp.

Minami, M., Lennert-Cody, C., Gao, W., and Román-Verdesoto, M. 2007. Modeling shark bycatch: the zero-inflated negative binomial regression model with smoothing. Fisheries Research 84:210–221.

McCandless, C.T. 2005. Large Coastal Shark 05/06 SEDAR Data Workshop Document: Relative abundance trends for juvenile sandbar sharks in Delaware Bay. LCS05/06-DW-30.

NRC (National Research Council). 1998. Improving fish stock assessments. National Academies Press, Washington DC, 177 pp.

Pennington, M. 1983. Efficient estimators of abundance for fish and plankton surveys. Biometrics 39:281-286.

Pennington, M. 1996. Estimating the mean and variance from highly skewed marine data. Fishery Bulletin 94:498-505.

Tobias, W., W. Troller, W. Ventura, and H. Rivera. 2002. St. Croix fisheries independent trap and line survey. SEAMAP-C USVI. Division of Fish & Wildlife, Department of Planning and Natural Resources, U.S. Virgin Islands, 27 pp.

Table 1. Results of the stepwise procedure for development of the catch rate model for juvenile age 1+ sandbar sharks captured by longline in Delaware Bay. %DIF is the percent difference in deviance/DF between each model and the null model. Delta% is the difference in deviance/DF between the newly included factor and the previous entered factor in the model. L is the log likelihood.

PROPORTION POSITIVE-BINOMIAL ERROR DISTRIBUTION

FACTOR	DF	DEVIANCE	DEVIANCE/DF	%DIFF	DELTA%	L	CHISQ	PR>CHI
NULL	279	384.4973	1.3781					
REGION	271	318.0197	1.1735	14.8465	14.8465	-159.0099	66.48	<.0001
YEAR	275	373.5627	1.3584	1.4295		-186.7813	10.93	0.0273
MONTH	278	379.7936	1.3662	0.8635		-189.8968	4.70	0.0301
DEPTH	246	381.0803	1.3807	-0.1887		-190.5401	3.42	0.3317
REGION +								
YEAR	267	303.8540	1.1380	17.4225	2.5760	-151.9270	14.17	0.0068
MONTH	270	311.9419	1.1553	16.1672		-155.9710	6.08	0.0137
DEPTH	268	310.1205	1.1572	16.0293		-155.0603	7.90	0.0481
REGION + YEAR +								
MONTH	266	297.4271	1.1181	18.8666	1.4440	-148.7135	6.43	0.0112
DEPTH	264	295.4705	1.1192	18.7867		-147.7352	8.38	0.0387
REGION + YEAR + MONTH +								
DEPTH	263	288.8211	1.0982	20.3106	1.4440	-144.4106	8.61	0.0350

FINAL MODEL: REGION + YEAR + MONTH + DEPTH

Akaike's information criterion -656.9

Schwartz's Bayesian criterion -658.7

(-2) Res Log likelihood 1311.8

Type 3 Test of Fixed Effects

	REGION	YEAR	MONTH	DEPTH
Significance (Pr>Chi) of Type 3 test of fixed effects for each factor	<.0001	0.0098	0.0132	0.0485
DF	8	4	1	3
CHI SQUARE	53.62	13.32	6.14	7.88

16

Table 1. continued

POSITIVE CATCHES-POISSON ERROR DISTRIBUTION

FACTOR	DF	DEVIANCE	DEVIANCE/DF	%DIFF	DELTA%	L	CHISQ	PR>CHI
NULL	155	1051.4106	6.7833					
REGION	147	961.8941	6.5435	3.5352	3.5352	1611.1794	89.52	<.0001
MONTH	154	1017.9457	6.6100	2.5548		1583.1536	33.46	<.0001
YEAR	151	1005.4245	6.6584	1.8413		1589.4142	45.99	<.0001
DEPTH	152	1029.8094	6.7751	0.1209		1577.2218	21.60	<.0001
REGION +								
MONTH	146	919.3252	6.2967	7.1735	3.6383	1632.4639	42.57	<.0001
YEAR	143	915.6459	6.4031	5.6049		1634.3035	46.25	<.0001
DEPTH	144	935.4449	6.4961	4.2339		1624.4040	26.45	<.0001
REGION + MONTH +								
YEAR	142	867.3226	6.1079	9.9568	2.7833	1658.4652	52.00	<.0001
DEPTH	143	895.5579	6.2626	7.6762		1644.3475	23.77	<.0001
REGION + MONTH + YEAR +								
DEPTH	139	840.1780	6.0444	10.8929	0.9361	1672.0375	27.14	<.0001

FINAL MODEL: REGION + MONTH + YEAR

Akaike's information criterion   -213.3

Schwartz's Bayesian criterion   -214.8

(-2) Res Log likelihood   424.7

	Type 3 Test of Fixed Effects		
Significance (Pr>Chi) of Type 3	REGION	MONTH	YEAR
test of fixed effects for each factor	0.2120	0.0103	0.1392
DF	8	1	4
CHI SQUARE	10.82	6.58	6.94

17

2.3. Discussion and Determination of Main Issues of Fishery Independent Surveys

Several issues were discussed concerning fishery-independent indices. One of the most discussed topics was of coefficients of variation (CVs). Theoretical CVs and GLM estimated CVs as a function of the number of samples for selected species and surveys are shown in Figure 1, and indicate the need for an increase in sampling effort with all surveys. The theoretical and GLM estimates may differ in any particular year owing to the unbalanced sampling designs used for the various surveys.
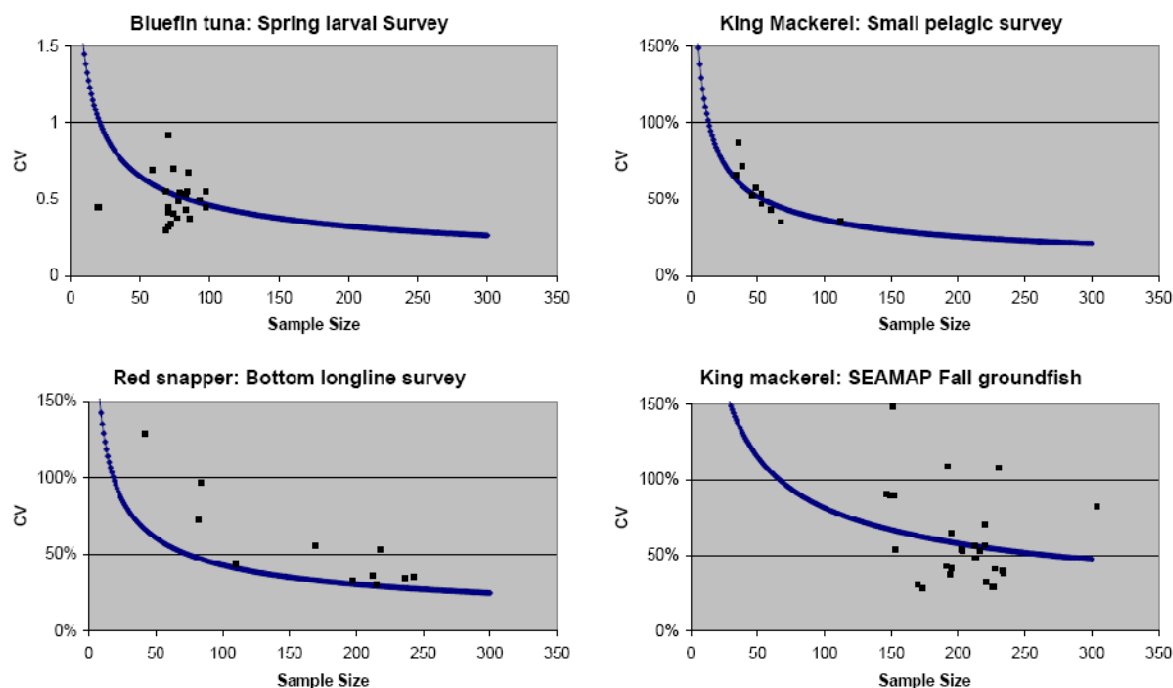
Figure 1. Theoretical CVs (lines) and GLM estimated CVs (points) as a function of the number of samples for selected species and surveys.

Many other topics, related to fishery independent indices, were discussed and are outlined below.

Large scale fishery independent surveys: Issues and Recommendations
- There may be vessel effects and paired samples may be needed.
- Need yearly maps of CPUE and sampling effort in submitted document.
- Amount of the spatial distribution of species covered in survey needs to at least be roughly estimate documented in contributed paper.
- When the SEDAR Indices Working Group has many options (i.e. many possible indices to choose from), we should look at length of time series to help pare down.
- More fishery independent indices in the U.S. South Atlantic and Gulf of Mexico are needed, and there needs to be enhancements in current surveys or new surveys created.

Small scale fishery independent surveys: Issues and Recommendations
- Issues
  — gear changes
  — area coverage

- — temporal coverage
- — selectivity
- — high proportion of zeros
- — uncontrollable environmental factors
- — missing years
- — short time series
- — variable effort
- — conflicting trends
- Recommendations
  - — There is a need for GLM standardization with most datasets.
  - — Environmental factors need to be considered.
  - — Researchers must be careful not to "go fishing" or include all factors just because they are recorded.
  - — Researchers need to be careful of inaccurate correlations and overparameterization.
  - — Some environmental factors need to be left in as year effects because want to see these decreases in the population for less mobile populations. This is not a problem for mobile or HMS populations.
  - — Researchers need to consider type of error distribution for the model – lognormal, delta lognormal, Poisson, negative binomial, or zero-inflated distributions.
  - — For conflicting trends, researchers need to look at reasons why the trends may conflict within sampling design. Often localized surveys have differing trends than those of larger-scaled surveys (e.g. red tides/high rainfall). Researchers need to look at the big picture to assess how index trends correspond.

3. Fishery Dependent Indices

3.1. Fisheries dependent commercial data and its use in constructing indices of abundance – Kevin McCarthy

For many species in the US south Atlantic, Gulf of Mexico, and Caribbean, fisheries independent data sets useful for the construction of indices of abundance may be few in number, of short time series, of limited spatial extent, or lacking altogether. Fisheries dependent data, therefore, are frequently required for stock assessments. A brief review of many of the commercial data sets and general characteristics of fisheries dependent commercial data, including advantages, disadvantages, known issues and possible solutions, is presented below. This review also highlights areas of concern that have been addressed repeatedly during numerous SEDARs.

Available Datasets

Two broad categories of fisheries dependent commercial data are available: data collected by trained observers or dockside interviewers and self-reported data. While not meant to be a comprehensive accounting, many of the data sets listed below have consistently used during SEDAR or ICCAT assessments.

Observer/interviewer data

Trip Interview Program (TIP) – dockside interviews
Pelagic longline – observers
Shark bottom longline (Gulf of Mexico and US south Atlantic) – observers
Shark gillnet – observers
Gulf of Mexico reef fish – observers
Shrimp observer program – observers

Self-reported data

Trip ticket (programs run by individual states)
Coastal logbook (trip based form required of all vessels with Federal permits)
Pelagic logbook (required of all US flagged vessels with Highly Migratory Species permits)
Puerto Rico sales tickets
US Virgin Islands landings reports
Golden Crab logbook
Wreckfish logbook
Gulf shrimp statistics
South Atlantic shrimp statistics

Data collected by observers or samplers are usually more detailed than self-reported data. Accurate species identification may be assumed for observer/sampler collected data. Such data sets will often include size information of the catch or landings. Observers collected information on the total catch for a trip or set, unlike most self-reported or interview data that are usually limited to landings information. Observer data provides fine scale spatial information, detailed effort information, detailed gear configuration information, as well as, additional details of fishing behavior and environmental conditions. Observer data also provides discard information when total catch is recorded.

Self-reported data will be limited in reliability by the honesty and accuracy of the reporting fishers. While completely dishonest reporting may be rare and limited to a few individuals, it is difficult to quantify. More problematic is inaccurate reporting, which may be more widespread and is also difficult to quantify. Species misidentification has occurred, e.g. systematic misreporting of gag and black grouper in the northern Gulf of Mexico because the local common name for gag grouper is black grouper. Self-reported data is usually less detailed than is observer data, often because of the limitations of the reporting forms. For example, in the US Virgin Islands data are reported by species group rather than by species. In many cases, spatial information is reported as occurring within large, predefined areas and not to the fine scale available in observer data. Self-reported data typically includes landings and not total catch. There are some exceptions, e.g. those vessels the voluntarily report discards in addition to landings and effort to the pelagic logbook program. Such discard information is limited, however, to voluntary reporting in the pelagic logbook program and to a 20% random selection of vessels reporting to the coastal logbook program. Collection of accurate effort information has been problematic for some self-reported data collection programs. The problem has been gear specific (e.g. traps, bottom longlines) in the coastal logbook program, but has been more widespread in the trip ticket, Puerto Rico, and US Virgin Islands data collection programs. In those cases, hours fished and the amount of gear fished (e.g. number of lines and hooks) are often unknown or unreported. Effort can only be defined as days at sea or per trip. In most cases, self-reported data are trip based, however in the pelagic logbook, data are set based. Finally, size information of individual animals in the landings is usually lacking in the self-reported data. The pelagic logbook program, however, does have size information of individuals for some species.

The time series for most of the fishery dependent commercial data sets listed above include more than 10 years of data (Figure 1). Often the initial years of data collection involved standardizing data collection procedures and changes in the universe of sampled vessels. For example the coastal logbook program began in the Gulf of Mexico in 1990, but, beginning in 1992 vessels fishing in the US south Atlantic were required to report. In addition, prior to 1993 only a 20% subsample of Florida vessels were required to report to the coastal logbook program, however all vessels with Federal permits were required to report beginning in 1993. The useful time series for some data sets may be less than the total time series contained within the data set.

The spatial range of data included in fishery dependent commercial data sets is usually extensive (Figure 2). Spatial resolution may be coarse in the case of self-reported data, however, pelagic logbook data are an exception to this generalization. Observer data are usually of finer spatial resolution. Additional spatial limitations are often associated with particular data sets. Most of the data sets listed above contain data from fishing trips in Federal waters, although individual states may have commercial data available from state waters, e.g. trip ticket data. Similarly, some gears are spatially restricted. Longlines in the eastern Gulf of Mexico are restricted to depths greater than 20 fathoms and in the western Gulf of Mexico a restriction limits longlines to depths greater than 50 fathoms. Fish traps have been banned from the Gulf of Mexico entirely. The pelagic logbook and observer data sets include data from all US flagged vessels landing Highly Migratory Species regardless of where the fishing occurred. All fisheries dependent data collection will be affected by closed areas, both seasonal and year-round closures. Some of the data sets (trip ticket and Puerto Rico sales ticket) do not record area fished information, but record the port and county landed. Depth fished is available in most of the data

21

sets listed at least for a portion of the time series, however, depth information is not available in the trip ticket data or the US Virgin Islands landings reports.

Fishery dependent commercial data may contain catch and effort information from a variety of fishing gears, all of which may be used to prosecute the same fishery. Those different fishing gears may harvest animals from different parts of the species range, e.g. bottom longlines in the Gulf of Mexico are limited to particular depths and may result in fishing farther offshore than vessels using other gears. Multiple gear specific indices of abundance may be constructed from those commercial data sets. In some cases, however, data are reported for only one or a few fishing gears, e.g. the pelagic longline observer data set.

Effort measures vary widely among commercial data sets. Hook hour fished may usually be determined from the available data in observer/interviewer and logbook data sets. Equally fine scale effort measures may be obtained for other gears where hook hour is not appropriate, e.g. gillnets. Some data sets, however, either do not record or are frequently missing data necessary fine scale effort measures. Trip ticket, Puerto Rico, and US Virgin Islands effort measures are often limited to days at sea or trip.

Many types of commercial fishing gears are likely poor methods of sampling the smallest individuals in a population. The size composition of the population will, therefore, not be reflected fisheries dependent data sets even in cases where total catch is known. Landings data will not include any size information of sub-legal sized individuals. Size of individuals in the catch/landings is typically not available in self-reported data. The size composition of the catch is, obviously, unknown other than the assumption that all individuals are of legal size. Even the number of individuals is unknown for most self-reported logbook data sets; landings are typically reported in weight. The pelagic logbook is again an exception. Individual weight of some species (e.g. swordfish, some tunas) is reported to the pelagic logbook. Unlike most self-reported data, observer/interviewer collected data does include size composition of the catch or landings.

Data availability for characterizing gears, vessels, areas fished, fishing behavior, etc. is highly variable among data sets. Observer collected data sets usually contain more detailed information than do self-reported data sets. The level of detail in self-reported data sets varies from no information (e.g. trip ticket data sets have no information regarding gear configuration) to the highly detailed information available in the pelagic logbook data set which includes hook size, type, and offset data, for example. Vessel information is frequently not available in self-reported data, but may be obtained from other sources such as permit records. Those records often contain vessel length and horsepower and may contain additional information such has hold capacity. Frequently, however, permit records lack vessel information. Information about the vessel captain's years of fishing experience for example, is generally not available in self-reported data and may not be directly available in any data set.

Prior SEDAR commercial data issues

Species misidentification/misreporting was recognized as a problem at the SEDAR 10 data workshop (Chih and Turner, SEDAR10-DW24). Gag grouper was frequently misreported as black grouper in the northern Gulf of Mexico primarily because black grouper was the local common name for gag grouper. The problem was addressed by determining area specific gag:black grouper landings ratios from TIP data. The observed TIP ratios were then applied to the combined gag and black grouper landings for each area to obtain more accurate landings totals. For CPUE calculations, the TIP ratio was applied to landings reported for individual trips.

A fundamental issue for index of abundance utility in stock assessments is the assumption that CPUE is correlated with abundance. This issue is especially problematic for fisheries dependent commercial indices given that knowledgeable fishers may maintain high CPUE even during declining stock size. One component of effort, search time, is usually not measured accurately, if at all. Total effort is, therefore, not included in most indices of abundance constructed using commercial fisheries data, rather effort spent actually fishing is included in index construction. An apparent lack of correlation between CPUE and abundance was a significant problem during the SEDAR 14 queen conch assessment. Conch yearly mean CPUE remained relatively constant over the period 1992-2006, however, landings increased approximately eight-fold during the same period. A combination of conch life history characteristics, trip limit regulations, and the unavailability of fine scale effort data including search time contributed to the construction of an index of abundance that was not correlated with conch stock abundance (Diaz, SEDAR14-AW-04). No assessment of queen conch was possible using a production model dependent upon indices of abundance and fisheries independent methods were recommended for future queen conch assessments.

Changing catchability may also affect indices of abundance constructed from fisheries dependent commercial data. Improvements in technology (e.g. GPS), gear changes (e.g. hook design), environmental conditions (e.g. red tides), and fisher experience will all potentially affect CPUE. Observer data sets are most likely to include some data relevant to a few potential causes of changing catchability. Other potential sources of changing catchability, such as fisher experience, are usually not recorded. Self-reported data frequently do not have the detail to examine possible sources of changing catchability. Many potential sources of changes in catchability are not measured at all in the available commercial data sets. As noted in the queen conch example above, a further limitation of fisheries dependent data is that search time, likely affected by technology and fisher experience, is frequently not recorded.

Among the more difficult tasks associated with constructing indices of abundance is identifying fishing effort that had targeted, but not caught, the species of interest. This issue is critical to the construction of indices of abundance. Targeting information may be available in observer collected data sets and, presumably, in species specific self-reported logbook data such as the golden crab logbook data set. Other self-reported data sets, including those that are frequently used in the SEDAR process, do not include targeting information, therefore, targeting must be determined by other methods. Most frequently in recent SEDARs, targeted effort has been identified using the Stephens and MacCall (2004) method. An alternative approach to identify targeted effort involves determining particular gear configurations that were used to catch the species of interest. This method is limited by the detail of the available data. More detailed data will potentially allow for identifying gear configurations specific to the species of interest, while less detailed information will likely result in identifying gear configurations that may be used when targeting a variety of species. Observer data has both the most detailed data and is most likely to have targeting information. Self-reported data usually has the least detailed data and has no specific targeting information.

Regulatory measures may directly affect CPUE and must be accounted for when construction indices of abundance from fisheries dependent data. The regulatory measures of concern include: regulatory geographic boundaries, minimum size limits, fishery closures, and bag/trip limits. Among species assessed under the SEDAR process, king mackerel was perhaps the species with the most numerous regulatory changes and will be used as an example of how regulatory measures may be accommodated when constructing indices of abundance.

Geographic boundaries may define regions with differing fishery regulations and those boundaries may have changed over time, may have been temporary or permanent, or may have changed seasonally. In the case of king mackerel, regulatory boundaries changed over time (the eastern Gulf of Mexico was subdivided into two zones in 2000) and also changed seasonally (regulatory jurisdiction of most of the east coast of Florida and Monroe County changes seasonally). If the area fished was known, the regulations appropriate to particular fishing trips will also be known. Some data sets contain only county or port where the catch was landed and identifying regulations applicable for some trips may be problematic.

Minimum size limits have been in place for most, if not all, species assessed under the SEDAR process. In the case of king mackerel, three minimum size limits had been implemented over a nine year period. Specific dates of implementation differed among regulatory region. Several approaches have been used to address changes in minimum size of landed fish when constructing indices of abundance. Separate indices may be constructed for the periods before and after change in a minimum size limit. Such an approach may result in indices with short time series. Multiple size limit changes, as with king mackerel, will result in the construction of multiple short time series indices. A second approach to the minimum size limit issue uses observer or TIP data to determine if size composition of landed animals changed after the implementation of or change to a size limit. If the size of landed individuals did not change after the size limit began, it may be assumed that the size limit had no affect on fishing behavior and indices need not be split at the time the size limit became effective. An "informed judgment" approach, using fisher observation and knowledge of the length composition of landings before and after size limit implementation, is another possible method for determining whether indices of abundance should be split before and after size limit regulations were imposed. Finally, some population models (e.g. Stock Synthesis 3) can account for changes in minimum size and the SEDAR 16 king mackerel data workshop recommended allowing the population model to address the minimum size issue.

The third regulatory measure, fishery closures, includes regular seasonal closures and closures due to quota limits. Regional closures are somewhat similar and may be addressed in a similar fashion. This type of regulation has been common among SEDAR assessed species. In the western Gulf of Mexico, the king mackerel fishery was closed on 24 separate occasions between 1983 and 2006. In the US south Atlantic, the king mackerel fishery was closed only twice during that period. In some regions and years, the king mackerel fishery was closed twice in a single year. The approach used in the construction of fisheries dependent commercial indices of abundance has been to exclude data from periods of fishery closures. So doing ensures that yearly mean CPUE was not lowered by erroneously including effort which, due to the closure, resulted in no landings.

A trip limit is the final regulatory measure which has presented problems during the construction of fisheries dependent indices of abundance. This regulation becomes a matter of concern when using data from self-reported data sets that are trip based. If vessels reach the trip or bag limit and then target other species on the same trip, effort cannot be apportioned unambiguously due to the nature of reporting effort in those data sets. Indices of abundance constructed from those data will be affected. King mackerel trip limits varied from 25 fish to 3,500 pounds of fish depending upon the management region and date of the trip. Within single regions, trip limits varied both among and within years. That complicated array of differing trip limits made index construction more difficult than usual. Since 1995, all regulatory regions had

king mackerel trip limits in effect; therefore, regions with trip limits could not simply be excluded from analyses.

A lengthy effort was made to determine the extent to which trip limits may have affected king mackerel fishing effort (McCarthy et. al, SEDAR16-AW-02). Similar data exploration should be repeated for other species before constructing fisheries dependent indices if trip limits were in place. Trip limits in effect for the date and region in which a trip occurred were identified. All trips reporting king mackerel landings were classified by the appropriate trip limit and by the percentage of king mackerel to total landings: <25% king mackerel, 25-50%, 50-75%, and >75%. The landings categories served as an approximation of king mackerel targeting. Potential trip limit effects were examined for each trip limit – targeting category.

For each trip limit – targeting category combination, scatter plots of trip specific landings were plotted by year to provide an overview of the range of landings among trips. Frequency histograms of trips by pounds landed were constructed to determine the frequency of trips that reported reaching or exceeding the trip limit. Plots of the distribution of trips across reported landings for the south Atlantic under a 3,500 pound trip limit are provided in Figure 3. The south Atlantic trip limit appeared to have had no effect on fishing effort. Vessels rarely if ever reached or exceeded the trip limit, regardless of the targeting category of the trip. Figure 4 contains plots of trip specific landings for the Florida Keys under a 500 pound trip limit. A large number of the >75% king mackerel landings trips reached or exceeded the trip limit. Trips classified in the other three targeting categories rarely reached the trip limit and the trip limit likely did not affect fishing effort for those trips. Including those data, therefore, will not adversely affect index construction. Those trips that fell within the >75% targeting category were assumed to catch king mackerel until the trip limit was reached and then returned to port with little or no additional fishing effort directed at other species. The recommendation following the examination of those data was that all trips in the Keys zone under a 500 pound trip limit may be included for index construction.

If trips reporting primarily king mackerel landings (>75%) reach or exceed the trip limit, a meaningful index may still be constructed provided hours fished was reported, as in the Keys example above. If effort is measured only by day or days, as is sometimes the case in the trip ticket data, then indices so constructed will be biased when trip limits are reached. For trips that report less than 75% king mackerel in the landings and also reach the trip limit, appropriately determining king mackerel effort is problematic with the available commercial data. Total hours fished is reported to the coastal logbook program (hours fished or days at sea are reported on trip tickets) and cannot be divided among the species landed. If, however, trips with landings containing less than 75% king mackerel do not reach the trip limit (as in the above examples) then those trips are presumed to have had a consistent effort across species. Data from particular targeting categories in which effort may have been affected by trip limits could be excluded from index construction. Indices could be constructed using data from trips in those targeting categories in which effort was not affected by trip limits.

Elements for SEDAR documents

Suggested elements for any SEDAR document describing the construction of a fisheries dependent commercial index include the following:

Identify the data set

Temporal and spatial range of the data
Identify the measure of effort (hook hours, set, day at sea, trip, etc.)
Describe how targeted trips (sets, trawls, etc.) were identified
Describe age/size structure of the catch, if know
Identify whether data landings or total catch
Include a table of factors/variable considered in the analysis (e.g. season, hook size, vessel length, etc.)
Include final model(s) used in the analysis
Describe methods used to address fishing regulations

<u>Literature cited</u>

Chih, C. and S. Turner. 2006. Estimation of species misidentification in the commercial landing data of gag groupers and black groupers in the Gulf of Mexico. SEDAR10-DW- 24. Sustainable Fisheries Division Contribution No. SFD-2006-005.

Diaz, G. 2007. On diver catch-per-unit-effort series as measures of the relative abundance of queen conch and their use in stock assessments for the islands of Puerto Rico and Saint Croix. SEDAR14-AW-04. Sustainable Fisheries Division Contribution SFD-2007-019.

McCarthy, K., S. Cass-Calay, M. Ortiz, and J. Walter. 2008. Effects of King Mackerel Fishing Regulations on the Construction of Fisheries Dependent Indices of Abundance. SEDAR16-AW-02. Sustainable Fisheries Division Contribution SFD-2008-015.

Stephens, A. and A. MacCall. 2004. A multispecies approach to subsetting logbook data for purposes of estimating CPUE. Fish. Res. 70:299-310.

4. Analytical Approaches and Issues

4.1. Spatial considerations for catch per unit effort standardization – John Walter

Incorporating spatial considerations into CPUE standardization presents an imperative to provide spatially-explicit indicators of population status as well as a challenge to avoid the pitfalls of improperly accounting for spatial changes in either the fishery or the resource (Walters 2003). The first task represents a higher order task necessary for spatial management that is more important for Tier II or III stock assessments under the Stock Assessment Improvement Panel hierarchy of assessment needs (NMFS 2001). The second task represents probably the greatest present challenge for SEDAR assessments because it directly affects the defensibility of current assessments. In this document we address the first task briefly but refer the reader to other references for greater details regarding obtaining spatially-explicit indices. We treat the second task in more detail by defining the problems of improperly accounting for space in CPUE standardization and providing a series of recommendations to minimize the potential for these problems to occur in SEDAR assessments. Concurrent with this meeting, the Interamerican Tropical Tuna Commission (IATTC) conducted a similar workshop dealing with spatial issues in CPUE standardization and the reader is referred to the workshop website for further details (http://www.fisheriesstockassessment.com/TikiWiki/tiki-index.php?page=IATTC+October+Stock+Assessment+Methodology+Workshops).

Task I Spatial index development

As stock assessment becomes increasingly spatially-explicit to meet management objectives or to account for stock structure, greater spatial definition of both fishery-dependent and independent indices will be required. Methods to obtain spatially-explicit indices range from simple post-stratification of available data to obtain regional indices to more elaborate spatial generalized linear models (GLMs), generalized additive models (GAMs), neural networks to full geostatistical models (Venables & Dichmont 2004, Maunder & Punt 2004). For fishery-independent data collection programs, construction of spatial indices will be rather straightforward as this data is generally collected according to a design-based sampling regime. For fishery data which lacks a sampling design, often possesses substantial and unknown sets of selection biases or lacks spatially-explicit catch and effort data, construction of spatial indices will be more difficult and limited to the spatial precision of the data.

For most species assessed by SEDAR, catch and effort data does not exist at the finest scale of spatial resolution, i.e., the tow, net haul or line drop location and is aggregated to the coarser resolution of shrimp grid, statistical area or county. Vessel monitoring system data, electronic logbooks and observer coverage data will likely improve the spatial resolution of catch and effort data, vastly increasing the quality and quantity of data for developing spatial indices.

The creation of spatially-explicit indices implies that this structure affects the population and should be accounted for within the modeling context. This can be performed either by constructing separate population models or by modeling movement or exchange between subpopulations (Punt et al. 2000). Increasing the spatial structure of assessment models will require extensive model development to handle this structure as well as data collection and

27

analysis to parameterize spatial biological parameters (growth, mortality, etc), and tagging, tracking and stock composition analysis to determine connectivity between subpopulations (Sibert et al. 2000, Cadrin et al. 2004).

Task II Properly considering spatial issues in SEDAR assessment: definition of the problem:

Walters (2003) details the problems associated with improperly accounting for changes in the spatial extent of the fishery which can produce incorrect CPUE trends and in constructing indices based upon the simple ratio estimator of total catch/total effort. Incorrect CPUE trends can occur as hyperdepletion, where the index declines faster than the overall population or hyperinflation, where the index remains high while the population declines. Both scenarios are dangerous for assessments as CPUE does not reflect population abundance, and both are often a result of spatial variation in the fishery. Walters (2003) recommends that analysts define the fishery as a series of spatial cells and calculate CPUE as a stratified mean of the spatial cells according stratified random sampling (StRS) theory (Cochran 1963). This follows from the recommendation of DeLury (1951) that while fishery-dependent data will never be random, it may be possible through stratification to create sub-populations that, though not strictly random, are reasonably representative to be treated as such. We second this recommendation; however, the spatial resolution of much fishery-dependent data presents a limitation for many SEDAR-assessed species.

The ratio estimator, the StRS estimator and most GLM-based CPUE standardization methods assume that catch and effort is essentially random within any spatial cell so that observations are independent and identically (iid) distributed. Non-random changes in fishing location within a cell, or over the entire area to target areas of high abundance or to avoid undesirable locations represents the most common violation of the assumption of iid observations. The violation of the iid requirement generally manifests itself as lower variances than if the samples had been collected according to a design-based survey. Such variances do not greatly diminish the utility of the indices when used in a relative capacity, assuming that the biases are the same from year to year.

However, the targeting of specific locations can impart a more problematic bias when different locations are fished in different years, and particularly when these changes in fishing location are due to sequential depletion of certain areas, avoidance of non-target or bycatch species or other factors that are not due to random movement of fish. Certain spatial cells may not be fished in all years and CPUE will solely be a function of the fished cells (Walters 2003). This may require either back-filling unfished cells with some value representing catch rates if the cells were fished or some making assumption that these changes in fishing location have no impact upon the abundance (Walters 2003).

Few definitive solutions to the non-random sample selection exist with the current CPUE standardization methods. For most SEDAR-assessed species no clear solution to these problems exists because so the most prudent solution for analysts is to understand the limitations of fishery-dependent CPUE data and to be cautious against the pitfalls of spatial changes in the fish or the fishery that may bias CPUE indices. Rather the workshop participants provide a series of

28

recommendations that can best ward against improperly accounting for spatial variation in the fish or the fishery.

Recommendations

1. Tables of sample sizes/percent positive by year and area
These provide a simple tabular representation of changes in fishing catch, effort and probability of success. They are particularly useful for observing imbalanced designs that may produce spurious space*time interactions or which may require decisions about backfilling missing observations, removing spatial cells from analysis or splitting up spatial regions. These tables can be simply constructed in R using either the *table()* or *tapply()* functions or with SAS with *proc freq*.

Table 1. Recorded sample sizes in each year for US Pelagic longline logbook data.

	Caribbean	Florida East	Gulf of Mexico	Mid-Atl. Bight	NE Coast	NE Distant	Sargasso Sea	South Atl. Bight
1986	76	379	448	239	209	51	2	93
1987	1605	3141	3003	1429	946	760	41	414
1988	2081	3340	2215	1282	957	1464	27	820
1989	1519	4142	2703	2069	1026	1577	186	970
1990	1391	3299	2672	2147	1617	1072	210	1386
1991	1043	2922	3145	2302	1794	1062	232	1152
1992	912	2539	3888	2456	1410	1189	385	1068
1993	1128	1962	3345	2476	1217	1067	606	1630
1994	1414	2108	3333	2821	1027	1007	642	1788
1995	1336	1963	3689	3061	1270	927	1191	1544
1996	1219	1830	4929	1429	1320	688	690	2729
1997	878	2254	5005	1719	1464	728	378	1706
1998	559	1785	3799	1741	1058	619	335	1411
1999	294	2097	4521	1814	757	430	194	1486
2000	420	1982	4432	1545	752	602	118	1325
2001	314	991	4459	1506	1003	332	172	1301
2002	288	934	4461	1283	681	493	206	834
2003	242	898	4425	903	556	534	297	948
2004	352	598	4542	958	547	456	156	1054
2005	201	569	3250	956	421	463	189	722
2006	84	578	2685	1088	456	383	209	760
2007	41	719	3078	1175	361	348	148	983
2008	51	156	285	17	NA	NA	44	123

2. Maps of catch observations for each year

Maps of individual catch/effort observations, to the finest spatial scale possible should be produced for every index for each year to explore for the potential that fishing catch and effort may show shifts that might be indicative in serial depletions, different abundance trends in different regions in different years or changes in the distribution of the species.
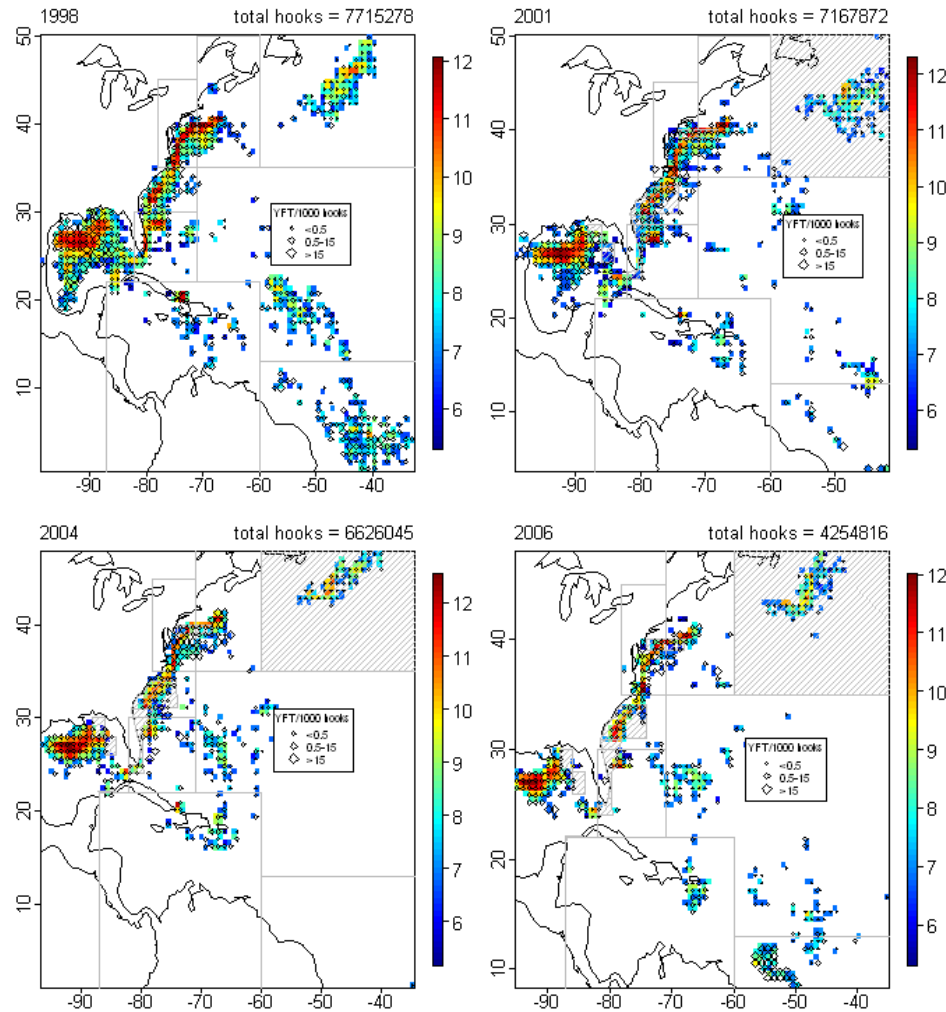
Figure 1. Map of catch rates and effort in US pelagic longline logbook. Shaded colors are effort in log(# of hooks). Diamonds represent catch of yellowfin tuna per 1000 hooks.

3. Plot the data

Plots of the nominal data, by factor level combinations, with standardized values, if possible, should be produced. These plots are critical to exploring and explaining space-time interactions as well as fishery dynamics and should be produced for each index.
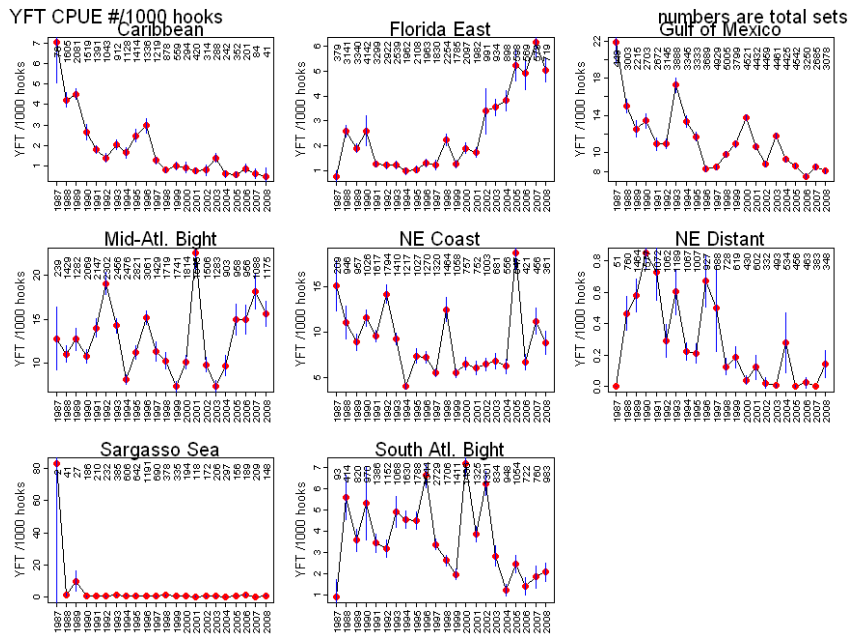
Figure 2. CPUE of yellowfin tuna from US pelagic longline logbook program.

4. Evaluate how representative the index is of population of interest

This evaluation is likely to be semi-quantitative and can be something as simple as summing up the number of spatial cells that the index covers in each year.
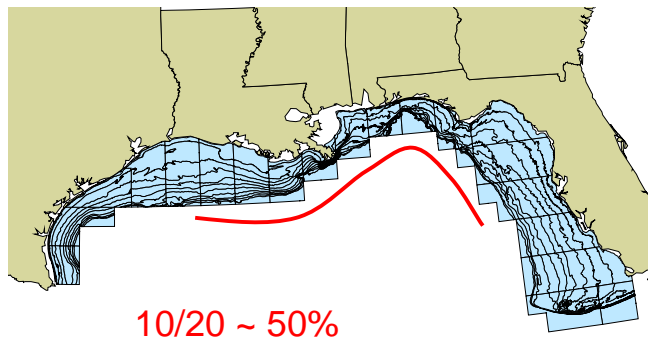
Figure 3. Map of shrimp grids showing sample coverage calculation for hypothetical survey covering 50% of the grid cells.

5. Test for and model space*time interaction in model standardization.

Significant fixed space*time interactions indicate that the resulting year effect differs in different areas and is no longer unique. Figure 2 illustrates a potential space*time interaction where the trend in CPUE differs between the Florida East Coast and the Gulf of Mexico. In a situation such as this the main effect of year is not the same in all locations. In general, space*time interactions

are modeled as random factors, usually with SAS Proc Mixed or the SAS glimmix macro. Similar mixed effect packages exist in R. Maunder et al (2008, *pers comm.*) has detailed four approaches to dealing with these interactions which the workshop participants also recommend:

a.   Explicitly ignore the interaction

Significant space*time random interactions can often occur with unbalanced sampling designs and are often observed with very large datasets. In certain cases it may be possible to argue that these interactions can be ignored, however this decision should be explicit and the interactions tested rather then implicitly ignoring them by not testing the interactions.

b. Average the year effects for each area

Construct a weighted average of the *year*area* effects. Assuming that each area provides some information on the overall trend, and that some weighting method (area, catch, etc) can be reasonably applied, then the *year*area* effects can be weighted to obtain an overall index (Punt et al. 2000).

c. Treat the interaction as a random effect

Include *year*area* interactions as random effects in the standardization model (Cooke 2007). This operates upon the assumption that the interaction is a product of random changes in the distribution of the fish. This approach is likely to be safer for pelagic resources rather then for benthic or structure-oriented fish. The result of including the *year*area* as a random effect is generally wider confidence intervals, however it can have an effect upon the overall CPUE trend. Interpretation of the main effect in the presence of random year*area effects remains problematic.

d.   Model separate populations

If significant interactions do occur and different regions clearly have different trends that are likely a result of relatively independent processes, then indices can be constructed separately. If the processes are completely independent, then separate models can be constructed, however if linkages exist between regions it will be necessary to parameterize the exchange between regions.

Concluding remarks

This document is meant to serve primarily as a set of guidelines for the development and analysis of CPUE indices. Most recommendations are also formally included in the CPUE report card.

References

Cadrin, S. X., K. D. Friedland, J. R. Waldman. 2004. Stock Identification Methods: Applications in Fishery Science. Academic Press, 2004. 736 pages

Cochran, W.G. 1963. Sampling techniques. John Wiley and Sons, New York. Cooke, J.G.

Cooke, J.G. 1997. A procedure for using catch-effort indices in bluefin tuna assessments. Col.Vol.Sci.Pap. ICCAT: 46 (2): 228-232.

DeLury, D. B. 1951. On the planning of experiments for the estimation of fish populations. Journal of the Fisheries Research Board of Canada, 8: 281–307.

Maunder, M. 2008. pers comm. http://www.fisheriesstockassessment.com/TikiWiki/tiki-index.php?page=Spatial+Analysis+for+Stock+Assessment+2008.

Maunder, M. N., and Punt, A. E. 2004. Standardizing catch and effort data: a review of recent approaches. Fisheries Research, 70: 141-159.

NMFS. 2001. Marine Fisheries Stock Assessment Improvement Plan. NOAA Technical Memorandum NMFS-F/SPO-56. Available online at: http://www.st.nmfs.noaa.gov/StockAssessment/index.html.

Punt, A.E., Pribac, F., Walker, T.I., Taylor, B.L. and J.D. Prince. 2000. Stock assessment of school shark Galeorhinus galeus based on a spatially-explicit population dynamics model. Mar. Freshw. Res. 51: 205-220.

Walters, C. J. 2003. Folly and fantasy in the analysis of spatial catch rate data. Canadian Journal of Fisheries and Aquatic Sciences, 60: 1433-1436.

Sibert, John, K. Holland, and D. Itano. 2000. Exchange rates of yellowfin and bigeye tunas and fishery interaction between Cross seamount and near-shore FADs in Hawaii. Aquatic Living Res., 13 (4): 225-232.

Venables W. N. and C. M. Dichmont. 2004. GLMs, GAMs and GLMMs: an overview of theory for applications in fisheries research. Fisheries Research 70: 319–337.

4.2. Methods for combining multiple indices into one – Paul Conn

P. Conn presented methods for combining multiple indices via Bayesian hierarchical analysis, with the goal of estimating a single time series of relative abundance.
In order to do so, he assumed that each index was attempting to measure the same quantity (relative abundance), but was subject to process and sampling errors. Although sampling errors can typically be approximated (as outputs of delta-GLM analyses or from knowledge of the sampling design), it is often the case that indices exhibit additional variation that cannot be explained by sampling errors alone. With one index alone, it is impossible to infer the appropriate level of process error. However, with multiple indices, there is some level of replication, and process errors become estimable – at least when conditioned on a particular error structure.

The particular approach that Conn advocated at the workshop involved gradient matching (i.e., modeling relative changes in abundance rather than absolute values). In this manner, one can ignore the possible differences in scale of each of the indices, an advantage that becomes apparent when trying to combine indices that don't all overlap at a single time point. These methods are presented in a SEDAR working paper (SEDAR17-AW06, available online at http://www.sefsc.noaa.gov/sedar) so are not reproduced here.

The presentation and subsequent discussion emphasized that the method appears promising, in that one time series can be generated from multiple, conflicting indices. On the other hand, consideration should be given to making sure selectivities are similar for the indices being combined, and that each index is truly reflective of abundance. Workshop participants also felt that extensive simulation testing was necessary to ensure that the method was producing reasonable results in typical fisheries scenarios.

P. Conn suggests that any prospective users of this approach contact him first. In particular, he has recently implemented some improvements to methods that do not appear in the SEDAR working paper. These improvements, together with simulation studies and an application to Spanish mackerel, are included in a manuscript that has been submitted for publication.

4.3. Notes on the Use of the Multispecies Regression in Subsetting Effort Data – Andi Stephens

Many statistical methods can be straightforward to use, but difficult to interpret. I present here some rules of thumb for using a multispecies logistic regression to predict habitat (Stephens and MacCall, 2004) with the goal of subsetting effort data.

It is very important to be able to relate the results of the method to a biological situation. First, the physical habitat must be one that lends itself to this type of analysis. The system should be one in which different species adhere to different habitats, and the target species in particular must exhibit a fair amount of habitat fidelity. Analyses should be confined to periods in which these conditions hold true.

Second, when we understand the response of regression metrics to physical parameters, then even if the regression fails it can provide information about the system that might otherwise be difficult to observe. I have highlighted here several effective performance measures, and described their behavior in the light of different plausible ecological conditions. These are the results of simulation analyses (not presented here).

My simulation study provided insight into the limitations of the method: it fails to perform well when data are too limited, when either the target species or the regressor species does not practice site fidelity, or when the regressors are predominantly negative or predominantly positive predictors of the target species. On the positive side, the regression is relatively robust to changes in regressor populations.

Note that the multispecies regression is not intended to inflate catch data.

Rules-of-thumb for regression diagnostics
The results of simulation experiments lend themselves to developing rules-of-thumb for interpreting regression diagnostics. First, note that one well-known rule of thumb, the metric that suggests regression deviance should be less than Pearson's χ^2 test for the significance of a proportion, is an indicator only that the model used is a good model for the data, but it doesn't correlate well with model performance.

Data:
> • The data suited to this type of analysis must be presence/absence data, and should present more than a few regressor species. Species with positive and negative co-occurrence with the target must be represented.
> • A classic problem in regressions of this type is colinearity among the regressors. In a logistic regression, this occurs when two regressors have the same pattern of occurrence. Data should be checked for colinearity, however many software packages such as R will report colinearity in the data.

• Related to colinearity is the problem that may occur if there are few species that co-occur with the target species, or if there are few patterns of co occurrence. In this case, the model may overfit the data, and reducing the number of parameters used may improve model performance.

Probabilities:

• The highest probability predicted by the model is a measure of how much faith we can put in the model's positive predictions. The lowest probability predicted is a measure of the certainty we have in predictions of absence of the target.

• The number of indeterminate probabilities is a measure of how well the model is able to distinguish the target's habitat.

• A lack of intermediate probabilities is a signal that either the regression is optimally segregating positive and negative predictions, or that the data may be too sparse, leading to overfitting the model. Other indicators of this type of failure are low numbers of significant regressors, and extreme values of regression coefficients.

Regressors:

• As the number of significant regressors increases, the accuracy of predictions increases, however the regression may still perform well with few regressors.

• A preponderance of either positive or negative regressors will tend to bias the model towards over- or under-predicting the target.

• The range of regression coefficients should be greater than one, but not extremely large. For this data, poor regressions had coefficient ranges either less than one, or greater than a hundred.

• Species that exhibit low habitat fidelity are poor regressors. Coefficients for these species will be close to zero.

Predictions:

• A regression that predicts all-negatives or all-positives is failing to predict the less-common case – the occurrence of a rare species or absence of an abundant one.

• The numbers of false positive and false negative predictions should be roughly the same. A predominance of either one indicates bias, and suggests that the predictive species are largely ones that co-occur with the target, or species that are seldom found with it. Removal of some regressors may balance the predictions.

Target characteristics:

• Species that exhibit low habitat fidelity or are migratory under different environmental regimes are poor targets.

Finally, when the data are collected over a period of time long enough for habitat associations to change, or when habitat use is not well understood, it is important to examine regression performance in temporal subsets of the data, especially with respect to change in sign or magnitude of the regression coefficients.

Conclusions

The species-based regression is quite robust except when the main assumption is violated: that species remain true to habitat. Thus the regression fails when habitats are indistinct, when either the target species or the regressor species use all habitats equally, and when the species change habitat use patterns, which is typical of many marine species under changing environmental regimes.

Regression deviance $< \chi^2$ indicates only that the model used is a good model for the data, and doesn't correlate well with model performance. While this is useful for catching extreme misfits, model performance is best judged using the criteria listed in the previous section, based on the probabilities and coefficients generated by the regression.

Regression coefficients should reflect biological facts in the fishery: species found rarely with the target species should have negative coefficients, and species often found with it should have positive ones. Ideally, one would prefer a variety of significant regressors, both positive and negative. Coefficients falling in the range[1..25] seem to be associated with good regressions, those much larger or smaller signify problems.

Probabilities should not be narrowly distributed, nor should they all be close to 0.5. Although it is intuitively satisfying to desire a probability distribution that perfectly distinguishes presence and absence of the target, this may indicate that the model is over fitting the data. False predictions should be examined for hints of bias in the regression. An unbiased regression will have approximately equal numbers of false positives and false negatives.

Although the basic assumptions underlying the use of CPUE as an abundance index are often questioned (e.g., Peterman and Steer, 1981; Swain and Sinclair, 1994; Harley, Myers, and Dunn, 2005), it continues to be widely used. Other methods of subsetting exist, but have their own shortcomings, including dependence on assumptions that may not hold.

The most commonly used method of subsetting data historically has been to include only those catches that contain the target; however this misses the empty catches in habitat, and it is changes in the proportion of these that reflect abundance changes. A common work-around for this has been to assume that the catch represents a constant proportion of the population, which doesn't take changes in fishing trends into account. However, fishing practices do change in ways that impact individual species (see, for example, Stephens and MacCall, 2004). Another historically employed method for subsetting recreational data relies on what the captain or fisher reports as the "target" species, but often this is the species most-frequently caught, designated in retrospect as the target.

In subsetting commercial data, some have suggested using only records from fishers whose data implies a reliable targeting of the species. However, the criteria for defining what this means is evolved ad-hoc (Taylor, 2003; Punt *et al.*, 2001b). Commercial data are often subset based on

vessel-type (Jimenez, *et al.*, 2004); Bishop, et al offer an interesting comparison of several of these in subsetting data in the Australian Northern Prawn fishing industry, using vessel and gear types (which are assumed constant in the recreational fishery) (Bishop, *et al.*, 2004).

Other GLM-based methods have involved the use of zero-inflated distributions, such as the zero-inflated Poisson (ZIP) that rely on a mixed probability distribution that accounts for zero-catches based on some presumption of a systematic reason for zeros, one that can be modeled with a known probability distribution (Maunder and Punt, 2004). Related to this is the concept of the delta-GLM, modeling the probability of obtaining a zero catch binomially, and the catch-rate for positive catches (of numbers or biomass of fish) separately, often using a log-normal distribution (Stefansson, 1996). These are sometimes called "hurdle models".

There are computationally intensive approaches such as Bayesian hierarchical analysis (another mixed-probability technique) (c.f. Argaez, *et al.*, 2005). These methods may not be well-suited to recreational datasets, either because of underlying assumptions or because they may be difficult to implement properly. For these reasons, the logistic regression will be the analytical tool of choice for many fisheries researchers.
Finally, it is important to note that the multispecies regression is not intended to inflate catch data. Using the regression to select relevant data and then also including the deselected positive catch is an inappropriate use of the method and provides biased results.

Literature
Argaez, J.A., Christen, J.A., Nakamura, M., Soberon, J., 2005. Prediction of potential areas of species distributions based on presence-only data. Environ. and Ecol. Statist., 12. 27-44.

Bishop, J., Venables, W.N., and Wang, Y.-G., 2004. Analysing commercial catch and effort data from a Penaeid trawl fishery: A comparison of linear models, mixed models, and generalised estimating equations approaches. Fish. Res. 70:179-193.

Harley S.J., Myers, R.A., and Dunn. A. 2001. Is catch-per-unit-effort proportional to abundance? Can. J. Fish. Aquat. Sci. 58: 1760–1772.

Maunder, M.N., Punt, A.E., 2004. Standardizing catch and effort data: a review of recent approaches. Fish. Res. 70. 141-159.

Peterman, R.M., and Steer, G.J. 1981. Relation between sportfishing catchability coefficients and salmon abundance. Trans. Am. Fish. Soc. 110: 585–593.

Punt, A.E., Pribac, R., Walker, T.I., Taylor, B.L., 2001. Population modeling and harvest strategy evaluation for school and gummy shark. Report of FRDC Project No. 99/102. Cited in: Maunder and Punt, 2004.

Stefansson, G. 1996. Analysis of groundfish survey abundance data: combining the GLM and delta approaches. ICES J. Mar. Sci. 53, 577-588.

Stephens, A., and MacCall, A. 2004. A Multispecies Method for Subsetting Logbook Data for Purposes of Estimating CPUE. J. Fish. Research 70, 299-310.

Stephenson, R.L., and Lane, D.E., 1995. Fisheries management science: a plea for conceptual change. Can.J. Fish. Aquat. Sci. 52: 2051-2056.

Swain, D.P., and Sinclair, A.F. 1994. Fish distribution and catchability: what is the appropriate measure of distribution? Can. J. Fish. Aquat. Sci. 51: 1046–1054.

Taylor, P.R., 2003. Standardized CPUE for the northwest Chatham Rise orange roughy fishery. NZ Fisheries Association Report No. 2003/32. Cited in: Maunder and Punt, 2004.

4.4. CPUE Model Diagnostics – Mauricio Ortiz

The common practice of model fitting involves the selection of a model class appropriate to the data, fitting and results presentation.   Comparisons between alternative models are normally related to "goodness" of fit to the data, and the selected model is commonly 'the best fit' according to some specific criteria, such as log likelihood, AIC or similar statistics. However, seldom there is a report of the model diagnostics and evaluation of "appropriateness" of the model implicit assumptions to the data structure.   Selection of the model class is usually the first step and decisions should take into account the nature of the data and objective(s) of the analysis mainly.   But, even after a "proper" selection of the class model, the data themselves may show inconsistencies in the fitting process. These inconsistencies can arise from: a) the data as whole show systematic departures of the model assumptions, and or b) that a subset of observations shows discrepant from the rest of the data.

Therefore model checking should be an integral part of the model fitting.   Based on diagnostics the modeler will introduce a verification step in the process assuring that the selected model is the appropriate one given the data analyzed. McCullagh and Nelder (1989) classified the model checking techniques into two groups: a) informal, those that relay on subjective human decision to determine patterns or better, departure from expected patterns. And b) formal techniques that imply a wider model (where the "selected current model" is a subset) with higher number of parameters. In this case, the current model passes the check if it can demonstrate that the extra parameter(s) in the wider model did not improve the fitting of the data.

Model checking elements

Model checking was originally developed for classical linear models, then McCullagh and Nelder (1989) extended it to Generalized Linear Models (GLMs).   Analysis of residuals are the primarily element for model checking, however other components in model checking include the fitted values, the linear predictors, the residual variance, the dispersion parameter and the elements of the projection ('hat') matrix.

The classical residual definition; $r = y - \mu$ is often called "response residual" in GLMs (R, S-Plus, SAS) and is simply the difference between the observed value and the corresponding predicted by the model.   Depending upon the GLM error model distribution, residuals are often transformed such its variance is constant and have a "normal like" behavior for model checking purposes, these residuals are called "standardized residuals" and include: i) the "Pearson residuals", where response residuals are scaled by the estimated standard deviation of the response variable $r_\mathrm{p} = \left. y - \mu \middle/ \sqrt{V(\mu)} \right.$. ii) The 'studentized' standardized residuals, where

response residuals are scaled by the corresponding element of the "hat" matrix .

$$r_\mathrm{s} = \left. y_i - \mu_i \middle/ \sqrt{1 - \boldsymbol{h_i}} \right.$$

 iii) And, the Anscombe residual, where the residual is replaced by a function A(y) that makes A(Y) "as normal as possible", where A is a transformation function specific for each type of error distribution. For example for the Poisson distribution the Anscombe residual is estimated as

$$r_A = \left. \frac{\frac{3}{2(\ )y^{\frac{2}{3}}} - \mu^{\frac{2}{3}}}{\mu^{\frac{1}{6}}} \right.$$

Other types of residuals are available from GLMs, including: iv) Deviance residuals, where each residual reflects the contribution to the total Deviance of the GLM fit, $r_D = sign(y - \mu)\sqrt{d_i}$, such

$$\sum r_D^2 = D$$

. Of course, for each error distribution the Deviance formulation varies (pp 34 McCullagh and Nelder 1989). In the case of the Poisson distribution, the deviance residuals are $r_D =$ $sign(y - \mu) \left\{ 2 \left(y \log\left(\frac{y}{\mu}\right) - y + \mu \right) \right\}^{\frac{1}{2}}$. Both the Pearson and Deviance residuals can be "studentized" scaling them by the variance and/or dispersion parameter, and the corresponding "hat" element, such

$$r_p' = \frac{y - \hat{\mu}}{\sqrt{\hat{\phi} V(\hat{\mu})(1 - h)}}$$

for the standardized Pearson residual, and

$$r_D' = \frac{r_D}{\sqrt{\hat{\phi}(1 - h)}}$$

for the standardized deviance residual.   v) Working residuals, these residuals correspond to the "linear" last step of the iterative fit solution of the GLM, where the y values are in the scale of the link function and the μ represent the linear predictors of the GLM model component. According to McCullagh and Nelder (1989) in general deviance residuals are preferred for model checking procedures because of its distributional properties are more similar to the linear regression models.

Systematic departures from model assumptions
        The following diagnostics try identifying inconsistencies between the 'model selected' and the data structure as a whole. We want to confirm the assumptions regarding a) the error or variance function, b) the link function selection, and c) possible omitted linear factors.
Residual Plots
    Residual plots generally include plots of standardized deviance residuals against the linear predictors ($\hat{\eta}$)of the model, or if not available, the fitted values transformed to the constant information scale of the error distribution assumed.   In the later case, the transform values are estimated as
- $\hat{\mu}$ for Normal error distribution (Gaussian)
- $\sqrt[2]{\hat{\mu}}$ for Poisson error distribution
- $2 \sin^{-1} \sqrt{\hat{\mu}}$ for Binomial error distribution
- $2 \log \hat{\mu}$ for gamma error distribution
- $-2\hat{\mu}^{-\frac{1}{2}}$ for inverse Gaussian error distribution.

41

The expected pattern of this plot is a distribution of residuals with constant range and with mean of zero. If there is large number of observations, it is recommended to use smoother functions to facilitate the interpretation of patterns. Departures of the expected pattern, for example curvature of the mean $\hat{\mu}$, may indicate incorrect choice of the link function, wrong scale of one or more covariates (linear components in model), or omission of higher order terms in the model. A non-constant range of residual indicates an incorrect variance function assumption in the model. This plot is not informative for binary data (0, 1).

A plot for checking the variance function assumption is the absolute residuals against fitted values in a constant-information scale. The expected pattern is no trend with constant distribution on the horizontal scale. A positive trend indicates that the variance function increases too slowly compare to the mean values of the data, in contrast a negative trend indicates that the variance function increases much faster that the mean values of the data.

A plot for checking the link function assumption is the adjusted dependent variable (y) against the linear predictor of the GLM model ($\hat{\eta}$). The expected pattern in a straight line, curvatures in the plot indicate a low or high power in the exponential link assumption. This plot is not informative for binary data.

A plot for checking the scale of factors/covariates is the standardized residuals against an explanatory variable(s) of the linear predictor formula. The expected pattern is also a constant range distribution with a mean distribution about zero. Departures indicate possible missing interactions or quadratic terms in some factor(s), incorrect scale for the explanatory variable or improper link function. An alternative plot is also the partial residual plot for each factor/covariate. Where the partial residual is estimated as $u = y - \hat{\eta} + \hat{\gamma}x$ where y is the adjusted dependent variable, $\hat{\eta}$ the linear predictor, and $\hat{\gamma}$ is the parameter estimate for the explanatory variable x. The expected pattern is a linear trend, departures may indicate the wrong scale of the factor, however it may also be due to wrong scale of another correlated factor, in this case multiple residual plots may help to identify common issues.

Isolated departures from the model

The prior section addressed mainly overall data inconsistencies with the selected model assumptions, once a final model can be confirm the next step is to evaluate particular points in the data especially those that appear in difference with the remain bulk of the observations. In this regard, residual analysis focus on the influence and/or "leverage" of a given observation, or their inconsistency with the general trend and the effect(s) on the parameters estimated by the model.

Leverage

For linear models the measure of leverage is given by the diagonal elements of the 'hat' matrix that project the data onto the fitted values. The leverages h_i represent the influence of a given point in the fit. In classical linear models they are a function of X only, but in GLM, they also depend on the weights assigned by the Iterative Re-weigthing Least Square (IRWLS) algorithms for solving the GLM models. Large value of h_i indicates that the fit may be sensitive to the response observation i. A plot of leverage values indicating those values of

$$h_i > \frac{2p}{n}$$

, or approximately

$$h_i' > 2, where\ h_i' = nh_i\frac{}{p}$$

is an informative tool.

Influence

Influence is normally measure as weighted combination of the changes of estimates with and without a given extreme point. Cooke (1977) introduce an statistic, 'Cooks distances' that approximates the residual scaled difference between the model fit with observation response for case *i* and the model fit without the observation *i*. The approximation by Williams (1987) is

$$(y - \hat{\mu})\sqrt{(1 - h_i)}r_{Ds}^2 + h_i r_{Ps}^2$$

where r_{Ds} is the studentized deviance residual, and r_{Ps} is the studentized Pearson residual.

A plot of Cook's distances will show the most influential observation(s) in the model fit. Q-Q plots of residuals are standard plots for checking linear models or GLM with normal error assumptions; however for other error distributions it should not expect a normal distributed residual pattern, particularly of the response residuals. For outlier detection or influential observations, half-normal plots are also informative. This type of plot will identify in one tail of the distribution those observations with large influence or leverage. Half normal plots of the leverages, Cook's distances and studentized residuals provide an informal check for observations diagnostics.

4.5. Review of Index Weighting Schemes for Stock Assessments – Clay Porch

Stock assessment algorithms typically incorporate data on stock structure and mortality from a variety of sources, including landings, length or age composition, and indices of relative abundance derived from fishery catches or fishery-independent surveys. Usually some data sets are less informative (more uncertain) than others and an important part of the stock assessment process is to determine how to take these differences into account. For example, one may wish to place greater emphasis on a large sample from the entire study population than on a small sample from a fraction of the study population. The challenge, of course, lies in objectively determining the degree to which each data source should be emphasized (weighted). At present, there is no universally accepted method for meeting this challenge, with the predictable result that the choice of weighting schemes often depends as much on who is participating in the assessment as it does on the types of data involved. As this has long been a contentious issue in many forums (Schnute and Hilborn 1993; ICCAT 1999), it is important for SEDAR participants to understand the strengths and weaknesses of the various weighting schemes before adopting one for a particular assessment. To that end, the chapter reviews the weighting schemes that have been developed specifically for indices of abundance and suggests criteria for selecting between those schemes during SEDAR data and assessment workshops.

Measuring uncertainty

Uncertainty can be thought of as a function of observation errors (degree to which the observed value of an index differs from its true value) and process errors (degree to which the true value of an index differs from that of the study population). Observation errors typically depend on factors such as sample size or randomness. Process errors arise, for example, when the fraction of the study population that occupies the area sampled for the index varies through time. If these errors are effectively random over the study period, then one may consider them to be unbiased representations of the study population (although they may still be highly imprecise). In that case it is advantageous to have multiple indices representing the same quantity or, when possible, to combine several uncertain indices into one (see Conn's chapter). Systematic observation or process errors, of course, can lead to bias.

Known biases of a substantive nature should obviously be corrected during the index standardization process or else the index should be discarded. The relevant question here is what to do with indices that are suspected of some unknown bias. In many assessments the available data are too sparse to afford the luxury of simply discarding suspicious indices. For such cases, we maintain that unknown biases among multiple independent data sets may be regarded (hopefully) as canceling out on average and therefore may be treated as another degree of imprecision. It is therefore appropriate simply to deemphasize such data during the fitting process in some proportion to the degree to which bias is suspected (more on this later). Given these caveats, the next order of business is to select a common metric for measuring the uncertainty of each candidate index of abundance.

Most modern stock assessment models are cast in a probabilistic framework, in which case the most natural measure of the uncertainty in an index is its variance. For example, consider two independent normally distributed indices X_1 and X_2 with the same expectation μ but different variances (σ^2 and σ^2). The probability of observing two given values x_1 and x_2 is

(1) $\qquad p(x_1, x_2) = \dfrac{1}{2\pi\sigma_1\sigma_2} \exp\left[-\dfrac{(x_1-\mu)^2}{2\sigma_1^2} - \dfrac{(x_2-\mu)^2}{2\sigma_2^2} \right]$

The maximum likelihood estimator for µ is the value that maximizes equation 1, or equivalently, minimizes its negative logarithm:

(2) $\qquad -\ln p(x_1, x_2) = \ln(2\pi\sigma_1\sigma_2) + \dfrac{(x_1-\mu)^2}{2\sigma_1^2} + \dfrac{(x_2-\mu)^2}{2\sigma_2^2}$

The solution when σ_1 and σ_2 are known is easily obtained by simple calculus:

(3) $\qquad \hat{\mu} = \left(\dfrac{x_1}{\sigma_1^2} + \dfrac{x_2}{\sigma_2^2} \right) \Big/ \left(\dfrac{1}{\sigma_1^2} + \dfrac{1}{\sigma_2^2} \right)$

As can be seen, the influence of each index series on the final estimate of µ is directly proportional to the variance associated with that index.

Some authors have adopted somewhat simpler notation where weight coefficients (w) are specified directly in the likelihood function, e.g.,

(4) $\qquad -\ln p(x_1, x_2) \propto w_1(x_1-\mu)^2 + w_2(x_2-\mu)^2$

In this example it is easy to see that the w's are merely the reciprocals of the variances and that minimizing (4) or (2) will yield identical results as long as the variances (and w's) are constant. In keeping with most statistical texts, we find the variance notation to be more intuitive, particularly when different error structures are mixed such that the relative values of the w coefficients are no longer comparable. Moreover, the weighting coefficient formulation is impractical when the values of the weights are to be estimated internally in the model.

Review of index weighting schemes

The weights applied to the indices of abundance used can have a profound effect on the outcome of an assessment. To illustrate this, consider the following simple model:

(5) $\qquad \begin{aligned} X_{1,t} &= q_1(N+Gt) + \varepsilon_{1,t} \\ X_{2,t} &= q_2(N+Gt) + \varepsilon_{2,t} \end{aligned}$

where $x_{1,t}$ and $x_{2,t}$ are independent indices of an underlying population that is growing linearly with time t from a known starting abundance N). If the errors ε_1 and ε_2 are approximately normal-distributed with equal variances, then maximum likelihood estimates for q_1, q_2, and G may be obtained by minimizing

45

$$(6) \qquad -\ln p(\vec{x}_1, \vec{x}_2) = \sum_t \ln(2\pi\sigma^2) + \frac{(x_{1,t} - q_1(N + Gt))^2}{2\sigma^2} + \frac{(x_{2,t} - q_2(N + Gt))^2}{2\sigma^2}$$

Otherwise, the appropriate likelihood function for unequal variances is

$$(7) \qquad -\ln p(\vec{x}_1, \vec{x}_2) = \sum_t \ln(2\pi\sigma_1\sigma_2) + \frac{(x_{1,t} - q_1(N + Gt))^2}{2\sigma_1^2} + \frac{(x_{2,t} - q_2(N + Gt))^2}{2\sigma_2^2}$$

In both (6) and (7) the over-arrow denotes the vector of time series observations for each index. It is easy to show that the estimates for q_1, q_2, and G obtained from the equal variance formulation (6) will be the same whether σ^2 is estimated or fixed to an arbitrary constant (which is the least squares solution). It is also generally true that the equal-variance (6) and bi-variance (7) estimators will perform similarly when applied to situations where the true variances of each index are the same (and the underlying model is correct). However, if the true variances are substantially different, the assumption of equal variances (application of 6) can lead to biased estimates.

Figure 1a and Table 1 compare the results when the equal-variance (6) and bi-variance (7) estimators were applied to simulated indices generated according to (5) when the true values of σ^2 and σ_2^2 were 64 and 13, respectively. The bi-variance estimator performed best, producing slightly better parameter estimates and a much lower value for the likelihood function. This would seem to confirm that there is some advantage to estimating separate variances for each index rather than assuming they are the same, however the experiment was predicated on having the correct underlying model. Consider now the same set of experiments when the indices of abundance were equally biased, but in opposite directions such that the annual rate of increase for index 1 is 0.2 and the annual rate for index 2 is 1.8 (compared to the true rate of 1.0). In that case the bi-variance estimator again produced a significantly lower likelihood function value than the equal-variance estimator (82.6 versus 96.6), and on that basis an investigator might be tempted to prefer it. However, closer inspection reveals that the bi-variance estimator produced this low likelihood value by severely down-weighting the less precise index 1 and focusing on the more precise index 2 (Figure 1b). As a result, the MLEs for q_2 and σ_2^2 were very close to the true values and the predicted values for index 2 matched the observations very well, whereas the MLE for σ^2 was double the true value and the predicted values for index 1 did not match the observations at all. Hence, the estimate of the primary parameter of interest (G) was 1.8; matching the true slope of index 2, but much higher than the underlying population. In contrast, the equal variance estimator gave equal credence to the two biased indicators and essentially averaged them, in this case resulting in a much less biased estimate of G.

Table 1. Parameter estimates from the equal-variance and bi-variance estimators applied to data generated from model 5.

Parameter	TRUE	Unbiased indices		Conflicting indices	
		Equal-var.	Bi-variance	Equal-var.	Bi-variance
G	1	0.98	1.01	1.11	1.81
q_1	2	2.04	1.99	0.57	0.37
q_2	1	1.02	0.99	1.54	1
σ_1^2	64	46	64	76	111
σ_2^2	13	46	13	76	13

The above example above was contrived to make the points that (1) weighting matters and (2) the estimation of variances is inextricably linked to the assumption that the model underlying the estimator is correct. While one variance scheme may appear better on the basis of the values of the likelihood function or other metric, it may in fact exacerbate potential biases by favoring solutions which disregard indices that are difficult to reconcile with the model but nevertheless accurate. For this, and other reasons that will become apparent, there is no widely accepted protocol for choosing a weighting scheme. Therefore it behooves the analyst to be familiar with both the indices being used and with the strengths and weaknesses of the various weighting options. The remaining discussion presents six basic approaches to index weighting. Most of the weighting schemes developed to date are essentially special cases of these six categories. The implementation of each scheme depends, of course, on the type of error structure assumed. The normal and lognormal error structures seem to be the most common choices in the literature; the latter in particular because it places less weight on outliers and because the variance of many indices tends to increase with the mean. Other error structures (e.g., gamma, Poisson, negative binomial) have also been considered, but the basic principles are the same. For the sake of brevity, equations are shown only for the normal and lognormal cases.

Equal Weighting.- The simplest weighting scheme is the so-called "equal weighting" option, where all observations are essentially assigned the same variance (i.e., all indices j in all years y) so that the likelihood function has the form:

$$(8) \qquad L = \begin{cases} \displaystyle\sum_{j,y} \frac{(x_{j,y} - \mu_{j,y})^2}{2\sigma^2} + \ln\sigma & \textit{normal} \\[4mm] \displaystyle\sum_{j,y} \frac{(\ln x_{j,y} - \mu_{j,y})^2}{2\sigma_{\ln x}^2} + \ln\sigma_{\ln x} & \log\textit{normal} \end{cases}$$

Note that μ is the expectation in the case of the normal distribution and the median in the case of the lognormal distribution. Also, the variance term in the normal case is on an arithmetic scale whereas in the lognormal case it is on a logarithmic scale (i.e., the variance of *ln x*). Hence, the implications of equal weighting are somewhat different between the two error structures; constant variance in the normal case and constant coefficient of variation (CV) in the lognormal case (recall $\sigma_{\ln x}^2 = ln(1+CV^2)$). It is also important to note that in the case of a normal distribution, the annual values for each index should be rescaled so that their means are similar (e.g., by dividing the original series by their respective means), otherwise series with larger absolute values will receive more weight in the fitting process than series with low values. In the lognormal case the scaling is implicit and the results will be the same whether the series is rescaled or not. In equal-weighting models where the uncertainty is limited to the indices of abundance (e.g., catch is assumed known without error), the estimates of μ will not depend on the value of the variance parameter (or whether it is estimated or not), but will depend on the choice between lognormal and normal errors.

Simple maximum likelihood weighting.- A simple extension of the equal weighting procedure is to assume the variance parameter for a given index is constant from year to year, but may differ among indices:

$$(9) \qquad L = \begin{cases} \sum_{j,y} \dfrac{(x_{j,y} - \mu_{j,y})^2}{2\sigma_j^2} + \ln \sigma_j & normal \\[4mm] \sum_{j,y} \dfrac{(\ln x_{j,y} - \mu_{j,y})^2}{2\sigma_{\ln x,j}^2} + \ln \sigma_{\ln x,j} & log\,normal \end{cases}$$

In this case the variance parameters may be fixed to some externally derived values (a special case of input variance weighting discussed below) or estimated internally in the model. Maximum likelihood estimates of the variances have been derived in three common ways. An early method was referred to as iterative re-weighting because it consisted of a two step process where the likelihood was minimized using fixed values for each variance parameter, after which the variances were recomputed using the standard MLEs.

$$(10) \qquad \begin{cases} \sigma_j^2 = \dfrac{1}{n} \sum_y (x_{j,y} - \mu_{j,y})^2 & normal \\[4mm] \sigma_{\ln x,j} = \dfrac{1}{n} \sum_y (\ln x_{j,y} - \mu_{j,y})^2 & log\,normal \end{cases}$$

The new estimates of variance would be substituted back into the likelihood and the minimization procedure reinitiated until the estimates of variance converged (essentially an application of the EM algorithm). Subsequently, it was realized that the concentrated likelihood approach, whereby the formulas in (10) are directly substituted into (9), was more efficient and yielded the same results. Alternatively, equivalent MLEs may be obtained by use of the search algorithm (minimizing 9) in the same manner as for the other estimated parameters (Porch 2002). Note that annual variances cannot be estimated using these techniques because the number of index parameters would exceed the number of index data points.

One potential pitfall of simple maximum likelihood weighting is that some index series may be severely down-weighted, as illustrated in the example. This is especially problematic if the assessment model is mis-specified and happens to match up (incorrectly) with some indices more than others. Simulations by Legault and Porch (2001) that were conditioned on data for Atlantic bluefin tuna (with various minor model mis-specifications) indicated no clear pattern in the bias or uncertainty of estimates between the equal weighting and maximum likelihood schemes.

Input-variance weighting.- This common alternative to equal weighting simply involves specifying the variance parameters based on considerations external to the stock assessment model. For example, estimates of the variance of an index may be derived for each year during the standardization process (e.g., by application of a general linear model). To the extent that they are reasonable estimates of year to year changes in index uncertainty, incorporating such externally derived variances can lead to more accurate assessments (Maunder and Starr 2003).

$$
(11) \quad L = \begin{cases} \displaystyle\sum_{j,y} \frac{(x_{j,y} - \mu_{j,y})^2}{2\sigma_{j,y}^2} + \ln \sigma_{j,y} & normal \\[3mm] \displaystyle\sum_{j,y} \frac{(\ln x_{j,y} - \mu_{j,y})^2}{2\sigma_{\ln x,j,y}^2} + \ln \sigma_{\ln x,j,y} & log\,normal \end{cases}
$$

Note that, in principle, it is possible to estimate the annual variance parameters for each index if the standardization procedure is conducted simultaneously with the stock assessment (Maunder 2001). In practice, however, the process is prohibitively time consuming.

Some scientists have advocated developing input weights based entirely on expert opinion, arguing that issues such as spatial coverage and potential for bias outweigh consideration of observation error. For example, Suzuki (2001) presented a framework where indices for bluefin tuna would be graded (on a scale of 1 to 5) on each of three factors: spatial coverage, temporal coverage, and the likely importance of changes in the operational characteristics of the index over time that are not accounted for during the index standardization process. Each index would then be weighted in accordance with the sum of those three scores (i.e., the variances would be set equal to the reciprocal of those scores).

Additional-variance weighting.- This alternative is based on the premise that the variance of an index may be decomposed into a component reflecting observation errors, such as might be obtained from the GLM standardization, and an additional component characterizing the uncertainty in the index as a reflection of the true abundance trends, as might occur when the proportion of the stock sampled by the index varies from year to year (Wade 1996, Punt et al. 1997, Geromont and Butterworth 2001). To the extent that these two error sources are independent, one might expect their variances to be additive:

$$
(12) \quad L = \begin{cases} \displaystyle\frac{1}{2}\sum_{j,y} \frac{(x_{j,y} - \mu_{j,y})^2}{(\sigma_{j,y}^2 + \chi_j^2)} + \ln(\sigma_{j,y}^2 + \chi_j^2) & normal \\[3mm] \displaystyle\frac{1}{2}\sum_{j,y} \frac{(\ln x_{j,y} - \mu_{j,y})^2}{(\sigma_{\ln x,j,y}^2 + \chi_j^2)} + \ln(\sigma_{\ln x,j,y}^2 + \chi_j^2) & log\,normal \end{cases}
$$

where χ_j^2 is the additional variance term which may vary among indices and can be estimated by maximum likelihood methods analogous to those discussed in connection with simple maximum likelihood weighting (equations 9 and 10). This method mitigates the pitfalls of equal weighting (e.g., giving equal credence to observations based on little data) and simple maximum likelihood weighting (e.g., all of the weight going to a few series), but still depends on the relative consistency of each series with the underlying model structure. The approach would be undesirable if there were strong *a priori* reasons for weighting one series more highly than another and there was uncertainty about model structure.

Multiplicative-variance weighting.- This approach is similar to the additional variance approach above except that the scaling parameter is multiplied by the year-specific input values rather than added to it:
.

$$(13) \quad L = \begin{cases} \dfrac{1}{2}\displaystyle\sum_{j,y}\dfrac{(x_{j,y}-\mu_{j,y})^2}{c_j\sigma^2_{j,y}}+\ln(c_j\sigma^2_{j,y}) & normal \\[3em] \dfrac{1}{2}\displaystyle\sum_{j,y}\dfrac{(\ln x_{j,y}-\mu_{j,y})^2}{c_j\sigma^2_{\ln x,j,y}}+\ln(c_j\sigma^2_{\ln x,j,y}) & \log normal \end{cases}$$

In many applications the year-specific variances are initially determined by GLM and then the multiplicative parameters c_j are adjusted to account for expert judgments on key factors not accounted for in the GLM (see Quinn and Deriso 1999; Sullivan et al. 1999; Parma 2000). In the most recent assessment of Atlantic and Gulf king mackerel, for example, the c_j values were adjusted such that the average coefficient of variation (CV) attributed to each index was the same, but the relative year to year changes in CV from the GLM were preserved (Cass-Calay et al., 2008).

The multiplicative-variance weighting approach is perhaps less intuitive than the additional variance approach in terms of distinguishing process and observation errors. It is useful, however, for cases where the available estimates of observation variance ($\sigma_{j,y}$) are trusted as measures of relative precision from one year to the next, but not as measures of the uncertainty across indices. It is of course also possible to estimate the c_j values analogous to the simple maximum likelihood method, subject to the same limitations (most notably that some indices may be severely down-weighted, particularly if the assessment model is mis-specified).

Variance as a function of the mean.- This last approach is based on the premise that the variance of catch per unit effort (or survey counts) tends to be correlated with the mean

.

$$(14) \quad L = \begin{cases} \dfrac{1}{2}\displaystyle\sum_{j,y}\dfrac{(x_{j,y}-\mu_{j,y})^2}{f_j(x_{j,y})}+\ln f_j(x_{j,y}) & normal \\[3em] \dfrac{1}{2}\displaystyle\sum_{j,y}\dfrac{(\ln x_{j,y}-\mu_{j,y})^2}{f_j(\ln x_{j,y})}+\ln f_j(\ln x_{j,y}) & \log normal \end{cases}$$

where f is some arbitrary function such as Taylor's power law: $f = ax_{j,y}{}^b$. In the trivial case where x follows a lognormal process with constant CV, the power function reduces to $b = 0$ and $a = ln(1+CV^2)$. More generally, the relationship (parameters a and b) could be established outside the stock assessment model by regressions of variance against the mean. However, it has been shown that such correlations tend to be spurious (see Porch 1998). Another possibility is to estimate the parameters internally, in which case one may expect similar issues to the multiplicative variance case. Regardless of whether the parameters of the mean-variance relationship are estimated internal or external to the assessment model, further bias may occur by overemphasizing observations $x_{j,y}$ that happen to be very low and it may be prudent to make the procedure more robust by adding a small constant.

Criteria for choosing between index schemes

There are no firm criteria for determining when to use one weighting scheme over another (but see McAllister et al. 2001 for a useful review). While several simulation studies have been conducted, they have all been rather limited in scope and tend to focus on comparing a few special cases of the weighting schemes discussed above. Nevertheless, there are several guidelines that can be posited.

A priori criteria

It is important for the choice of weighting schemes to be rooted in arguments that are based on the nature of the data at hand rather than how they affect the outcome of the assessment. Otherwise, the working group may be accused of succumbing to divisive arguments predicated on producing a desired stock status rather than the best available science. To this end, the working groups should discuss the following points:

1.   Are year to year variations in uncertainty likely to be substantial and measurable for a given index? If so:
   - Are they reliable only in a relative sense (as is sometimes the case with model-based estimators such as GLM standardizations)?
   - Are they also reliable in terms of absolute magnitude (e.g., a designed-based estimate from a survey that covers the entire range of the stock)?

2.   Does the level of uncertainty likely vary substantially among indices? If so:
   - Is the available expertise sufficient and able to reach consensus ranking?
   - Is the candidate index-weighting estimation method well-understood (widely-practiced, simulation-tested, and easily applied in the current assessment framework)?

A posteriori criteria

While it is important that the initial choice of a weighting method be made before running the stock assessment model, it is also important to ensure that the weighting method behaves sensibly once applied. In particular, the working group should ask the following questions once the method has been applied:

1.   Does the method accord unrealistically high or low variance to some indices?
   - What level should be deemed unrealistic?
   - Do the estimates of variance strongly disagree with expert judgment on the relative reliability of each series as an index of abundance
2.   Are estimates statistically defensible?
   - Examine goodness of fit statistics (e.g., Chi-square deviance statistic)
   - Examine model selection criteria (e.g., AICc)

A flow chart is offered in Figure 2 to help facilitate discussions relating to the above criteria. Of course in any given application there may be gray areas. Accordingly, the recommendations offered in the flow chart are intended to be viewed as guidance rather than hard and fast rules.

Acknowledgments

Literature Cited

Butterworth, D.S., Punt, A.E. and Smith, A.D.M. 1996. On plausible hypotheses and their weightings, with implications for selection between variants of the Revised Management Procedure. Rep. int. Whal. Commn 46: 481-491.

Cass-Calay, S., Ortiz, M., and Restrepo, V.R. 2008. Notes on the weighting of indices for the king mackerel VPA. SEDAR 16-AW-09 *in:* SEDAR 16 South Atlantic and Gulf of Mexico King Mackerel.

Geromont, H.F. and Butterworth, D.S. 2001. Possible extensions to the ADAPT VPA model applied to western north Atlantic bluefin tuna, addressing in particular the need to account for "additional variance". ICCAT Doc. SRCS/00/33.

ICCAT. 1999. Collective Volume of Scientific Papers. XLIX (2).

Legault, C. M. and Porch, C. E. 2001. Comparisons of index weighting schemes for tuned virtual population analyses. Col. Vol. Sci. Pap. ICCAT, 52 (5) : 1729-1738.

Maunder, M. N. 2001. A general framework for integrating the standardization of catch-per-unit-effort into stock assessment models. Can. J. Fish. Aquat. Sci. 58:795−803.

Maunder, M. N. and Starr, P. J. 2003. Fitting fisheries models to standardised CPUE abundance indices. Fisheries Research 63: 43–50

McAllister, M., Babcock, E. A., and Pikitch, E. K. 2001. Evaluating the relative merits of alternative methods to weight different time series of abundance indices in stock assessment. Col. Vol. Sci. Pap. ICCAT, 52 (3) : 1094-1115.

Parma, A.M. 2000. Bayesian approaches to the analysis of uncertainty in the stock assessment of Pacific halibut. In J. M. Berkson, L.L. Kline and D.J. Orth (eds.) Incorporating uncertainty into fisheries models. (accepted).

Porch, C. E. 1998. Some intrinsic limitations of sample variances in stock assessment models. Pages 385-398 *in* Funk, F., T.J. Quinn, J. Heifetz, J.N. Ianelli, J.E. Powers, J.F. Schweigert, P.J. Sullivan, and C.I. Zhang (ed.) Fishery Stock Assessment Models. Alaska, Sea Grant College Program Report No. AK-SG-98-011997. University of Alaska Fairbanks.

Porch, C. E. 2002. VPA-2BOX Version 3.0 User's guide. National Marine Fisheries Service, Sustainable Fisheries Division Contribution SFD-01/02-175 (also published internationally on ICCAT web page)

Punt, A.E., J.G. Cooke, D.L. Borchers, and S. Strindberg. 1997. Estimating the extent of additional variance for Southern Hemishere minke whales from the results of the IWC/IDCR cruises. Rep. int. Whal. Comm. 47:431-434.

Quinn II, T. J. and Deriso, R.B. 1999. Quantitative fish dynamics. Oxford University Press. Oxford.

Schnute, J.T. and Hilborn, R. 1993. Analysis of contradictory data sources in fish stock assessments. Can. J. Fish. Aquat. Sci. 52: 2063-2077.

Sullivan, P.J., Parma, A.M., and Clark, W.G. 1999. The Pacific halibut stock assessment of 1997. International Pacific Halibut Commission Scientific Report 79, 84 pp.

Suzuki, Z. 2001. Comparisons Qualitative evaluation of CPUE series used for West Atlantic bluefin stock assessment. Col. Vol. Sci. Pap. ICCAT, 52 (3) : 977-983.

Wade, P.R. 1996. Gray whale stock assessment. Document SC/48/AS8 presented to the June 1996 meeting of the IWC Scientific Committee meeting. 17pp.
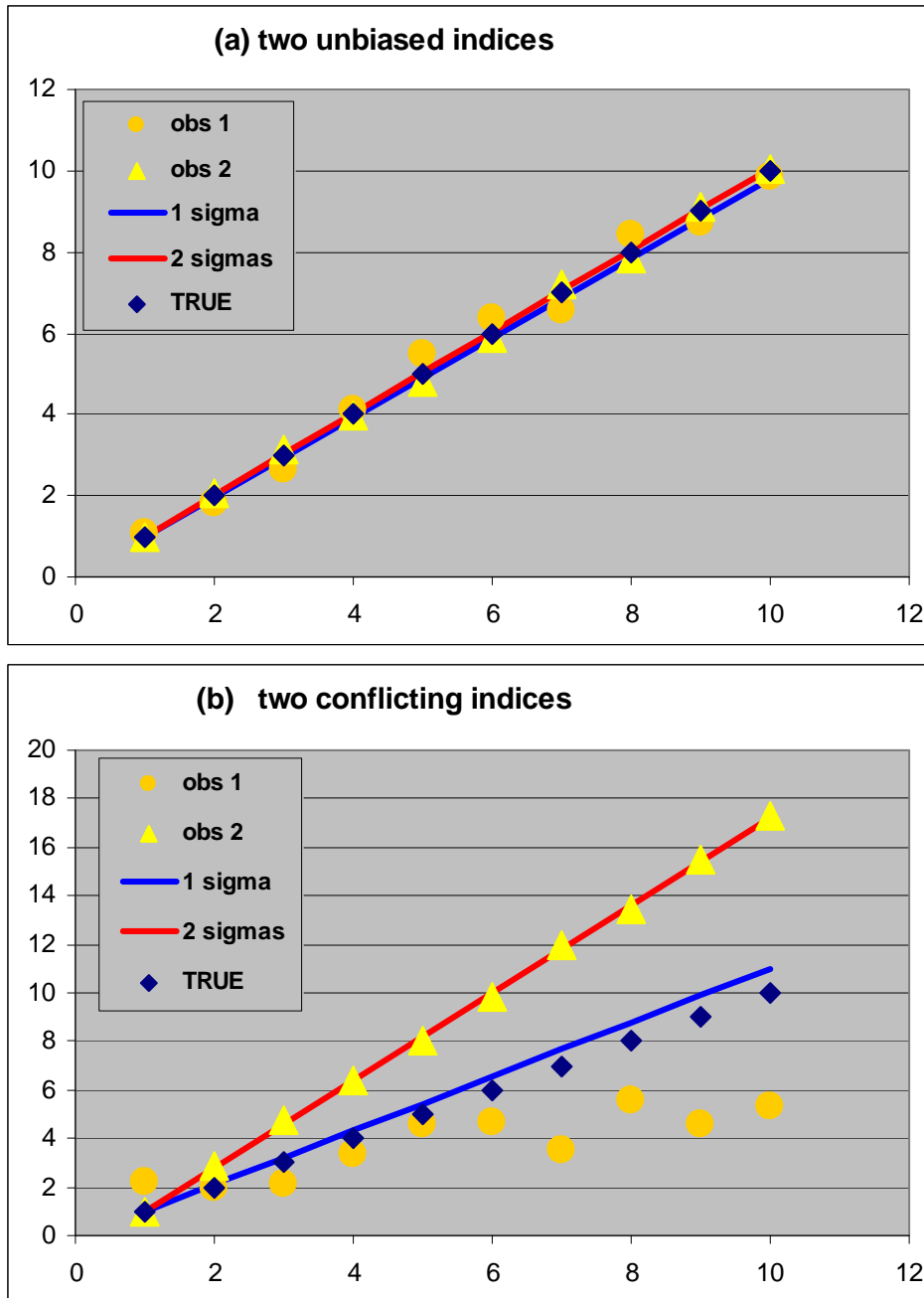
Figure 1. Predicted trends in population growth using the equal-variance (red) and bi-variance estimators (blue) compared to the true values (diamonds) and 'observed' values for index 1 (circles) and index 2 (triangles). The top and bottom panels refer to the cases when the 'observed' values are generated without bias (Fig. 1a) or generated with equal but opposite biases (Fig. 1b).

Index weighting key

Estimates of variance reliable for all years ………………….….…………………………….(11) *input variance weighting*

Estimates of variance not available or do not fully represent primary sources of uncertainty

 Overall uncertainty levels deemed similar for all indices used in the assessment model

  Uncertainty levels similar for all years………………………………..…………………………(8) *equal weighting[1]*

  Uncertainty levels differ among years

   Annual variance estimates reliably measure relative uncertainty between years…….(13) *multiplicative variance[1,2]*

   Annual variance estimates reliably measure observation errors only.…………………...(12) *additional variance[1,2]*

   Annual variance estimates unavailable or uninformative……………………………………(8) *equal weighting[1]*

 Overall uncertainty levels differ between indices used in the assessment model

  Expert consensus reached concerning relative weights to be assigned for each index

   Annual variance estimates reliably measure relative uncertainty between years....(13) *multiplicative variance[3]*

   Annual variance estimates reliably measure observation errors only.…………..……(12) *additional variance[3]*

   Annual variance estimates unavailable or uninformative……………………….(11) *input variance weighting*

  Expert consensus not reached (or not preferred) concerning relative weights to be assigned for each index

   Annual variance estimates reliably measure relative uncertainty between years....(13) *multiplicative variance[4]*

   Annual variance estimates reliably measure observation errors only.…………….…(12) *additional variance[4]*

   Annual variance estimates unavailable or uninformative

    Relationship between variance and mean is well defined[5]….……………(14) *variance as function of mean*

    Relationship between variance and mean is unclear………..…(9) *simple maximum likelihood weighting*

Figure 2. Key depicting the recommended chain of decisions with various levels of information on the uncertainty in the candidate indices of abundance. Numbers in parentheses represent the equations associated with each method in the text above. Note that the key is intended for general guidance towards developing initial weighting schemes. Assessment scientists should also consider *a posteriori* diagnostics (as discussed earlier) as well as the interactions between selected index weighting schemes and the weighting schemes applied to other types of data used in the assessment model (e.g., total catch or age composition).

---

[1] The single variance parameter representing all indices may be estimated or fixed based on expert consensus.
[2] Scaled so that the mean variance is the same for all index series.
[3] Scaled so that the mean variance of each index is consistent with expert opinion.
[4] Mean variance of each index estimated by maximum likelihood.
[5] In the trivial case of a lognormal process with constant CV, this option is equivalent to a lognormal model with simple maximum likelihood weighting.

5. Minimum Requirements for Submission of Indices of Abundance to SEDAR Workshop Meetings – SEDAR Procedures Workshop 1 Panel – Shannon Cass-Calay

This worksheet is intended to inform authors of SEDAR indices of abundance of the minimum requirements for submission recommended by the SEDAR CPUE workshop panel (Miami, FL October, 2008). The SEDAR CPUE Workshop Panel encourages the development of improved techniques for the construction of CPUE series and attempts to construct indices for data-poor species. We acknowledge that this worksheet may not be appropriate for such indices, and request that CPUE working group panels use this worksheet *and* expert judgment to determine the acceptability of each CPUE series.

The spreadsheet is intended to be used as follows:

1) No less than 2 months prior to a SEDAR data workshop, the SEDAR coordinator and the CPUE working group rapporteur will insure that all participants who plan to submit an index of abundance receive this worksheet. The CPUE working group rapporteur will communicate with all CPUE authors to ensure that they are aware of the new recommendations.

2) CPUE authors will attempt to meet the minimum requirements outlined in the worksheet. When these cannot be met, the author will submit a written justification to the CPUE working group. CPUE indices will not be rejected solely due to a failure to meet all minimum requirements.

3) The CPUE Series Worksheet is intended to be filled out by the CPUE working group at the SEDAR Data Workshop. Any missing data requirements will be noted and explained in the column "Working Group Comments". If the working group determines that an index cannot be recommended for use because the minimum requirements have not been met, the working group rapporteur will communicate with the author and determine a schedule for the submission of a revised document. Failure to submit a revised document before the negotiated deadline may result in rejection of the CPUE series.

CPUE Series Worksheets will be included in the SEDAR Data Workshop report and will therefore become a permanent, reviewable record.

DESCRIPTION OF THE DATA SOURCE

	Not Applicable	Absent	Incomplete	Complete

1. Fishery Independent Indices

A. Describe the survey design (e.g. fixed sampling sites, random stratified sampling), location, seasons/months and years of sampling.				
B. Describe sampling methodology (e.g. gear, vessel, soak time etc.)				
C. Describe any changes in sampling methodology (e.g. gear, vessel, sample design etc.)				
D. Describe the variables reported in the data set (e.g. location, time, temperature, catch, effort etc.).				
E. What species or species assemblages are targeted by this survey (e.g. red snapper, reef fish, pelagic).				
F. Describe the size/age range that the index applies to. Include supporting figures (e.g. size comp) if available.				

2. Fishery Dependent Indices

A. Describe the data source and type of fishery (e.g. commercial handline, commercial longline, recreational hook and line etc.).				
B. Describe any changes to reporting requirements, variables reported, etc.				
C. Describe the variables reported in the data set (e.g. location, time, temperature, catch, effort etc.).				
D Describe the size/age range that the index applies to. Include supporting figures (e.g. size comp) if available.				

METHODS

1. Data Reduction and Exclusions

A. Describe any data exclusions (e.g. gears, fishing modes, sampling areas etc.). Report the number of records removed and justify removal.				
B. Describe data reduction techniques (if any) used to address targeting (e.g. Stephens and MacCall, 2004; gear configuration, species assemblage etc).				
C. Discuss procedures used to identify outliers. How many were identified? Were they excluded?				

Working Group Comments:

	Not Applicable	Absent	Incomplete	Complete

2. Management Regulations (for FD Indices)

A. Provide (or cite) history of management regulations (e.g. bag limits, size limits, trip limits, closures etc.).

B. Describe the effects (if any) of management regulations on CPUE

C. Discuss methods used (if any) to minimize the effects of management measures on the CPUE series.

3. Describe Analysis Dataset (after exclusions and other treatments)

A. Provide tables and/or figures of number of observations by factors (including year, area, etc.) and interaction terms.

B. Include tables and/or figures of number of positive observations by factors and interaction terms.

C. Include tables and/or figures of the proportion positive observations by factors and interaction terms.

D. Include tables and/or figures of average (unstandardized) CPUE by factors and interaction terms.

E. Include annual maps of locations of survey sites (or fishing trips) and associated catch rates *OR* supply the raw data needed to construct these maps (Observation, Year, Latitude, Longitude (or statistical grid, area), Catch, Effort).

F. Describe the effort variable and the units. If more than one effort variable is present in the dataset, justify selection.

G. What are the units of catch (e.g. numbers or biomass, whole weight, gutted weight, kilograms, pounds).

4. Model Standardization

A. Describe model structure (e.g. delta-lognormal)

B. Describe construction of GLM components (e.g. forward selection from null etc.)

C. Describe inclusion criteria for factors and interactions terms.

D. Were YEAR*FACTOR interactions included in the model? If so, how (e.g. fixed effect, random effect)? Were random effects tested for significance using a likelihood ratio test?

E. Provide a table summarizing the construction of the GLM components.

F. Summarize model statistics of the mixed model formulation(s) (e.g. log likelihood, AIC, BIC etc.)

G. Report convergence statistics.

Working Group Comments:

MODEL DIAGNOSTICS

Comment: Other model structures are possible and acceptable. Please provide appropriate diagnostics to the CPUE indices working group.

	Not Applicable	Absent	Incomplete	Complete

Working Group Comments:

1. Binomial Component

A. Include plots of the chi-square residuals by factor.

B. Include plots of predicted and observed proportion of positive trips by year and factor (e.g. year*area)

C. Report overdispersion parameter and other fit statistics (e.g. chi-square / degrees of freedom).

2. Lognormal/Gamma Component

A. Include histogram of log(CPUE) or a histogram of the residuals of the model on CPUE. Overlay the expected distribution.

B. Include plots describing error distribution (e.g. Studentized residuals vs. linear predictor.

C. Include QQ-plot – (e.g. Student deviance residuals vs. theoretical quantiles), Overlay expected distribution.

D. Include diagnostic plot for variance function (e.g. square root of std residuals vs. fitted values). Overlay expected distribution.

E. Include diagnostic plot for link function (e.g. linear response variable vs. linear predictor). Overlay expected distribution.

F. Include plots of the residuals by factor

3. Poisson Component

A. Report overdispersion parameter and other fit statistics (e.g. chi-square / degrees of freedom).

B. Include plots describing error distribution (e.g. Studentized residuals vs. linear predictor.

C. Include QQ-plot – (e.g. Student deviance residuals vs. theoretical quantiles), Overlay expected distribution.

D. Include diagnostic plot for variance function (e.g. square root of std residuals vs. fitted values). Overlay expected distribution.

E. Include diagnostic plot for link function (e.g. linear response variable vs. linear predictor). Overlay expected distribution.

The feasibility of this diagnostic is still under review.

4. Zero-inflated model

A. Include ROC curve to quantify goodness of fit.

B. Include plots describing error distribution (e.g. Studentized residuals vs. linear predictor).

C. Include QQ-plot (e.g. Student dev. residuals vs. theoretical quantiles), Overlay expected distribution.

	Not Applicable	Absent	Incomplete	Complete

Working Group Comments:

MODEL DIAGNOSTICS (CONT.)

D. Include diagnostic plot for variance function (e.g. square root of std residuals vs. fitted values). Overlay expected distribution.

E. Include diagnostic plot for link function (e.g. linear response variable vs. linear predictor). Overlay expected distribution.
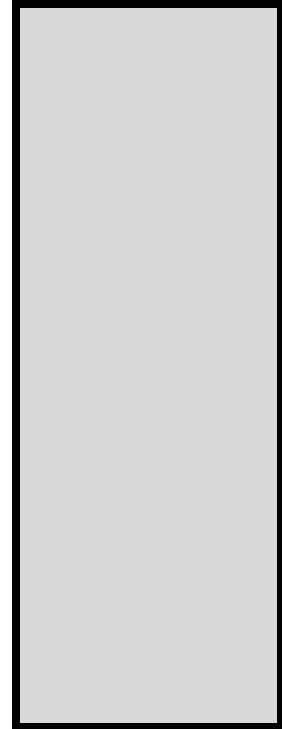
MODEL RESULTS

A. Tables of Nominal CPUE, Standardized CPUE, Observations, Positive Observations, Proportion Positive Observations and Coefficients of Variation (CVs). Other statistics may also be appropriate to report

B. Figure of Nominal and Standardized Indices with measure of variance (i.e. CVs).

IF MULTIPLE MODEL STRUCTURES WERE CONSIDERED:

(*Note: this is always recommended but required when model diagnostics are poor*.)

1. Plot of resulting indices and estimates of variance

2. Table of model statistics (e.g. AIC criteria)

	Date Received	*Workshop Recommendation*	*Revision Deadline* ***	*Author and Rapporteur Signatures*
First Submission				
Revision				

*The revision deadline is negotiated by the author, the SEDAR coordinator and the CPUE rapporteur. The author **DOES NOT** commit to any **LEGAL OBLIGATION** by agreeing to submit a manuscript before this deadline. The maximum penalty for failure to submit a revised document prior to the submission deadline is rejection of the CPUE series.*

Justification of Working Group Recommendation

62

This page intentionally left blank.

6. SEDAR Data Workshop Report Requirements for the Indices Working Group

Documentation of Working Group decisions is critical to the SEDAR process. It provides a record of the discussions and recommendations of each group not just to the Assessment Workshop Panel, but also to the Review Workshop panel, the Councils, and individual stakeholders as the assessment moves from the scientific phase to management action based on that science.

Terms of Reference are provided at each SEDAR workshop to provide information regarding what tasks are expected by the individual working groups and the Workshop Panel as a whole. The general Terms of Reference which pertain to the Indices Working Group are listed below:

General Indices Working Group Terms of Reference:
Provide measures of population abundance that are appropriate for stock assessment.

Consider all available and relevant fishery dependent and independent data sources.

Document all programs evaluated, addressing program objectives, methods, coverage, sampling intensity, and other relevant characteristics.

Provide maps of survey coverage.

Develop CPUE and index values by appropriate strata (e.g., age, size, area, and fishery); provide measures of precision and accuracy.

Evaluate the degree to which available indices adequately represent fishery and population conditions.

Recommend which data sources are considered adequate and reliable for use in assessment modeling.

The final two tasks highlighted in the above Term of Reference for the Indices Working Group as they (or the content they request) have been commented on by past SEDAR independent peer review panels, and were one of the primary driving forces for conducting this workshop.

SEDAR independent peer review panels have commented that objective and consistent criteria for evaluating indices are currently lacking. The criteria used to select particular indices for inclusion in a particular assessment have not been clearly documented, and those indices selected, along with the weighting scheme applied to them, for inclusion by the Data Workshop have been questioned. The Panels have questioned why indices of abundance displaying conflicting trends have been included in the base assessment and it has appeared to some Reviewers that little or no evaluation of the indices with respect to their relationship and/or ability to track overall abundance has been considered in the selection process. Lack of clear documentation explaining why such choices had been made has left it to the lead analyst to try and explain the rationale, despite the fact that the analyst in question may not have participated in the Indices Working Group.

Acknowledging that these are difficult tasks to complete, these issues need to be discussed by the Indices Working group at the Data Workshop, and index recommendations, along with their reasoning, need to be documented. The Index Checklist developed by this workshop will serve two purposes. First, it will provide a clear record of the issues discussed and the criteria used in recommending or rejecting particular indices and 2) it should provide the working group more time to focus and document their discussions regarding the somewhat less objective aspects of the indices being reviewed, such as how well the index may track abundance.

Summary Tables
Summary tables are useful for providing many details about the individual indices in a consolidated form. Although various table formulations have been used in SEDAR reports in the past, SEDAR is continually striving towards consistency in report elements when practicable and thus request that at least the following two tables be present in the Indices Working Group report section: Indices Description Summary Table and Indices Recommendation Summary Table. Examples of these tables can be seen in Table 1 and 2.

Suggested Fields for Indices Description Summary Table:

Series Name	Document #	Data Source (e.g. longline obs; angler interview)
Area	Years	Season
Biomass/number	Units	Sampling Design
Fishery Type (FI, FD-R, FD-C)		Standardization Method (e.g. GLM; Lo method; none)
Selectivity info	Age range	Size range

Suggested Fields for Indices Recommendations Summary Table:

Series Name	Document #
Positive Aspects	Negative Aspects
Issues Addressed	Utility for Assessment (Base, sensitivity, not)

Index Values Summary Table:

As the Data Workshop report is supposed to document all of the data reviewed, not just those recommended for use in the assessment, SEDAR requests a table containing information on all of the indices reviewed during the workshop. This table will serve as a summary of the indices reviewed, and will provide all of the values of all indices in one location, should the decision to include or exclude an index change during the process. An example of an Index Values Summary Table can be seen in Table 3.

Summary Figures

Index Coverage Maps

Summary maps displaying the area/survey coverage of the indices being considered are already required by the Terms of Reference but have often been overlooked. These maps do not need to be highly accurate but are provided to provide a quick overview of the areas being covered by the data presented. They allow for discussions regarding coverage over the species range, and may provide insight into possible concerns regarding limited spatial coverage or whether a particular index may need to be included since it covers a portion of the range missed by others indices. Additionally, it is important to remember that most CIE reviewers are international assessment biologists and not familiar with the geographic coverage of the surveys they are reviewing. An example can be seen in Figure 1.

Index Summary plots

Providing summary plots of the indices being considered for an assessment is an effective way of displaying multiple sources of information in one plot. Plotting multiple indices on the same figure requires that the individual series be scaled in some manner so that they may appear on a common scale. Series can be scaled by dividing each series by the mean of the years within that series which overlap between all series. Plotting similar indices (fishery-dependent, fishery independent, juvenile indices, etc) in a common figure allows for examination of the overall trends, along with the potential outliners or conflicting trends which may require further discussion and documentation. Examples of summary indices plots can be found in Figures 2 and 3.

Table 1: Example of an Indices Description Summary Table. Information extracted from SEDAR 11 when available.

Series	Reference	Data Source	Area	Years	Season	Biomass/ Number	Units	Fishery Type	Sampling Design	Standardization Method	Selectivity Info	Age Range	Size Range
NC #	LCS05/06-DW-01	6 Directed longline boats	North Carolina	1988-1992	All	Biomass/Number		Commercial		GLM	None	None	
LPS	LCS05/06-DW-09	Angler interviews	Virginia-Mass.	1986-2004	June-October	Number		Recreational		Lo method	Length frequency	None	
Gillnet Observer	LCS05/06-DW-11	Shark drift gillnet fishery	Florida, Georgia	1993-1995, 1998-2004	All	Number		Commercial		Lo method	Length by mesh size (based on fishery independent study)	None	
Gillnet Observer	LCS05/06-DW-11	Shark drift gillnet fishery	Florida, Georgia (Atl)	1993-1995, 1998-2004	All	Number		Commercial		Lo method	Length by mesh size (based on fishery independent study)	None	
PC LL	LCS05/06-DW-12	PC NMFS Longline Survey	NW Florida	1993-2000	Spring, Summer, Fall	Number		Independent		Lo method	Length frequency	None	
NC #	LCS05/06-DW-01	6 Directed longline boats	North Carolina	1988-1992	All	Biomass/Number		Commercial		GLM	None	None	
Gillnet Observer	LCS05/06-DW-11	Shark drift gillnet fishery	Florida, Georgia	1993-1995, 1998-2004	All	Number		Commercial		Lo method	Length by mesh size (based on fishery independent study)	None	

Table 2: Example of an Indices Recommendation Summary Table. Information extracted from SEDAR 11 when available.

Series	Reference	Positive Aspects	Negative Aspects	Issues addressed	Utility for Assessment
NC #	LCS05/06- DW-01	Historic	Not species specific, low sample sizes, possible changes in fishing methodology not accounted for		Not recommended
LPS	LCS05/06-DW-09			revisit cv calculation	Usable
Gillnet Observer	LCS05/06- DW-11	standardized		Rerun with new effort calculation, need to attempt calcs for all LCS scenarios	
Gillnet Observer	LCS05/06- DW-11	standardized		Rerun with new effort calculation, restricted to Atlantic only	
PC LL	LCS05/06- DW-12			need to attempt calcs for all LCS scenarios	Recommended

Table 3: Example of an Index Values Summary Table. Information extracted from SEDAR 11.

Table 3.2 Available catch rates series for the large coastal shark complex. The index is the relative (divided by the overall mean) estimated mean CPUE and the CV is the estimated precision of the mean value. Type refers to whether the index is fishery – independent (FI) or fishery-dependent (FD), recreational (R) or commercial (C). Observations with a CV of 1.0 are nominal data for which no measure of the precision of the estimate was available. Recommendation refers to the recommendation by the Indices Working Group to include the particular index as a base index (Base), use it for sensitivity runs (Sensitivity) or not recommended for use in the assessment (NR).

Document Number	Series Name	Type	Recommendation	Year	Index	CV
LCS05/06-DW-01	NC #	FD - C	NR	1988	0.758	0.422
				1989	1.242	0.232
LCS05/06-DW-11	Gillnet Observer	FD - C	Base	1993	0.338	1.026
				1994	1.050	0.132
				1995	0.299	0.779
				1998	1.088	0.177
				1999	1.336	0.079
				2000	1.239	0.073
				2001	1.179	0.070
				2002	1.077	0.116
				2003	1.112	0.150
				2004	1.281	0.082
LCS05/06-DW-12	PC Longline	FI	Sensitivity	1993	0.816	0.730
				1994	0.386	0.894
				1995	1.272	0.610
				1996	0.858	0.583
				1997	0.926	0.539
				1998	0.725	0.967
				1999	1.174	0.564
				2000	1.844	0.508
LCS05/06-DW-12	PC Gillnet	FI	Base	1996	0.511	0.241
				1997	1.637	0.132
				1998	0.607	0.310
				1999	0.969	0.297
				2000	0.811	0.326
				2001	1.549	0.211
				2002	0.936	0.201
				2003	1.072	0.186
				2004	0.908	0.220
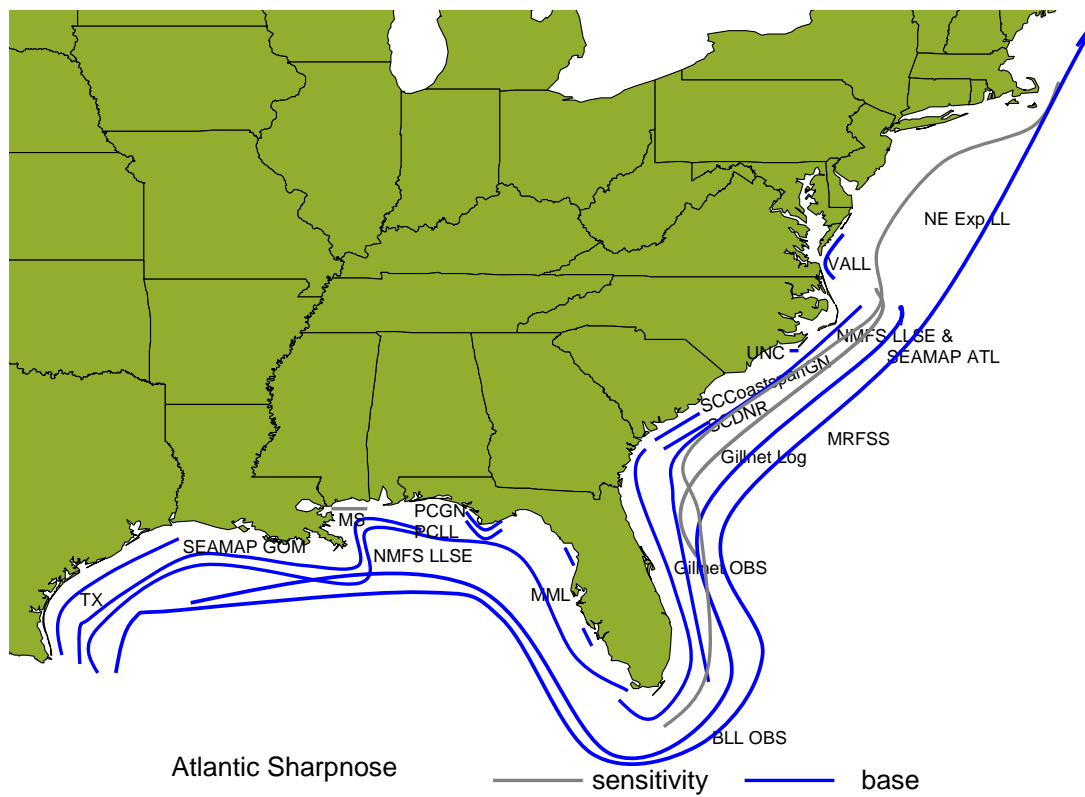
Figure 1: Example of the "General geographic coverage of relative abundance indices reviewed at the Data Workshop." (SEDAR 13)

Blacknose Shark (Fishery Independent)
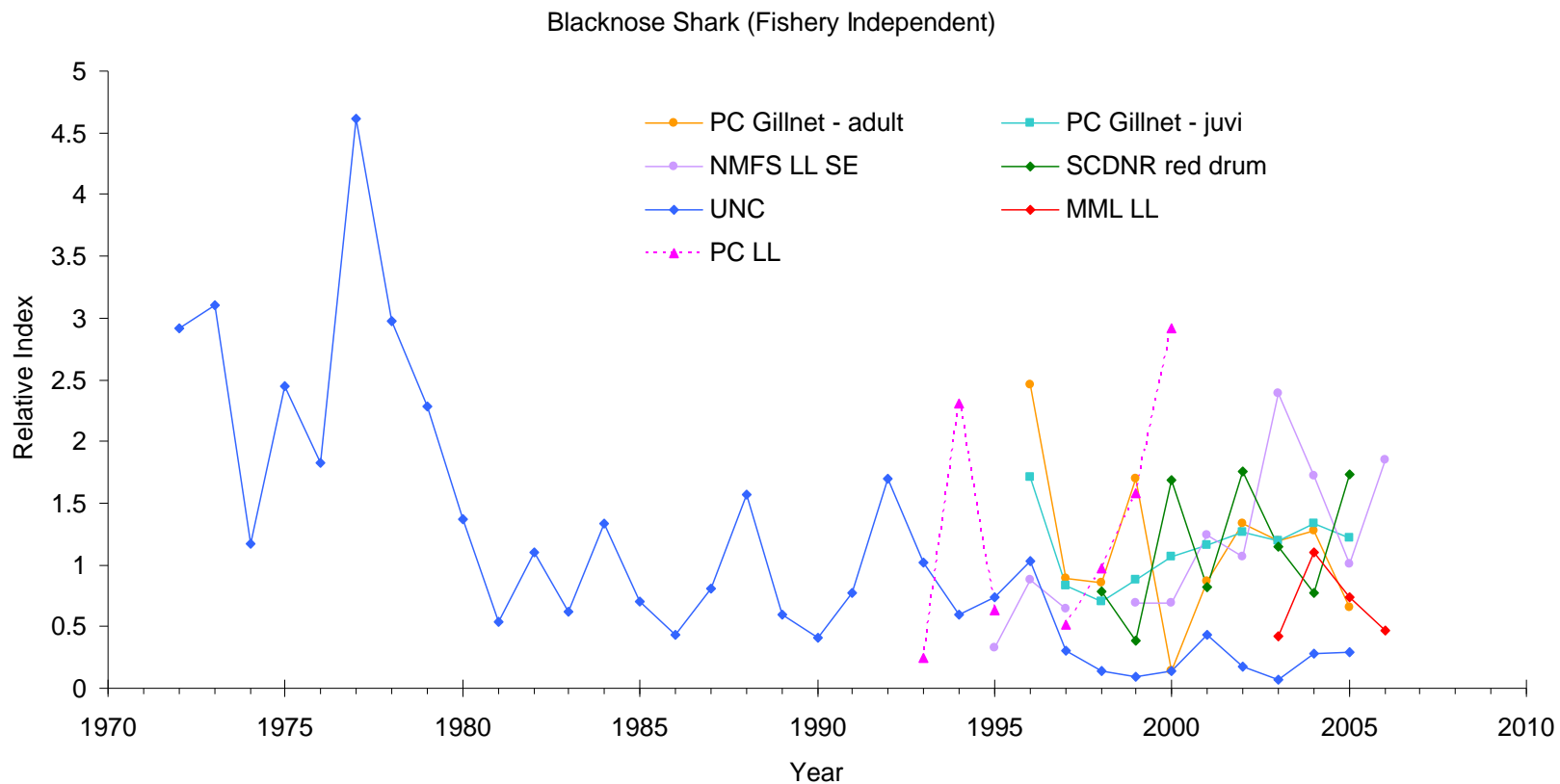
Figure 3.9. Fishery-independent catch rate series for **blacknose sharks**. Solid lines indicate base case indices while dashed lines are for series to be used in sensitivity analysis. Series are scaled (each series is divided by the mean of the years within that series which overlap between all series) to appear on a common scale.

Blacknose Shark (Fishery Dependent)
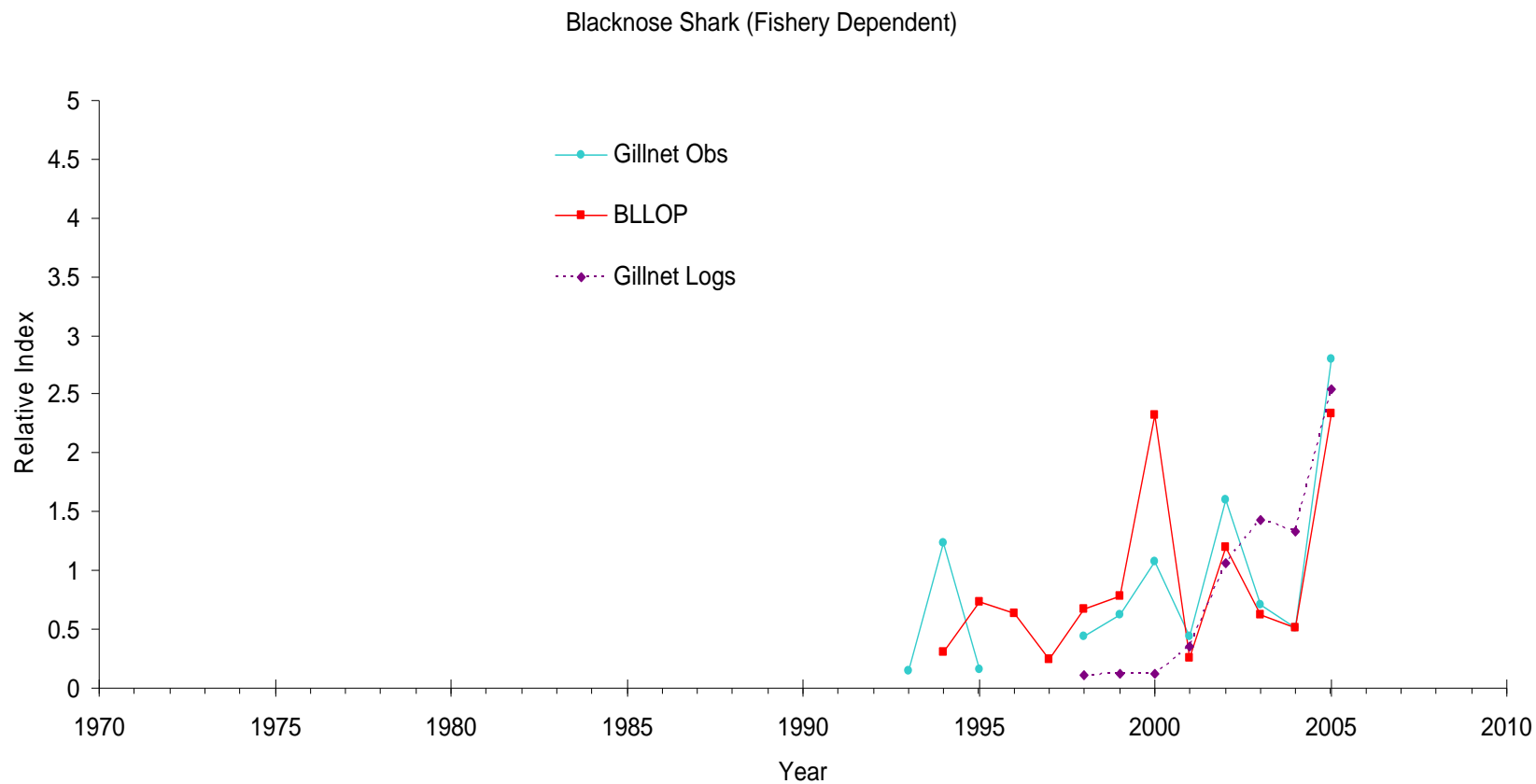
Figure 3.10. Fishery-dependent catch rate series for **blacknose sharks**. Solid lines indicate base case indices while dashed lines are for series to be used in sensitivity analysis. Series are scaled (each series is divided by the mean of the years within that series which overlap between all series) to appear on a common scale.

Appendix: Computer Code for Analyses Commonly Used In the SEDAR Process

The following appendix consists of a compilation of computer code, applicable to numerous software types (e.g. SAS, R, etc.) for analyses, which are commonly used to develop abundance indices for the SEDAR process. Each section of code will be labeled by the type of analysis that it represents. Also, this part of the document will be periodically updated with new code.

Title: Steps-Macro
Software required: SAS
Application:

```
/***************************\
* Define the macro 'steps'  *
\***************************/

%MACRO  steps ;


/********************************************************************************\
* STEP 1: Run the base model against which all subsequent models will be compared  *
\********************************************************************************/


 /* Run the model */

 PROC GENMOD  DATA=analysis &procopts ;
  CLASS  &cbase ;
  MODEL  &response = &mbase / &modopts ;
  MAKE  'MODELFIT'  OUT = modelfit  ;
  MAKE  'TYPE3'     OUT = type3fit  ;
 RUN ;


 /* Strip the required statistics (degrees of freedom, deviance,
    deviance-per-degree-of-freedom, and log-likelihood value) from the GenMod
    output tables MODELFIT and TYPE3; RETAIN and a sub-setting IF are used to
    combine the output on different lines into a single obs, and only one obs
    is created in the corresponding output data set (mfstats or t3fstats) when
    all the stats have been stripped. Finally, merge the statistics into a
    single file and append it to the cumulative output file 'outstats'. */

 DATA  mfstats ;
  SET  modelfit ;
  RETAIN  degf deviance devpdf loglike ;
  IF criterion = 'Deviance' THEN
    DO ;
      degf     = df ;
      deviance = value ;
      devpdf   = valuedf ;
    END ;
  ELSE  IF  criterion = 'Log Likelihood' THEN
    loglike = value ;
  ELSE
    DELETE ;
  factor = TRIM(UPCASE(SYMGET('mbase'))) ;
put 'factor=[' factor ']' ;
  IF  ( factor = "" )  THEN  factor = "NULL" ;
  IF  loglike NE . ;
  KEEP  degf deviance devpdf loglike factor ;
 RUN ;
```

73

```
%put &mbase ;
 %IF  &mbase = %STR()  %THEN  %DO ;
    DATA  outstats ;
     SET  outstats mfstats ;
    RUN;
   %END ;
 %ELSE  %DO ;
    DATA  t3fstats ;
     SET  type3fit ;
put 'source=[' source ']' ;
%put &mbase ;
     IF  TRIM(UPCASE(source)) = TRIM(UPCASE(SYMGET('mbase'))) ;
     KEEP  chisq ProbChiSq ;
    RUN ;
    DATA  stpstats;
     MERGE  mfstats t3fstats ;
    RUN ;
    DATA  outstats ;
     SET  outstats stpstats ;
    RUN;
   %END ;




/*************************************************************\
* STEPS 1, 2, ..., i, ... N: Run the models with one factor each  *
*   added to the base model run in step 0                 *
\*************************************************************/

/* Step through the factors one at a time using the do-loop index
  'i', which runs from 1 to the number of potential factors 'nsteps'. */

%DO  i = 1 %TO &nsteps ;

 /* Check to see if the flag for the ith factor is on or off.
   If it is on the execute the following code, otherwise jump
   to the end of the loop */

 %IF  (&&flag&i = on)  %THEN  %DO ;


 /* FIRST TASK: run the model */

 PROC GENMOD  DATA=analysis &procopts ;
  CLASS  &cbase &&cfactor&i ;
  MODEL  &response = &mbase &&mfactor&i / &modopts ;
  MAKE  'MODELFIT'  OUT = modelfit  ;
  MAKE  'TYPE3'    OUT = type3fit  ;
 RUN ;


 /* SECOND TASK: strip the required statistics (degrees of freedom, deviance,
   deviance-per-degree-of-freedom, and log-likelihood value) from the GenMod
   output tables; RETAIN and a sub-setting IF are used to combine the output
```

74

```
     on different lines into a single obs, and only one obs is created in the
     output data set 'out1' when all the stats have been stripped */

  DATA  mfstats ;
    SET  modelfit ;
*  ATTRIB  factor LENGTH=$80 ;
    RETAIN  degf deviance devpdf loglike ;
    IF criterion = 'Deviance' THEN
      DO ;
        degf    = df ;
        deviance = value ;
        devpdf  = valuedf ;
      END ;
    ELSE  IF  criterion = 'Log Likelihood' THEN
      loglike = value ;
    ELSE
      DELETE ;
    factor = UPCASE(SYMGET('mfactor' || LEFT(&i))) ;
    IF  loglike NE . ;
    KEEP  degf deviance devpdf loglike factor ;
  RUN ;
proc summary data=type3fit;
var chisq;
output out=last n=lasts;
run;
data type3fit;set type3fit;merg=1;run;
data last;set last;merg=1;drop _freq_ _type_;run;
data type3fit;merge type3fit last;by merg;run;


  DATA  t3fstats ;
    SET  type3fit ;
         last+1;
          if last=lasts;
    *IF  UPCASE(source) = UPCASE(SYMGET('mfactor' || LEFT(&i))) ;
    KEEP  chisq ProbChiSq ;
  RUN ;


  DATA  stpstats;
    MERGE  mfstats t3fstats ;
  RUN ;

  DATA  outstats ;
    SET  outstats stpstats ;
  RUN;

 %END ;

%END ;


%MEND  steps ;
```

End: Steps-Macro

Title: GLIMMIX-Macro
Software required: SAS
Application: a SAS macro for fitting generalized linear mixed models using Proc Mixed and the Output Delivery System (ODS). This macro requires SAS/STAT Version 8 or greater.

```
/*---------------------------------------------------------
TITLE
-----
GLIMMIX: a SAS macro for fitting generalized linear mixed
models using Proc Mixed and the Output Delivery System (ODS).
Requires SAS/STAT Version 8.

SUPPORT
-------
Russ Wolfinger, SAS Institute Inc.  The original version was
written by Jason Brown, formerly of SAS Institute Inc.
Please send email to russ.wolfinger@sas.com with any suggestions
or corrections.

HISTORY
-------
initial coding                        01Jun92   jbb
a few changes and additions           09Oct92   rdw
corrections from Dale McLerran, FHCRC          16Feb94   rdw
suggestions from David Murray, U. Minnesota    21Sep95   rdw
suggestions from Ken Goldberg, Wyeth-Ayerst    27Oct95   rdw
various minor updates                 06Apr96   rdw
per suggestions from Ken Goldberg, INITIAL
   option changed, INTERCEPT= option dropped,
   and FITTING, NOPREV, and NOTEST options
   added.                             12Mar97   rdw
more Goldberg ideas: NOTES option added,
   PARMS specification is only used in the
   first iteration unless you also specify
   NOPREV, some clean up              19May97   rdw
7.01 conversion                       01Jul97   rdw
switched XBETA= and PRED=             14Nov97   rdw
save spatial coordinates as suggested by
   Michael O'Kelly, Quintiles Dublin         01Dec97   rdw
eliminated LSMEANS / OM check         05Dec97   rdw
titling code from Dale McLerran, FHCRC        20Feb98   rdw
made PRINTLAST and FITTING the default        25Mar98   rdw
fixed problem with TYPE=SP(EXP)               30Apr98   rdw
allowed METHOD=MIVQUE0 to persist as
   suggested by Svetlana Rudnaya, Ford        14Aug98   rdw
changed output data set as suggested
   by Carol Gotway-Crawford, CDC             25Sep98   rdw
Mauricio Ortiz Oct-2000 CHANGES IN OUTPUT LSMEASN TO
GIVE SE FOR MU


DESCRIPTION
```

The macro uses iteratively reweighted likelihoods to fit the
model; refer to Wolfinger, R. and O'Connell, M., 1993,
``Generalized Linear Mixed Models: A Pseudo-Likelihood
Approach,'' Journal of Statistical Computation and
Simulation, 48.

By default, GLIMMIX uses restricted/residual psuedo
likelihood (REPL) to find the parameter estimates of the
generalized linear mixed model you specify. The macro calls
Proc Mixed iteratively until convergence, which is decided
using the relative deviation of the variance/covariance
parameter estimates. An extra-dispersion scale parameter is
estimated by default.

There are a few macros at the beginning; all are used in the
main macro, GLIMMIX. This macro will work on any type of
model with the error distributions and link functions given
in the ERRLINK macro. In addition, you can specify your
own error and/or link functions. In order to do this, you
must specify error=user and/or link=user in conjunction with
the errvar=, errdev=, linku=, linkud=, linkui=, and linkuid=
options.

The relevant information is saved using the MAKE statement
of Proc Mixed, which is a part of ODS.

The following are reserved variable names and should not be
used in your input SAS data set:

  col, deta, dmu, eta, lowereta, lowermu, mu, pred, resraw,
  reschi, stderreta, stderrmu, uppereta, uppermu, var, _offset,
  _orig, _w, _wght, _y, _z

The following data sets are created by the macro and exist
after completion unless certain options exclude them:

  _class, _con, _cov, _diff, _dim, _ds, _est, _fitstats, _lsm,
  _model, _pred, _predm, _slice, _soln, _solnr, _tests3

To see how each of these data sets are created, search the macro
code below. If the data set you want is not one of these, add
an appropriate MAKE statement to your STMTS= specification.

CAUTION: This macro can produce biased results for repeated
binary data with few repeats on each subject. Refer to
Breslow and Clayton (1993, JASA, 9-25).

SYNTAX

Syntax for the macro is similar to that of Proc Mixed.
There are other options that are macro-specific, however.

%glimmix(data=,
  procopt=,

78

```
stmts=,
weight=,
freq=,
error=,
errvar=,
errdev=,
link=,
linkn=,
linknd=,
linkni=,
linku=,
linkud=,
linkui=,
linkuid=,
numder=,
cf=,
converge=,
maxit=,
offset=,
out=,
outalpha=,
options=
)
```

where

data   specifies the data set you are using. It can either be a regular input data set or the _DS data set from a previous call to GLIMMIX. The latter is used to specify starting values for GLIMMIX and should be accompanied by the INITIAL option described below.

procopt specifies options appropriate for a PROC MIXED statement. Refer to the Proc Mixed documentation for more information.

stmts   specifies Proc Mixed statements for the analysis, separated by semicolons and listed as a single argument to the %str() macro function. Statements may include any of the following: CLASS, MODEL, RANDOM, REPEATED, PARMS, ID, CONTRAST, ESTIMATE, and LSMEANS. Syntax and options for each statement are exactly as in the Proc Mixed documentation. If you wish to use the OM option with the LSMEANS statement, you should specify OM=dataset to avoid conflicts with weights.

weight specifies a weighting variable for the analysis This allows you to construct your own weights which can modify or replace the ones constructed by GLIMMIX.

freq   specifies a frequency variable for the analysis. It replicates observations with the number of replicates being equal to the value of the FREQ variable.

error   specifies the error distribution. Valid types are:

> binomial|b, normal|n, poisson|p, gamma|g,
> invgaussian|ig, and user|u

> When you specify error=user, you must also provide
> the errvar= and errdev= options. The default
> error distribution is binomial.

errvar   specifies the user-defined variance function. It
         must be expressed as a function the argument "mu"
         (see examples).

errdev   specifies the user-defined deviance function. It
         must be expressed as a function the arguments
         "_y", which is the response variable, and "mu",
         which is the mean. You are allowed to use "_wght"
         also, which corresponds to the denominator of a
         binomial response. Typical deviance functions are
         as follows:

> normal   (_y-mu)**2
> poisson   2*_y*log(_y/mu);
> binomial   2*_wght*(_y*log(_y/mu)+
>   (1-_y)*log((1-_y)/(1-mu)))
> gamma   -2*log(_y/mu)
> invgaussian   (((_y-mu)**2)/(_y*mu*mu))

> The default deviance is binomial.

link   specifies the link function. Valid types are

> logit, probit, cloglog, loglog, identity,
> power(), log, exp, reciprocal, nlin, and user.
> (warning: nlin has not been tested, and it currently
> uses an MQL-type estimation scheme.)

> When you specify link=nlin, you must also provide
> the linkn=, linknd=, and linkni= options. When
> you specify link=user, you must also provide the
> ulink=, dulink=, and iulink= options. The default
> link is different for each error distribution and
> is as follows:

Distribution	Default Link
Binomial	Logit
Poisson	Log
Normal	Identity
Gamma	Reciprocal
Invgaussian	Power(-2)

linkn   specifies a nonlinear link function. It must be
        enclosed in %str() and assign a value to "mu" by
        using parameters "b1" - "bk".

80

linknd   specifies the derivative of the nonlinear link
         function.

linkni   specifies the initial values for the nonlinear
         link function.

linku   specifies a user-defined link function. It must
         be expressed as a function with the argument "mu".

linkud   specifies the derivative of the user-defined link
         function with respect to mu. It must be expressed
         as a function with argument "mu". For an
         approximation, use the formula

              (u(mu+h)-u(mu-h))/(2*h)

         where u() is the link and h is a small number.

linkui   specifies the inverse of the user-defined link.
         It must be expressed as a function with argument
         "eta".

linkuid specifies the derivative of the inverse of the
         user-defined link. It must be expressed as a function
         with argument "eta".

numder   specifies the tolerance used to numerically differentiate
         certain link functions (e.g. probit and power). It has
         a default value of 1e-5.

cf   specifies the correction factor added to the data
         in order to avoid singularities in the initial
         iteration. It has a default value of 0.5.

converge sets the convergence criterion for the GLIMMIX
         macro. This is not the convergence criteria used
         for each internal Proc Mixed call, but rather the
         criterion used to assess convergence of the entire
         macro algorithm. It has a default value of 1e-8.

maxit   specifies the maximum number of iterations for the
         GLIMMIX macro to converge. It has a default value of
         20.

offset   specifies the offset variable. By default no offset
         is used.

out   specifies a name for an output data set. This data
         set is the predicted value data set from Proc Mixed with
         the following additional variables:

         eta   = linear predictor (xbeta) + offset
         stderreta = approximate std err of eta
         lowereta = lower confidence limit for eta
         uppereta = upper confidence limit for eta

81

```
mu      = inverse link transform of eta
dmu     = derivative of mu with respect to eta
stderrmu = approx std err of mu via delta method
lowermu  = lower cl for mu, inv link transform of lowereta
uppermu  = upper cl for mu, inv link transform of uppereta
var     = variance
resraw  = raw residual, y - mu
reschi  = scaled residual, (y-mu)/sqrt(phi*var)
deta    = derivative of eta with respect to mu
_w      = weight used in final Proc Mixed call
_z      = dependent variable used in final Proc Mixed call
```

If none is given, then a default name
of _OUTFILE is used.

outalpha specifies an alpha level for the confidence limits
in the out= data set.

options specifies GLIMMIX macro options separated by
spaces:

INITIAL   specifes that the input data set is actually
the _DS data set from a previous call to
GLIMMIX.   This allows you to restart a
problem that stopped or to specify starting
values.

MQL   computes MQL estimates (see Breslow and
Clayton, 1993, JASA, 9-25). The default
is PQL with an extra-dispersion
parameter.

NOPREV   prevents use of previous covariance parameter
estimates as starting values for the next
iteration.

NOPRINT   suppresses all printing.

NOITPRINT   suppresses printing of the iteration
history.

NOTES   requests printing of SAS notes, date, and page
numbers during macro execution. By default,
the notes, date, and numbers are turned off
during macro execution and turned back on after
completion.

PRINTALL   prints all Proc Mixed runs.

PRINTDATA   prints the pseudo data after each
iteration.

OUTPUT

The output from this macro is a printout of selected tables

82

from the final iteration of Proc Mixed. All of these tables
are stored in data sets whose names begin with an
underscore; you can scan the macro code to find the name of
the data set you wish to use.

EXAMPLE SYNTAX
----
1) Both of the following examples specifiy the same
   analysis: logistic regression, no random effects

```
%glimmix(data=ingots,
  stmts=%str(
     class soak;
     model nready/ntotal=soak heat;
   )
)

%glimmix(data=ingots,
  stmts=%str(
     class soak;
     model nready/ntotal=soak heat;
   ),
  error=user,errvar=mu*(1-mu),
  errdev=2*_wght*(_y*log(_y/mu) +
       (1-_y)*log((1-_y)/(1-mu))),
  link=user,
  linku=log(mu/(1-mu)),
  linkud=1/(mu*(1-mu)),
  linkui=exp(eta)/(1+exp(eta)),
  linkuid=-exp(eta)/(1+exp(eta))**2;
)
```

   Here _wght corresponds to ntotal and _y to nready/ntotal.

2) This example uses the random, lsmeans, and options
   arguments:

```
%glimmix(data=salaman1,
  stmts=%str(
     class fpop fnum mpop mnum;
     model y = fpop|mpop;
     random fpop*fnum mpop*mnum;
     lsmeans fpop|mpop / cl;
   )
  options=noitprint
)
```

3) This example uses the procopt, random, and offset
   arguments:

```
%glimmix(data=ship,
  procopt=order=data,
  stmts=%str(
     class type year period;
     model y=type;
```

```
      random year period;
    ),
    error=poisson,link=log,offset=service
  )
```

4) This example uses the repeated argument:

```
  %glimmix(data=salaman1,
    stmts=%str(
      class fpop fnum mpop mnum;
      model y = fpop|mpop;
      repeated / type=ar(1) sub=fpop*fnum;
      lsmeans fpop|mpop / cl;
    )
  )
```

DISCLAIMER

THIS INFORMATION IS PROVIDED BY SAS INSTITUTE INC. AS A
SERVICE TO ITS USERS. IT IS PROVIDED "AS IS". THERE ARE NO
WARRANTIES, EXPRESSED OR IMPLIED, AS TO MERCHANTABILITY OR
FITNESS FOR A PARTICULAR PURPOSE REGARDING THE ACCURACY OF
THE MATERIALS OR CODE CONTAINED HEREIN.

---*/


```
/*---------------------------------------------------------*/
/*                                                         */
/*   %mvarlst                                              */
/*   Make a variable list from the class list, model       */
/*   specification, and random specification.              */
/*                                                         */
/*---------------------------------------------------------*/
%macro mvarlst;
%let varlst =;
%let mdllst = &mdlspec;

/*---get response variable---*/
%if %index(&response,/) %then
  %let varlst = %scan(&response,1,/) %scan(&response,2,/) &varlst;
%else %let varlst = &response &varlst;

/*---get fixed effects---*/
%if %index(&mdllst,@) %then %do;
  %let j = 1;
  %let mdl = &mdllst;
  %let mdllst=;
  %do %while(%length(%scan(&mdl,&j,' ')));
    %let var=%scan(&mdl,&j,' ');
    %if %index(&var,@) %then %do;
      %let b = %eval(%index(&var,@)-1);
      %let mdllst = &mdllst %substr(%quote(&var),1,&b);
    %end;
```

```
      %else %let mdllst = &mdllst &var;
        %let j = %eval(&j+1);
      %end;
  %end;

  %let iv = 1;
  %do %while (%length(%scan(&mdllst,&iv)));
    %let varlst = &varlst %scan(&mdllst,&iv);
    %let iv = %eval(&iv + 1);
  %end;

  /*---get random effects---*/
  %let iv = 1;
  %do %while (%length(%scan(&rndlst,&iv)));
    %let temp = %scan(&rndlst,&iv);
    %if &temp ne INT and &temp ne INTERCEPT %then
      %let varlst = &varlst &temp;
    %let iv = %eval(&iv + 1);
  %end;

  /*---get repeated effects---*/
  %let iv = 1;
  %do %while (%length(%scan(&replst,&iv)));
    %let temp = %scan(&replst,&iv);
    %if &temp ne DIAG %then %let varlst = &varlst &temp;
    %let iv = %eval(&iv + 1);
  %end;

  %let varlst = &varlst &class &id &freq &weight;
%mend mvarlst;


/*------------------------------------------------------------*/
/*                                                            */
/*   %trimlst                                                 */
/*   Get rid of repetitions in a list                        */
/*                                                            */
/*------------------------------------------------------------*/
%macro trimlst(name,lst);
%let i1 = 1;
%let tname =;
%do %while (%length(%scan(&lst,&i1,%str( ))));
  %let first = %scan(&lst,&i1,%str( ));
  %let i2 = %eval(&i1 + 1);
  %do %while (%length(%scan(&lst,&i2,%str( ))));
    %let next = %scan(&lst,&i2,%str( ));
    %if %quote(&first) = %quote(&next) %then %let i2=10000;
    %else %let i2 = %eval(&i2 + 1);
  %end;
  %if (&i2<10000) %then %let tname = &tname &first;
  %let i1 = %eval(&i1 + 1);
%end;
%let &name = &tname;
%mend trimlst;


/*------------------------------------------------------------*/
```

```
/*                                              */
/*   %remove                                    */
/*   Remove a word from a string                */
/*                                              */
/*-------------------------------------------------------------*/
%macro remove(sntnc,wrd);
%let sentence=%str( )%nrbquote(&sntnc);
%if &sentence^=%str( ) %then %do;
  %let word=%str( )%nrbquote(&wrd);
  %let answer=;
  %let i=%index(&sentence,&word);
  %if &i and &word^=%str( ) %then %do;
    %if &i>1 %then %let answer=%qsubstr(&sentence,1,&i-1);
    %let j=%eval(&i+%index(%qsubstr(&sentence,&i+1),%str( )));
    %if &j>&i %then
    %let answer=&answer%qsubstr(&sentence,&j);
  %end;
  %else %let answer=&sentence;
  %unquote(&answer)
%end;
%mend remove;



/*-------------------------------------------------------------*/
/*                                              */
/*   %errlink                                   */
/*   Create macro variables that contain the link, invlink,    */
/*   derivative, and the variance funtion for the following    */
/*   error distributions and link functions:               */
/*                                              */
/*   error distn: normal, binomial, poisson, gamma, and       */
/*              inverse gaussian                    */
/*   link func:   logit, probit, complementary log-log,       */
/*              log-log, identity, power(), log, exp, and    */
/*              reciprocal.                       */
/*                                              */
/*   The user-defined specification is given by leaving the    */
/*   error distribution field blank and then giving the link,  */
/*   the derivative of the link, the inverse link, and the     */
/*    variance function.  The parameters for each are:       */
/*                                              */
/*       mu:   variance function, link, and the derivative   */
/*            of the link                      */
/*       eta:   inverse link;                      */
/*                                              */
/*-------------------------------------------------------------*/
%macro errlink;

/*---error distributions: set variance and deviance functions---*/
%let exiterr = 0;
%if %length(&error)=0 %then %do;
  %if %length(&errvar) and %length(&errdev) %then %let error=USER;
  %else %let error=BINOMIAL;
%end;
%if %length(&linkn) and %length(&link)=0 %then %let link=NLIN;
%if %length(&linku) and %length(&link)=0 %then %let link=USER;
```

```
%if &error=BINOMIAL or &error=B %then %do;
  %let errorfn=BINOMIAL;
  %let varform=mu*(1-mu);
  %let devform=_wght*2*(_y*log(_y/mu) + (1-_y)*log((1-_y)/(1-mu)));
  %if %length(&link)=0 %then %let link=LOGIT;
%end;
%else %if &error=POISSON or &error=P %then %do;
  %let errorfn=POISSON;
  %let varform=mu;
  %let devform=_wght*2*_y*log(_y/mu);
  %if %length(&link)=0 %then %let link=LOG;
%end;
%else %if &error=NORMAL or &error=N %then %do;
  %let errorfn=NORMAL;
  %let varform=1;
  %let devform=_wght*(_y-mu)**2;
  %if %length(&link)=0 %then %let link=IDENTITY;
%end;
%else %if &error=GAMMA or &error=G %then %do;
  %let errorfn=GAMMA;
  %let varform=mu**2;
  %let devform=_wght*-2*log(_y/mu);
  %if %length(&link)=0 %then %let link=RECIPROCAL;
%end;
%else %if &error=INVGAUSSIAN or &error=IG %then %do;
  %let errorfn=INVGAUSSIAN;
  %let varform=mu**3;
  %let devform=_wght*(((_y-mu)**2)/(_y*mu*mu));
  %if %length(&link)=0 %then %let link=INVGAUSSIAN;
%end;
%else %if &error=USER or &error=U %then %do;
  %let errorfn=USER;
  %if %length(&errvar) %then %let varform=&errvar;
  %else %let exiterr = 1;
  %if %length(&errdev) %then %let devform=&errdev;
  %else %let exiterr = 1;
  %if %length(&link)=0 %then %let link=LOGIT;
%end;

/*---truncate link function, so we can match if a power link---*/
%if %length(&link)>5 %then %let trlink=%substr(&link,1,5);
%else %let trlink=&link;

/*---link functions; set eta, mu, and derivative formulas---*/
%if &trlink=LOGIT %then %do;
  %let linkfn=LOGIT;
  %let etaform=log(mu/(1-mu));
  %let detaform=1/(mu*(1-mu));
  %let muform=exp(eta)/(1+exp(eta));
  %let dmuform=exp(eta)/(1+exp(eta))**2;
%end;
%else %if &trlink=PROBI %then %do;
  %let linkfn=PROBIT;
  %let etaform=probit(mu);
  %let detaform=(probit(mu+&numder)-probit(mu-&numder))/
    (2*&numder);
```

87

```
  /*
  %let detaform=(probit(min(1,mu+&numder*(1+abs(mu)))) -
    probit(max(1,mu-&numder*(1+abs(mu))))) /
    (2*&numder*(1+abs(mu)));
  */
  %let muform=probnorm(eta);
  %let dmuform=(probnorm(eta+&numder)-probnorm(eta-&numder))/
    (2*&numder);
%end;
%else %if &trlink=CLOGL %then %do;
  %let linkfn=COMPLEMENTARY LOG LOG;
  %let etaform=log(-log(1-mu));
  %let detaform=-1/((1-mu)*log(1-mu));
  %let muform=1-exp(-exp(eta));
  %let dmuform=exp(-exp(eta))*exp(eta);
%end;
%else %if &trlink=LOGLO %then %do;
  %let linkfn=LOG LOG;
  %let etaform=-log(-log(mu));
  %let detaform=-1/(mu*log(mu));
  %let muform=exp(-exp(-eta));
  %let dmuform=exp(-exp(-eta))*exp(-eta);
%end;
%else %if &trlink=IDENT %then %do;
  %let linkfn=IDENTITY;
  %let etaform=mu;
  %let detaform=1;
  %let muform=eta;
  %let dmuform=1;
%end;
%else %if &trlink=LOG %then %do;
  %let linkfn=LOG;
  %let etaform=log(mu);
  %let detaform=1/mu;
  %let muform=exp(eta);
  %let dmuform=exp(eta);
%end;
%else %if &trlink=EXP %then %do;
  %let linkfn=EXPONENTIAL;
  %let etaform=exp(mu);
  %let detaform=exp(mu);
  %let muform=log(eta);
  %let dmuform=1/eta;
%end;
%else %if &trlink=RECIP %then %do;
  %let linkfn=INVERSE;
  %let etaform=1/mu;
  %let detaform=-1/mu**2;
  %let muform=1/eta;
  %let dmuform=-1/eta**2;
%end;
%else %if &trlink=POWER %then %do;
  %let linklen = %eval(%length(&link)-7);
  %let expon=%substr(&link,7,&linklen);
  %let linkfn=POWER(&expon);
  %let etaform=mu**(&expon);
```

```
  %let detaform=(&expon)*mu**(&expon-1);
  /*
  %let detaform=((mu+&numder)**(&expon)-(mu-&numder)**(&expon))/
    (2*&numder);
  %let detaform=((mu+&numder*(1+abs(mu)))**(&expon) -
    (mu-&numder*(1+abs(mu)))**(&expon))/
    (2*&numder*(1+abs(mu)));
  */
  %let muform=eta**(1/(&expon));
  %let dmuform=(1/(&expon))*eta**(1/(&expon)-1);
%end;
%else %if &trlink=INVGA %then %do;
  %let linkfn=POWER(-2);
  %let etaform=mu**(-2);
  %let detaform=-2*mu**(-3);
  %let muform=eta**(-1/2);
  %let dmuform=(-1/2)*eta**(-3/2);
%end;
%else %if &trlink=BOXCO %then %do;
  %let linkfn=BOX-COX;
  %let linklen = %eval(%length(&link)-8);
  %let expon=%substr(&link,8,&linklen);
  %let etaform=(mu**(&expon)-1)/(&expon);
  %let detaform=mu**((&expon)-1);
  %let muform=((&expon)*eta + 1)**(1/(&expon));
  %let dmuform=((&expon)*eta + 1)**(1/(&expon)-1);
%end;
%else %if &trlink=USER %then %do;
  %let linkfn=USER;
  %if %length(&linku) %then %let etaform=&linku;
  %else %let exiterr = 1;
  %if %length(&linkud) %then  %let detaform=&linkud;
  %else %let exiterr = 1;
  %if %length(&linkui) %then %let muform=&linkui;
  %else %let exiterr = 1;
  %if %length(&linkuid) %then %let dmuform=&linkuid;
  %else %let exiterr = 1;
%end;
%else %if &trlink=NLIN %then %do;
  %let linkfn=NONLINEAR;
  %if %length(&linkn) %then %let nlinform=&linkn;
  %else %let exiterr = 1;
  %if %length(&linknd) %then %let nlinder=&linknd;
  %else %let exiterr = 1;
%end;

%if %index(&options,DEBUG) %then %do;
  %put options = &options;
  %put intopt = &intopt;
  %put varlst = &varlst;
  %put error = &errorfn;
  %put variance = &varform;
  %put deviance = &devform;
  %put link:  eta = &etaform;
  %put dlink:  deta = &detaform;
  %put invlink:  mu = &muform;
```

```
%end;

%mend errlink;



/*------------------------------------------------------------*/
/*                                        */
/*   %init                          */
/*   Sets the initial values for the iterations.        */
/*                                        */
/*------------------------------------------------------------*/
%macro init;

%let off = &offset;

%if %index(&intopt,NLIN) %then %do;
   /*---determine number of parameters---*/
   %let nb = 0;
   %let i = 1;
   %do %while(%index(&linkni,B&i));
      %let nb = %eval(&nb + 1);
      %let i = %eval(&i + 1);
   %end;
   %let nu = 0;
   %let ns = 0;
   %let nus = 0;
   %let varlst = &varlst one mu;
%end;

data _ds;
   set %unquote(&data);
   %if not %index(&options,INITIAL) %then %do;
      /*---move away from parameter space boundary for the binomial
           error situation---*/
      %if %index(&response,/) %then %do;
         mu = (%scan(&response,1,/) + &cf)/(%scan(&response,2,/) +
            2*&cf);
         _wght = %scan(&response,2,/) ;
      %end;
      %else %if &errorfn=BINOMIAL %then %do;
         mu = (&response + &cf)/(1 + 2*&cf);
         _wght = 1;
      %end;
      %else %do;
         mu = &response + &cf;
         _wght = 1;
      %end;
      %if %length(&weight) %then %do;
         _wght = &weight * _wght;
      %end;
      _y     = &response;
      var    = &varform;
      _offset = &off;
      %if %index(&intopt,NLIN) %then %do;
         array b{&nb} b1-b&nb;
```

90

```
      array db{&nb} db1-db&nb;
      one = 1;
      %do i = 1 %to &nb;
        %let idx = %index(&linkni,B&i);
        b&i = %scan(%substr(&linkni,&idx),2,'=' ' ');
      %end;
      &nlinform
      _z = _y - mu;
      &nlinder
      do i = 1 to &nb;
        _z = _z + db{i}*b{i};
      end;
      _w = _wght / var;
    %end;
    %else %do;
      eta  = &etaform ;
      deta = &detaform ;
      _w = _wght / ((deta**2)*(var));
      _z = (_y-mu)*deta + eta - _offset;
    %end;
    %if %length(&freq) %then %do;
      do i = 1 to &freq;
        if i=1 then _orig='y';
        else _orig='n';
        output;
      end;
    %end;
    %else %do;
      _orig='y';
    %end;
    if (_w = .) then _w = 1;
    /*
    keep _y _z _w _offset _wght _orig &varlst;
    */
  %end;
run;

%if %index(&options,PRINTDATA) %then %do;
  proc print;
  run;
%end;

%let iter = 0;

%mend init;


/*----------------------------------------------------------*/
/*                                                          */
/*   %newdata                                               */
/*   Create the new data set with the updated values        */
/*                                                          */
/*----------------------------------------------------------*/
%macro newdata;
/*---save previous parameter estimates---*/
data _oldsoln;
```

```
   set _soln;
run;

/*---save previous estimates of covariance matrix---*/
data _oldcov;
  set _cov;
  %let covsaved = 1;
run;

%if %index(&options,DEBUG) %then %put Creating new pseudo data.;

/*---create new data set---*/
data _ds;
  %if %index(&intopt,NLIN) %then %do;
    set _ds;
    array b{&nb} b1-b&nb;
    array db{&nb} db1-db&nb;
    /*
    array u{&nu,&ns} u1-u&nus;
    array du{&nu,&ns} du1-du&nus;
    array _r{&nu} _r1-_r&nu;
    */
    &nlinform
    _z = _y - mu;
    &nlinder
    do i = 1 to &nb;
      _z = _z + db{i}*b{i};
    end;
    /*
    do i = 1 to &nu;
      _r{i} = du{i,subject};
      _z = _z + _r{i}*u{i,subject};
      end;
    end;
    */
    var = &varform;
    _w = _wght / var;
    /*
    keep _y _z _w _offset _wght _orig &varlst db1-db&nb one x;
    */
  %end;
  %else %do;
    set _pred;
    eta = pred + _offset;
    mu = &muform;
    deta = &detaform;
    var = &varform;
    %if %index(&options,HYBRID) %then %do;
      eta = (pred + _offset)/1.5;
      mu = &muform;
      deta = &detaform;
      eta = pred + _offset;
      mu = &muform;
      /*
      mu = &muform;
      var = &varform;
```

92

```
      */
    %end;
    _w = _wght / ((deta**2)*(var));
    _z = (_y - mu)*deta + eta - _offset;
    /*
    keep _y _z _w _offset _wght _orig &varlst;
    */
    %end;
  run;

%if %index(&options,PRINTDATA) %then %do;
  proc print;
  run;
%end;

%mend newdata;




/*------------------------------------------------------------*/
/*                                                            */
/*   %mixed                                                   */
/*   Calculate parameter estimates using Proc Mixed.          */
/*                                                            */
/*------------------------------------------------------------*/
%macro mixed;

%if %index(&options,DEBUG) %then %put Calling Proc Mixed.;

%let mivque0 = 0;

%again:

/*---get rid of predicted data set---*/
/*proc datasets lib=work nolist;
  delete _pred;
run;*/

/*---use mivque0 if did not converge the first time---*/
%if (&mivque0 = 1) %then %do;
  /*---save the original method---*/
  %let procopt0 = &procopt;
  %let procopt = &procopt METHOD=MIVQUE0;
%end;

proc mixed data=_ds &procopt;
  %if %length(&class) %then %do;
    class &class ;
    %if not %index(&procopt,NOCLPRINT) %then %do;
      ods output classlevels=_class;
    %end;
  %end;
  model _z = %unquote(&mdlspec) %unquote(&mdlopt);
  weight _w;
  ods output modelinfo=_model;
  ods output dimensions=_dim;
```

93

```sas
    ods output covparms=_cov;
    ods output fitstatistics=_fitstats;
    ods output solutionf=_soln;
    %if %length(%scan(&mdlspec,1)) and not %index(&mdlopt,NOTEST)
      %then %do;
      ods output tests3=_tests3;
    %end;
    %if %length(&spec) %then %do;
      %unquote(&spec)
      %if %index(&intopt,SOLNR) %then %do;
        ods output solutionr=_solnr;
      %end;
      %if %index(&spec,ESTIMATE) %then %do;
        ods output estimates=_est;
      %end;
      %if %index(&spec,CONTRAST) %then %do;
        ods output contrasts=_con;
      %end;
      %if %index(&spec,LSMEANS) %then %do;
        ods output lsmeans=_lsm;
        %if %index(&intopt,LSMDIFF) %then %do;
          ods output diffs=_diff;
        %end;
        %if %index(&intopt,LSMSLICE) %then %do;
          ods output slices=_slice;
        %end;
      %end;
    %end;
    %if &covsaved=1 and not %index(&options,NOPREV) and
      not %index(&procopt,MIVQUE0) and (&mivque0 = 0) %then
      %str(parms / pdata=_oldcov &parmopt2;);
    %else %if (%length(&parmspec) or %length(&parmopt)) and
      (&mivque0 = 0) %then %do;
      parms &parmspec &parmopt ;
    %end;
    id _y _offset _wght _orig &varlst;
run;

%if %index(&options,PRINTDATA) %then %do;
  proc print data=_pred;
  run;
%end;

/*---check for convergence, if not, then run again
    with method=mivque0---*/
%if (&mivque0 = 1) %then %do;
  %let procopt = &procopt0;
  %let mivque0 = 0;
%end;
%else %do;
  %let there = no;
  data _null_;
    set _pred;
    call symput('there','yes');
  run;
  %if ("&there" = "no") %then %do;
```

```
      %if not %index(&options,NOPRINT) %then %do;
        %put Computing MIVQUE0 estimates in iteration &iter because;
        %put %str(   )Proc Mixed did not converge.;
      %end;
      %let mivque0 = 1;
      %goto again;
    %end;
%end;

/*---set up for hybrid Taylor series---*/
%if %index(&options,HYBRID) %then %do;
  data _predm;
    set _predm;
    predm = pred;
    keep predm;
  run;
  data _pred;
    merge _predm _pred;
  run;
%end;

/*---merge in new estimates of b and u for nonlinear link---*/
%if %index(&intopt,NLIN) %then %do;
  %if (&nb) %then %do;
    proc transpose data=_soln out=_beta;
      var estimate;
    run;
    data _beta;
      set _beta;
      array b{&nb} b1-b&nb;
      array col{&nb} col1-col&nb;
      do i = 1 to &nb;
        b{i} = col{i};
        end;
      one = 1;
      keep one b1-b&nb;
    run;
    data _ds;
      set _ds;
      drop b1-b&nb;
    run;
    data _ds;
      merge _ds _beta;
      by one;
    run;
  %end;
  /*---this isn't finished---*/
  /*
  %if (&nu) %then %do;
    proc transpose data=_solnr out=_blup;
      var estimate;
    run;
    data _blup;
      set _blup;
      array u{&nus} u1-u&nus;
      array col{&nus} col1-col&nus;
```

95

```
      do i = 1 to &nus;
        u{i} = col{i};
        end;
      one = 1;
      keep one u1-u&nus;
    run;
  %end;
  */
%end;

%mend mixed;



/*----------------------------------------------------------*/
/*                                            */
/*   %compare                                 */
/*   Compare the last two parameter estimates to check for    */
/*   convergence if no random components, else compare        */
/*   estimates of covariance matrix.                */
/*                                            */
/*----------------------------------------------------------*/
%macro compare;

/*---Use relative difference of parameter estimates and
     covariance matrix as a measure of convergence---*/
%let crit = 0;
%compit(soln,estimate);
%compit(cov,estimate);

data _null_;
  crit = &crit;
  if (crit<&converge) then conv = 1;
  else conv = 0;
  call symput('conv',conv);
run;

%mend compare;



%macro compit(type,est);

/*---save convergence information in conv and crit---*/
data _compare;
  merge _old&type(rename=(&est=oldest)) _&type end=last;
  retain crit &crit;
  denom = (abs(oldest) + abs(&est))/2;
  if (denom > &converge) then do;
    reldiff = abs(oldest - &est) / denom;
    crit = max(crit,reldiff);
  end;
  output;
  if last then do;
    call symput('crit',left(crit));
  end;
run;
```

96

```
%if %index(&options,DEBUG) %then %do;
  proc print data=_compare;
  run;
%end;

%mend compit;




/*----------------------------------------------------------*/
/*                                          */
/*   %iterate                            */
/*   Iteration process                       */
/*                                          */
/*----------------------------------------------------------*/
%macro iterate;

%let conv = 0;
%let iter = 1;

%do %while(&iter <= &maxit);
  %newdata
  %mixed
  %compare
  %if not %index(&options,NOPRINT) and
     not %index(&options,NOITPRINT) %then %do;
    %if (&iter=1) %then %do;
      %put %str(   ) GLIMMIX Iteration History;
      %put;
      %put Iteration    Convergence criterion;
    %end;
    %if (&iter<10) %then
      %put %str(   ) &iter          &crit;
    %else %if (&iter<100) %then
      %put %str(  ) &iter            &crit;
    %else
      %put %str( ) &iter            &crit;
  %end;
  %let iter = %eval(&iter+1);
  %if (&conv=1) %then %let iter=%eval(&maxit+1);
%end;

%mend iterate;




/*----------------------------------------------------------*/
/*                                            */
/*   %compile                            */
/*   Compile the macro results.                 */
/*                                          */
/*----------------------------------------------------------*/
%macro compile;

/*---get variance estimate---*/
%let scale = 1;
```

97

```
data _null_;
  set _cov;
  if covparm='Residual' then call symput('scale',estimate);
run;

/*---calculate deviance and Pearson Chi-Squared---*/
%if %index(&intopt,NLIN) %then %do;
  data _pred;
    set _pred;
    one = 1;
  run;
  data _pred;
    merge _pred _beta;
    by one;
  run;
%end;

data _stats;
  set _pred end=last;
  retain deviance 0 pearson 0;
  if ((_y ne .) and (pred ne .)) then do;
    _y  = _y + 1e-10*(_y=0) - 1e-10*(_y=1);
    eta = pred + _offset;
    %if %index(&intopt,NLIN) %then %do;
      &nlinform;
    %end;
    %else %do;
      mu = &muform;
    %end;
    deviance + &devform;
    pearson  + _wght * ((_y-mu)**2/(&varform));
  end;
  if last;
  keep deviance pearson;
run;

data _stats;
  length descript $35;
  set _stats;
  descript = 'Deviance';
  value = deviance; output;
  descript = 'Scaled Deviance';
  value = deviance / &scale; output;
  descript = 'Pearson Chi-Square';
  value = pearson; output;
  descript = 'Scaled Pearson Chi-Square';
  value = pearson / &scale; output;
  descript = 'Extra-Dispersion Scale';
  value = &scale; output;
  keep descript value;
  label descript = 'Description' value = 'Value';
run;

/*---ESTIMATE statement results---*/
%if %index(&spec,ESTIMATE) %then %do;
  data _est;
```

98

```
    set _est;
    %if not %index(&intopt,NLIN) %then %do;
      eta = estimate;
      mu = &muform;
      label mu = 'Mu';
      drop eta;
      %if %index(&intopt,ESTCL) %then %do;
        eta = lower;
        lowermu = &muform;
        eta = upper;
        uppermu = &muform;
        label lowermu = 'LowerMu' uppermu = 'UpperMu';
      %end;
    %end;
  run;
%end;

/*---least squares means---*/
%if %index(&spec,LSMEANS) %then %do;
  data _lsm;
    set _lsm;
    %if not %index(&intopt,NLIN) %then %do;
      eta = estimate;
          stderreta = stderr;
      mu = &muform;
      dmu = &dmuform;
      stderrmu = abs(dmu)*stderreta;
      label mu = 'Mu';
      %if %index(&intopt,LSMCL) %then %do;
        eta = lower;
        lowermu = &muform;
        eta = upper;
        uppermu = &muform;
        label lowermu = 'LowerMu' uppermu = 'UpperMu';
      %end;
      drop eta;
    %end;
  run;
%end;

%mend compile;


/*------------------------------------------------------------*/
/*                                                  */
/*   %printout                               */
/*   Print out the macro results.                */
/*                                                  */
/*------------------------------------------------------------*/
%macro printout;

&titlen 'GLIMMIX Model Statistics';
proc print data=_stats noobs label;
  format value 10.4;
run;
&titlen;
```

99

%mend printout;

```
/*----------------------------------------------------------*/
/*                                          */
/*   %outinfo                           */
/*   Make an output data set.                 */
/*                                          */
/*----------------------------------------------------------*/
```
%macro outinfo(outfile);

```
 data &outfile;
   set _pred;
   eta = pred + _offset;
   stderreta = stderrpred;
   lowereta = lower + _offset;
   uppereta = upper + _offset;
   mu = &muform;
   dmu = &dmuform;
   stderrmu = abs(dmu)*stderreta;
   eta = lowereta;
   lowermu = &muform;
   eta = uppereta;
   uppermu = &muform;
   eta = pred + _offset;
   var = &varform;
   resraw = _y - mu;
   reschi = (_y-mu)/sqrt(&scale * var);
   deta = &detaform;
   _w = _wght /((deta**2)*(var));
   _z = (_y - mu)*deta + eta - _offset;
   if _orig='y';
 run;
```

%mend outinfo;

```
/*----------------------------------------------------------*/
/*                                          */
/*   %glimmix                           */
/*   Put it all together                    */
/*                                          */
/*----------------------------------------------------------*/
```
%macro glimmix(data=,procopt=,stmts=,weight=,freq=,
   error=BINOMIAL,errvar=,errdev=,link=,linku=,linkud=,
   linkui=,linkuid=,linkn=,linknd=,linkni=,numder=**1e-5**,cf=**0.5**,
   converge=**1e-8**,maxit=**20**,offset=**0**,out=,outalpha=**0.05**,options=);

%let options = %qupcase(&options);
%if %index(&options,DEBUG) %then %put Initializing.;
%else %if not %index(&options,NOTES) %then %do;
   options nonotes nodate nonumber;
%end;

/*---default data set---*/

100

```
%if %bquote(&data)= %then %let data=&syslast;
%let exiterr = 0;
%let covsaved = 0;
%let there = no;
/*---check that it is there---*/
data _null_;
  set &data;
  call symput('there','yes');
run;
%if ("&there" = "no") %then %do;
  %let exiterr = 1;
  %goto exit;
%end;

/*---change to uppercase---*/
%let data    = %qupcase(&data);
%let procopt  = %qupcase(&procopt);
%let weight   = %qupcase(&weight);
%let freq    = %qupcase(&freq);
%let stmts   = %qupcase(&stmts);
%let error   = %qupcase(&error);
%let errvar  = %qupcase(&errvar);
%let errdev   = %qupcase(&errdev);
%let link    = %qupcase(&link);
%let linkn   = %qupcase(&linkn);
%let linknd   = %qupcase(&linknd);
%let linkni  = %qupcase(&linkni);
%let linku   = %qupcase(&linku);
%let linkud   = %qupcase(&linkud);
%let linkui  = %qupcase(&linkui);
%let linkuid  = %qupcase(&linkuid);
%let offset  = %qupcase(&offset);
%let outfile  = %qupcase(&out);
%let outalpha = %qupcase(&outalpha);
%let options  = %qupcase(&options);

/*---title---*/
%let ntitle=0;
%let titlesas=;
data _null_;
  set sashelp.vtitle;
  if (number=1) and (text="The SAS System") then
    call symput("titlesas",right(text));
  else call symput("ntitle",left(put(number,2.)));
  if number=10 then call symput("title10",text);
run;
%if &ntitle=10 %then %let titlen = title10;
%else %let titlen = title%eval(&ntitle + 1);

/*---loop through statements and extract information---*/
%let spec = ;
%let class = ;
%let parms = ;
%let id = ;
%let rndlst = ;
%let replst = ;
```

```
%let intopt = ;
%let iv = 1;
%do %while (%length(%scan(&stmts,&iv,;)));

  %let stmt = %qscan(&stmts,&iv,%str(;));
  %let first = %qscan(&stmt,1);
  %let fn = %eval(%index(&stmt,&first) + %length(&first));

  /*---check RANDOM options and extract random effects list---*/
  %if %index(&first,RANDOM) or %index(&first,REPEATED) %then %do;
    %let intopt = &intopt COVMOD;
  %end;
  %if %index(&first,RANDOM) %then %do;
    %let i = %index(&stmt,/);
    %if &i = 0 %then %let i = %length(&stmt);
    %else %do;
      %let rndopt = %substr(&stmt,&i);
      %if %index(&rndopt,%str( S )) or %index(&rndopt,SOLUTION) or
        %index(&rndopt,%str( CL )) or %index(&rndopt,ALPHA) %then %do;
        %let intopt = &intopt SOLNR;
      %end;
    %end;
    %let rndlst = %substr(&stmt,&fn,%eval(&i-&fn+1));
  %end;

  /*---check REPEATED options and extract repeated effects list---*/
  %if %index(&first,REPEATED) %then %do;
    %let i = %index(&stmt,/);
    %if &i = 0 %then %let i = %length(&stmt);
    %else %do;
      %let repopt = %substr(&stmt,&i);
      %let j = %index(&repopt,EXP);
      %if &j ne 0 %then %do;
        %let k = %index(&repopt,EXP);
        %let repexp = %bquote(%substr(&repopt,&k));
        %let k1 = %index(&repexp,%str(%());
        %let k1 = %eval(&k1 + 1);
        %let repexp1 = %bquote(%substr(&repexp,&k1));
        %let k2 = %index(&repexp1,%str(%)));
        %let replst = &replst %substr(&repexp,&k1,
          %eval(&k2-1));
      %end;
      %let j = %index(&repopt,TYPE=SP);
      %if &j ne 0 %then %do;
        %let k = %index(&repopt,TYPE=SP);
        %let repexp = %bquote(%substr(&repopt,&k));
        %let k1 = %eval(%index(&repexp,%str(%))) + 1);
        %let repexp = %bquote(%substr(&repexp,&k1));
        %let k2 = %eval(%index(&repexp, %str(%() + 1);
        %let k3 = %eval(%index(&repexp, %str(%))) - 1);
        %let replst = &replst %substr(&repexp,&k2,
          %eval(&k3-&k2+1));
      %end;
    %end;
    %let j = %eval(&i-&fn);
    %if &j > 0 %then
```

```
      %let replst = &replst %substr(&stmt,&fn,&j);
    %end;

  /*---check ESTIMATE and LSMEANS options---*/
  %if %index(&first,ESTIMATE) and
     (%index(&stmt,CL) or %index(&stmt,ALPHA)) %then %do;
     %let intopt = &intopt ESTCL;
  %end;
  %if %index(&first,LSMEANS) %then %do;
     %if %index(&stmt,CL) or %index(&stmt,ALPHA) %then %do;
        %let intopt = &intopt LSMCL;
     %end;
     %if %index(&stmt,DIFF) or %index(&stmt,ADJ) %then %do;
        %let intopt = &intopt LSMDIFF;
     %end;
     %if %index(&stmt,SLICE) %then %do;
        %let intopt = &intopt LSMSLICE;
     %end;
  %end;

  /*---save statements---*/
  %if %index(&first,CLASS) %then %do;
     %let class = %qsubstr(&stmt,&fn);
  %end;
  %else %if %index(&first,MODEL) %then %do;
     %let model = %qsubstr(&stmt,&fn);
  %end;
  %else %if %index(&first,ID) %then %do;
     %let id = %qsubstr(&stmt,&fn);
  %end;
  %else %if %index(&first,PARMS) %then %do;
     %let parms = %qsubstr(&stmt,&fn);
  %end;
  %else %let spec = &spec &stmt %str(;) ;

  %let iv = %eval(&iv + 1);
%end;

/*---get response, model specification, and model options---*/
%let response = %scan(&model,1,=);
%let eqidx = %eval(%index(&model,=)+1);
%if (&eqidx > %length(&model)) %then %let mdl = %str();
%else %let mdl = %str( ) %qsubstr(&model,&eqidx);
%if %index(&mdl,/) %then %do;
  %let mdlspec = %qscan(&mdl,1,/);
  %let mdlopt = / %qscan(&mdl,2,/);
  %if %index(&mdlopt,%str( S )) or %index(&mdlopt,SOLUTION) or
     %index(&mdlopt,CL) or %index(&mdlopt,ALPHA) %then %do;
     %let intopt = &intopt SOLNF;
  %end;
  %if %index(&options,HYBRID) %then %do;
     %let mdlopt = &mdlopt S OUTPM=_predm OUTP=_pred;
  %end;
  %else %if %index(&options,MQL) %then %do;
     %let mdlopt = &mdlopt S OUTPM=_pred;
  %end;
```

103

```
    %else %do;
      %let mdlopt = &mdlopt S OUTP=_pred;
    %end;
  %end;
  %else %do;
    %let mdlspec = &mdl;
    %if %index(&options,HYBRID) %then %do;
      %let mdlopt = / S OUTPM=_predm OUTP=_pred;
    %end;
    %else %if %index(&options,MQL) %then %do;
      %let mdlopt = / S OUTPM=_pred;
    %end;
    %else %do;
      %let mdlopt = / S OUTP=_pred;
    %end;
  %end;
  %let mdlopt = &mdlopt alphap = &outalpha;

  /*---add an @ sign if it is missing---*/
  %if %index(&mdlspec,|) %then %do;
    %let mdl=&mdlspec;
    %let mdlspec=;
    %let i=1;
    %do %while(%length(%scan(&mdl,&i,' ')));
      %let mdlterm = %scan(&mdl,&i,' ');
      %if %index(&mdlterm,|) and %index(&mdlterm,@)=0 %then %do;
        %let j=1;
        %do %while(%length(%scan(&mdlterm,&j,|)));
         %let j = %eval(&j+1);
        %end;
        %let atvalue = %eval(&j-1);
        %let mdlterm = &mdlterm.@&atvalue;
      %end;
      %let mdlspec = &mdlspec &mdlterm;
      %let i = %eval(&i+1);
    %end;
  %end;

  /*---get parms specification, and parms options---*/
  %let parmopt2 = ;
  %if %index(&parms,/) %then %do;
    %let parmspec = %scan(&parms,1,/);
    %let parmopt = / %scan(&parms,2,/);
    %if not %index(&options,NOPREV) %then %do;
      %let parmopt2 = %scan(&parms,2,/);
      %let parmopt2 = %remove(%quote(&parmopt2),PARMSDATA);
      %let parmopt2 = %remove(%quote(&parmopt2),PDATA);
      %let parmopt2 = %remove(%quote(&parmopt2),OLS);
      %let parmopt2 = %remove(%quote(&parmopt2),RATIOS);
    %end;
  %end;
  %else %do;
    %let parmspec = %qupcase(&parms);
    %let parmopt  = ;
  %end;
```

```
%if %length(&linkn) %then %let intopt = &intopt NLIN;

/*---create local variables---*/
%local varlst errorfn linkfn varform devform etaform detaform
   muform dmuform nlinform nlinder deviance scale n nb nu ns nus
   crit conv cf intopt iter;

/*---get variable list and trim it---*/
%mvarlst
%trimlst(varlst,&varlst);

/*---set error and link function macro variables---*/
%errlink

/*---print header---*/
%if not %index(&options,NOPRINT) %then %do;
   %if %index(&data,.)=0 %then %let data=WORK.&data;
   %put;
   %put %str(     ) The GLIMMIX Macro;
   %put;
   %put Data Set          : &data;
   %put Error Distribution : &errorfn;
   %put Link Function     : &linkfn;
   %put Response Variable  : &response;
   %if %length(&weight) %then
     %put Weight           : &weight;
   %if %length(&freq) %then
     %put Frequency        : &freq;
   %put;
   %put;
%end;

/*---initialize iteration starting values---*/
%init

/*---turn off printing---*/
%if not %index(&options,PRINTALL) %then %do;
   ods exclude all;
%end;

/*---run first estimates for convergence tests---*/
%mixed

/*---iterate until convergence---*/
%iterate

/*---turn the printing back on---*/
%if not %index(&options,PRINTALL) %then %do;
   ods select all;
%end;

/*---final Proc Mixed run---*/
%if not %index(&options,NOPRINT) %then %do;
   %put;
   %put Output from final Proc Mixed run:;
   %mixed
```

```
%end;

/*---stop if did not converge---*/
%if &conv ne 1 %then %do;
   %if not %index(&options,NOPRINT) %then %do;
      %put GLIMMIX did not converge.;
   %end;
%end;
/*---otherwise compile and print results---*/
%else %do;
   %compile
   %if not %index(&options,NOPRINT) %then %do;
      %printout
   %end;
   /*---create output data set---*/
   %if %length(&outfile) %then %do;
      %outinfo(%quote(&outfile));
   %end;
%end;

%exit:;
%if not %index(&options,NOTES) %then %do;
   options notes date number;
%end;

/*---title cleanup---*/
&titlen;
%if %length(&titlesas) %then %do;
   title "The SAS System";
%end;
%else %if &ntitle=10 %then title10 &title10;

%mend glimmix;
```

End: GLIMMIX-Macro