DERIVATION OF RED SNAPPER TIME SERIES FROM SEAMAP AND GROUNDFISH TRAWL SURVEYS

Scott Nichols
NMFS Pascagoula

Introduction:

This document provides an entry point for the Gulf of Mexico Groundfish and SEAMAP trawl survey data. The content here emphasizes definition of variables, survey designs, and a few analytical implications of historical practices. Users more interested in the nuts and bolts of field procedures should consult the SEAMAP Field Operations Manual, available at this workshop.

Although the data available should be useful for many applications, this document focuses on the characterization of interannual variations of year class strength, in particular explaining the actual survey series that have been extracted from the larger base of multipurpose cruises. Although we can make the data available in most any form (the entire base is now being transferred to Oracle), this document tends to make reference to SAS variables and programming. SAS is the primary analytical software for most of the SEAMAP partners. For the red snapper SEDAR, SAS data files at two different levels of summarization will be available.

Results for some 'Base Indexes,' using some very simple analytical structures are included, but these simple indexes take no account of missing stations, nor are they automatically comparable across the separate time series. The indexes actually recommended for use in stock assessment are derived in a companion paper (Nichols 2004), using the same data introduced here.

History:

Pascagoula has conducted trawl surveys since the mid-1950s, but the 'Modern Era' actually began in 1972. At that time, the laboratory began surveys in the north central Gulf of Mexico to quantify seasonal, spatial, and interannual variations of groundfish resources. Differing from all previous work, these surveys used a probability-based (random) sampling design, and all portions of the catch were worked up and recorded. (Prior to 1972, survey designs were 'exploratory,' and only the species of interest at the time might be recorded.) Groundfish surveys made at one time or another have covered most of US Gulf of Mexico, but only a small area of the north central Gulf (88° to 91°30′, 5-50 fm; named the 'Primary Area') was covered consistently over years. Surveys were conducted in some years in winter, spring, and summer, but the fall season was selected as the main season to monitor interannual variation. Sampling in the October November period has been conducted every year since 1972.

A second survey program based on random sampling began in 1982 with the establishment of SEAMAP, a state-federal partnership to coordinate fishery research vessel activity in the region. SEAMAP's first survey project was a summer trawl survey from Pensacola to Brownsville, 5-50 fm, using a design adapted from a survey to assess the Texas Closure shrimp management measure in 1981. The SEAMAP summer trawl survey continues today.

Geographic coverage in the fall groundfish program was expanded tentatively in 1985 and 1986. In 1987, the SEAMAP design was adopted in the fall as well as summer, unifying the programs.

Principles:

There are many applications of resource surveys data, but one goal stands as the most vital to our mission, and also requires the most rigor in sampling design: provide information on year to year variations in abundance of the living marine resources of the southeast. We expand that goal into providing interannual information for as many stocks as possible, at two points in the life history of each stock: (pre-)recruits, and adult spawning stock. There are five principles we strive for in the development of our surveys:

1) Synoptic
2) Stock-wide
3) Well-defined sampling universe
4) Useful Precision
5) Control Bias

Obviously, no survey can meet all five principles completely; some trade-offs are necessary.

In the southeast, trawl surveys are very effective for many FMP species, but generally only on new recruits. Other techniques are required for indexing spawning stocks. These other survey techniques are not covered in this document.

In the Gulf of Mexico, where peak shrimping F approaches 1.0 per month, 'synoptic' for trawl surveys is always compromised – differences shrimp catch rates are detectable statistically even over about 2 weeks. The stock-wide principle is easier to approach – ranges for the stocks of many species are almost contained within the SEAMAP survey area. However, some stocks do extend outside the survey area, either in depth, into Mexico, or into the essentially untrawlable areas to the east. (The small size of the 'Primary Area' of the groundfish program compared to the ranges of the stocks of interest was the main reason the program was replaced by SEAMAP.) For stock assessment, the distribution of a stock over space can be considered an annoying impediment – one must integrate over the spatial distribution to extract information at the stock level. A well-defined sampling universe addresses this, making each sample 'representative' of some known area, allowing integration over space. Precision is usually judged by the coefficient of variation (CV) for the mean – defined as standard error over the mean. (CV as standard deviation over mean is more a property of nature; CV as standard error over mean becomes a property of the survey.) With dependence between variance and mean, CVs usually improve both with increased sample sizes, and increased catches. 'Control bias' is a bit of a catch-all phrase. For trawl surveys, it would include things like maintaining a geographic pattern to sample collection over years, to minimize confounding of interannual variation with the variation due to compromises in synopticity. (After all, a truly random sampling design would occupy stations in random order – no one can do that at $10k per day, and synopticity would be totally hopeless.)

The Database:

We are just now completing some significant changes to our database. Previously, the organization of historical data was totally fragmented. Separate years existed on separate tapes; often with different formats and codings. Potential users had to contact the data manager, and a time-consuming assembly of data from requested cruises was begun. Automated editing checks were almost non-existent. We are now combining all data to a single ORACLE base with consistent structure and coding, which will be addressable directly by users. All our cruises are multi-purpose. Previous versions of the data contained little or no information about which data belonged to which projects. We are now adding identifiers to individual sampling events, indicating which data belong to which projects, and thus should be analyzed as a unit. Stratum designations and weightings (usually spatial areas) are being attached. An expanded list of editing checks is being applied. As these enhancements are completed, key projects are being identified through SQL-plus logic into ORACLE 'Views,' saving potential users considerable time (and uncertainty) in screening the data to extract observations that belong to a particular project. At present, we have nearly completed these enhancements for two of the most significant sets of trawling data: the Fall portion of the Groundish Program, and the SEAMAP Trawl Surveys Program. The data for red snapper from stations specifically identified as part of the designed abundance surveys have been extracted for this report, and will be available to the SEDAR.

Continuity:

One of the most powerful advantages of 'fishery independent' (research vessel) surveys is the ability to keep things constant from year to year – in principle. In practice, there is always some year to year variation in survey capability. Sometimes, available dollars dictated changes. Some years, bad weather or vessel malfunctions compromised coverage. Vessels do change uncontrollably over time – the Oregon II got increasingly noisy (easily noticed, but never quantified) until some major repairs several years ago. We would have been foolish not to have added more modern navigational equipment as advances became available. For example, we have gone from no direct knowledge of vessel speed, through a series of improvements, to very precise determination now based on GPS capability. Vessel operators change. Fishing skills of an individual operator may change over time. Some of the largest changes have been changes in design, necessary to move the survey program closer to the basic principles supporting stock assessment objectives.

When all anomalies to 'identical' surveys are listed, one may start to despair, but as a practical matter, many of the changes likely had only minor impact, and really need to be listed only for rigor in documentation. For some other changes, it will be possible to derive appropriate analytical adjustments. What I have done is to identify 'time series' – sequences of years for which a proportionality between a mean CPUE statistic and actual abundance could be reasonably treated as constant for all years within the series. (Assuming, of course, that even a varying proportionality might exist in general). Some variations do exist within time series (as documented below), but the time series identified are intended to be conservative – remaining variations among years within time series should be of small impact, whereas variations among time series could be more substantial. I present 4 multi-year time series:

Fall Groundfish	(FG)	1972-1986
Early SEAMAP	(ES)	1982-1986
Summer SEAMAP	(SS)	1987-ongoing
Fall SEAMAP	(FS)	1988-ongoing;

then extend the 'time series' concept to 4 single-year samplings that might be attachable to the main time series analytically.

Data analysis can then proceed in two directions. 1) Simple indexes consistent with the survey design can be calculated for each time series. These indexes will be suitable as candidates for assessment tuning, which should return an estimate of proportionality (q) between index and estimated abundance. 2) Statistical modeling approaches can be developed to link the separate time series together to obtain indexes of longer history. These model-based indexes become alternative tuning candidates. There are pro's and con's to each approach. It should also be possible to compare statistical model performance with q's returned from tunings using the shorter time series, evaluating for consistency among approaches. Of course, modeling approaches may also be explored to increase analytical rigor within single time series.

Groundfish Trawl Survey Design:

A semiballon shrimp trawl with a 42 ft headrope was the selected as the standard gear for the groundfish design. For historical reasons, the gear is often referred to (and coded) as the '40 ft trawl.' A complete description of the standard gear is available from the Pascagoula lab.

With exception of a special case in 1986, all groundfish-design sampling currently included in our database was conducted from the NOAA research vessel Oregon II. There was some non-continuous usage of other vessels in peripheral areas, not suitable for time series derivation. These data from other vessels may be added to our data base in the future, but they are unlikely to have any stock assessment application.

The potential survey area was divided into 3 regions: a 'Primary Area' between 88° and 91° 30′ longitude, and two 'Secondary Areas' of indefinite extent on either side. The range of the area to be worked differed among cruises. Most cruises included coverage of the Primary Area. (Only the samples from the primary area are included in the fall groundfish (FG) time series. Samples from other areas are treated as separate 'projects' in the database.)

The nominal survey area has been described as 5 and 50 fm. This is an approximation, however. The actual sampling universe is based on latitude and longitude, and includes depths that extend slightly outside this range in both directions. The 5 fm depth was minimum that the NOAA vessel Oregon II was usually allowed to work, for safety reasons. The 50 fm depth was the expected outer limit of significant catches of the sciaenids and other target species of groundfish. The 50 fm depth also appears to be a practical outer limit of the range for shrimp, and for a good many coastal finfish species not generally considered part of the groundfish complex.

Once the area to be surveyed was decided, sampling sites were selected: a 10 minute latitude by 10 minute longitude 'block' was selected at random, with replacement. Within each 10 by 10 minute block, a 2.5 minute latitude by 2.5 minute longitude 'grid' was selected at random. A few grids were excluded because the area was known to be untrawlable. This 'block/grid' procedure results in a completely randomized selection of 'grids,' without replacement. Three replicate, 10 minute tows (minutes of time) were to be taken at each site. Time of day was not considered – each station was to be occupied whenever the vessel reached the site. The total number of grids selected was adjusted to the number of days available for each survey. Stations were selected independently for each cruise.

In the early years, stations were selected by hand using a random number table (table source unknown). A complete set of potential stations was considered, and if judged to be too operationally difficult, or too time consuming, or too unbalanced with respect to area and depth, stations were reselected until the potential cruise track was considered viable. In the early 1970s, this judgement was apparently made by eye. The small amount of written material from the time seems to imply a complete reselection if a cruise track were rejected (as one would want), but the amount of labor required makes this seem unlikely. In some years (which years, now unknown), the alongshore and depth distributions of the randomly selected sites were examined more formally. If a Chi-square test indicated that the number of samples in alongshore and depth divisions departed from the number expected from completely randomized selection, the station list was rejected, and a new draw was made. Starting in 1978, a computerized station selection program (known locally as the 'travelling salesman program') was implemented (Leming and Holley 1978). This program again randomly selected blocks and grids, and suggested an optimal cruise track through the selected sampling sites. The program was customarily run about 50 times, and the runs with the 10 shortest cruise tracks were identified. A subjective choice was made among those 10 shortest tracks for the particular set of stations to be sampled that cruise. Leming and Holley (1978) report a list of grids included in the sampling universe, but it does not match the distribution of the data exactly. Apparently, a handful of grids were added or removed over the years (when, unknown) due to expected difficulty in trawling. Leming and Holley also mention excluding some areas of 'known low groundfish densities,' but these exclusions are not really evident, as their list is fairly continuous in space. The best list we could derive (based on the union of the Leming and Holley listing, and the block/grids in the data records after our editing) contains about 2% more grids in the Primary Area than the Leming and Holley listing. There were a handful of grids that never appeared in any block/grid-based cruise, but did appear in the Leming and Holly listing. I recommend our current list (available in the database as the Primary_Area_Universe table) be used for all analyses for the Primary Area for all years.

In operational practice, triplicate 10 minute tows were not always taken. How many tows were actually taken depended on how much time was left in the survey. Surveys generally progressed east to west, so the 1- and 2-tow stations, and missed stations, were more common in the western part of the survey area. There were a few instances where more than 3 tows were taken, due to misunderstanding by the vessel operators. These tows have been retained as part of the time series. In 1976, the cruise report stated that the number of tows per grid was reduced while at sea if the first tow produced a low total catch. This could cause a bias. The potential for bias was apparently soon recognized, as that procedure was not mentioned in any earlier or later cruise reports, and indications of decisions determined by catch are not obvious in the data.

Tows targeted to 15 minutes were sometimes made in deeper water, or in areas where catches were expected to be low. Tow times recorded sometimes varied a few minutes from the 10 or 15 minute targets. (There is even one 30 minute tow in the base that appears to a standard groundfish sample in all other respects; it has been retained as a legitimate observation.)

Choice of trawling direction within grids was left to the vessel operator, who was presumed to base his choice on wind and sea conditions, and on the direction to the next sampling site. Most trawls were probably conducted roughly parallel to the depth contours, which is the custom in the local fishery.

For convenience, station selection was actually done separately for each scheduled leg of the cruise. Usually, this meant separate selections for the Primary Area east and west of 89°30′. From the results, it appears that some care was taken to keep the sampling fraction nearly the same in both areas (maintaining a completely randomized design overall), but as cruises proceeded east to west, any operational problems that developed during a cruise preferentially impacted the western part of the Primary Area. (Chi-square analysis of station distributions detected some depletion in the west when all cruises were included in a single analysis, but little evidence of structure beyond that finding. Analysis courtesy of S. Diamond, Texas Tech). I recommend dividing the Primary Area into Eastern and Western strata (at 89°30′), thus using the number of samples in each stratum independently in the calculations.

There is one description of groundfish survey design extant (Kemmerer *et al.* 1982) that has some conflicts with the descriptions given in this paper. The Kemmerer *et al.* paper suggests some depth stratification in the design, and a 2 minute instead of 2.5 minute grid dimension. It is clear from the data that the Kemmerer *et al.* paper is not accurate on these issues. Actually, the block/grid design is not very amenable to depth stratification. Some individual grids have surprisingly large depth ranges, as determined by accumulation of individual trawl records over time. We have examined the records for grids with the largest ranges (over 6 fm). Only a very few were found to be due to erroneous entry (since corrected); we have no reason to doubt most of the wide ranges reported. Most appear in areas where large gradients, holes, or channels could be expected. As a practical matter, for any analysis looking at variation over depth, it would usually be better to use the depth measure recorded for each tow individually, rather than assign a depth value to a grid.

Actual spatial areas in latitude and longitude grids are not constant over change in latitude. Within the Primary Area, this effect is small (about a 2% difference between smallest and largest) and can probably be ignored. The variations in grid size increases to about 4% for cruises extending into Florida or Texas. This still seems too small to bother with for most applications, but we have included calculated values of area (hectares) in the data records.

Secondary area sampling usually depended on time available after Primary Area sampling was completed, and there are major differences in areal extent and density of sampling (both intended, and realized) among years. There were also exceptions to the block/grid selection in some years outside the primary areas. For example, a transect station patterns were clearly used in the western Secondary Area one year. As the Secondary Area data really do not form a continuous time series, we have not attempted to distinguish design differences among years, and have lumped various types of Secondary Area sampling in this data base, separating only by location. Secondary Area samplings are not included in the Fall Groundfish data presented here.

In 1985, an attempt was made to extend the groundfish survey throughout the Gulf, 5-100 fm, using the block/grid design. Replicates tows within a grid were eliminated (i.e. one 15-minute tow per grid). There was also an attempt to repeat stations day vs night. Operations were highly disrupted by hurricanes, so only the northern gulf was covered, and the day /night repeats were incomplete. In 1986, the day / night repeat strategy was abandoned, and sampling did take place throughout the northern and western Gulf. There was also some participation by other SEAMAP partners in the 1986 survey. Because the design within the Primary Area was largely the same as previous surveys, Primary Area stations for 1985 and 1986 have been included in the Fall Groundfish time series.

The Primary Area stations for the Fall Groundfish Survey, and the red snapper catch data, are collected in the SAS file FGCATCH made available on the SEDAR workshop LAN. Data are for 1972-1986. Two cruise numbers occur for 1972, with some sampling occuring in September that year. All other years except 1986 have only a single cruise number, with all sampling in the October / November time period. Table 1 summarizes the each year's collection. In 1987, the SEAMAP design replaced the groundfish design. Data under the SEAMAP design are treated as separate time series, described below.

Many other trawl projects were conducted aboard fall groundfish survey cruises. None have the consistency, continuity, or random design to qualify as fishery-independent time series. The data for trawls conducted for other purposes are available in our database, but are not included in the data for the fall groundfish time series submitted to this workshop.

SEAMAP Trawl Survey Design:

The SEAMAP survey design had roots in a 1981 summer survey of the Texas coast conducted to evaluate the Texas Closure shrimp management measure, the first management action implemented by the Gulf of Mexico Fishery Management Council. This survey adopted the same standard gear as the Groundfish program. The design used a very fine structure of stratification in depth, and implemented cross-stratum trawling. Towing speed was set at 3 knots. The survey area was expanded in 1982 as the first survey of the new SEAMAP program, covering the area between Pensacola and the US / Mexico border, 5-50 fm. The 1982 survey was the first to incorporate cooperating state and university vessels, as well as the NOAA vessel Oregon II. Summer surveys have been conducted every year since 1982. The SEAMAP design was adopted for the Fall trawl survey beginning in 1987. Within the SEAMAP design, there have been some significant variations, so several separate time series are described.

The 1981 Texas Closure survey was developed after examining data collected in the 1960s from a systematic design aboard the chartered vessel Gus III. Maximizing information on the abundance, size composition and distribution of shrimp was the original goal. Depth was seen to be the primary 'organizing' gradient in the community, with rapid and predictable variations of catch rate in space expected as depth changed. Regular alongshore changes were detectable, but the rate of change with distance was much slower than the rate of change with depth. Day / night differences were profound for some species, but not others. We suspected that trawling alongshore (the usual fishery, and survey, practice) would enhance the effect of patchiness compared to trawling parallel to the depth gradient. In summary, we wanted to integrate over depth, but retain information with depth for management alternatives. The solution chosen was to stratify in depth far beyond what was justifiable for precision gain (to 1 fm widths in nearshore areas), and trawl from the inner depth boundary to the outer depth boundary in each stratum (at the cost of accepting variable tow times). Still concerned about highly variable tow times, we did impose minimum and maximum times for individual tows, but towing in a stratum was to continue (by immediately starting a new haul) until the stratum was crossed. There were no alongshore divisions of strata in 1981 – the entire Texas coast was one stratum. In 1981, sampling was conducted only at night, as catch efficiency for brown shrimp is much higher at night. Number of stations per stratum were assigned approximately as a Neyman allocation based on variance of shrimp catch rates with depth extracted from the Gus III surveys, ignoring the systematic nature of the Gus III surveys. Some strata were allocated only a single sample. Red snapper data from this Texas Closure project are available in the TCCATCH file. Only Oregon II data are included in these views. There was some sampling conducted by other vessels, but their activities were not part of the original design, and in any case are not available within our database. With only the Texas coast covered, the range of the survey was much smaller than the range of most stocks, so it would take a major assumption about constancy of distribution to attach this year to stockwide abundance indexes of succeeding years.

In 1982, the survey was extended throughout the northern and western Gulf, in the June / July period. With this extended area, an alongshore component to stratification was added, based on the commercial shrimp statistical zones (generally 2 zones per stratum). Sampling continued only at night, primarily due to lack of funds. As per the philosophy underlying SEAMAP, several state-operated vessels joined in the sampling, and a program of side by side trawling for calibration was started. Sampling began each season off Mississippi and Alabama, then continued as a counterclockwise sweep through the Gulf. After this sweep, an additional set of samples was allocated to the Mississippi / Alabama area, but any operational problems during the cruises tended to cause this second sampling in the eastern zone to be cut short. (Because this second sampling was often incomplete, while coverage during early sampling was nearly complete, the second sampling was not included as part of the standard time series.) Planned number of stations per stratum was set at two in 1982, and reduced to one upon subdivision of alongshore strata into single shrimp statistical zones in 1983. Enough strata were missed in 1983 that the strata were recast to the broader 1982 alongshore strata after the fact. Sometimes, state vessels were assigned stations within strata sampled by the Oregon II for 'enhancement;' at other times stations were divided up among available vessels to try to complete the coverage. (The approach varied depending on the sea days available from the participating vessels.) Sampling under this strategy took place from 1982-1986. This time series is designated ES, for "Early SEAMAP," and red snapper data are available in the file ESCATCH.

We quickly became concerned about the duration of the survey efforts – the principle of synopticity over such a large area with active shrimping. In 1984, an additional 'clockwise sweep' was conducted after the then-standard counterclockwise sweep to south Texas, to determine how quickly detectable differences arose. There was a clear progression of difference in shrimp catch rates with temporal separation around the Gulf. Red snapper data from this second sweep in 1984 are available in the C2CATCH file. We considered ways to compress the sampling in time, by combining some strata, and cutting the number of stations per stratum, basically reducing to one sample per stratum by 1987. We also began to encounter problems with completing the surveying in Texas by the time of the Texas opening (a requirement of the original Texas Closure survey, and still an important goal under SEAMAP), as the State of Texas began changing the opening dates from year to year. So by 1987, we had several reasons to modify the design. Sampling direction was changed, first starting off Alabama and Mississippi as before, then proceeding immediately to south Texas, beginning a clockwise sweep. As before, a second sampling off Mississippi and Alabama was scheduled for the end of the survey period, but was often incomplete. First sweep sampling east of the river was more complete than the second, so the first ('early') sweep was taken as the standard survey. Because early coverage was not as complete after 1987 as it was before, samples from the late sweep were included, but only for strata missed during the early sweep. This convention improves coverage, at some cost to synopticity. Sampling around the clock became possible financially, so sampling was doubled, with day vs night added as an additional stratification. This time series is designated SS, for "Summer SEAMAP." Red snapper data, from 1987 on, are available in the file SSCATCH. This survey continues.

In 1987, we decided to include the Fall survey under the SEAMAP umbrella, and adopted the SEAMAP design. We started with a counterclockwise sweep in 1987, but ran into severe weather problems, resulting in incomplete coverage. We realized we could reduce weather problems by sampling off Texas first, so we switched to a clockwise sweep in 1988, which continues. Because of the sweep difference, and the incomplete coverage, the 1987 survey was isolated here as FF, for "First Fall," and red snapper data are available in the file FFCATCH. Surveys from 1988 on are designated FS, for "Fall SEAMAP," and red snapper data are collected in FSCATCH.

The stratification structure of SEAMAP design has been varied slightly over the years, mostly to try to improve synopticity, but the total geographic area (sampling universe) targeted has remained the same, with one exception. Due to the accumulation of artificial reefs in statistical zone 10, we had to abandon trawling there in 1989, part way into the SS and FS time series. By convention, the east-of-river alongshore strata are considered to be 12-10 in the ES time series, and 12-11 throughout the SS and FS time series, even though some samples actually came from stat zone 10 in the earlier years of SS and FS. SS, FF, and FS strata are assigned weights based on spatial area excluding zone 10, but samples from zone

10 in these time series are considered representative of strata based on 12-11, even though they lie outside the 12-11 area. Starting in Fall 1989, we no longer selected stations east of 88° W.

The original intent was to use 'the method of collapsed strata' (collapsing adjacent strata together) to obtain multiple samples per stratum for variance estimates. We soon realized that the approach could be generalized to collapsing all strata, effectively treating the SEAMAP design as if it were a completely randomized design (CRD), but with samples weighted by the spatial area of the strata they represented. This approach is computationally simple, and avoids the ambiguity of just which strata to collapse. The probable cost is increased variance: because the stratified selection is more evenly spaced than a true CRD, variance estimates for most species (usually patchy in space) would be higher than with a true CRD. Of course, the main intent of such a fine stratification in the first place was the guarantee of coverage in alongshore and depth space, and not a gain in precision.

Summary of the Time Series:

The following is a summarization of the coverages, and major differences among time series presented at the SEDAR. The tabulation also includes anomalies and variations among years. These variations were considered important enough to document, but not significant enough to warrant splitting a time series into two or more separate series.

Fall Season

FG ('Fall Groundfish'). Cruise dates, and number of samples are collected in Table 1.

1972-1986
October / November
'Primary Area' covered each year
Groundfish block/grid design
Stations generally worked east to west in primary area
Nominally 3 tows per grid; reduced when time became short
Numbers of grids drawn for sampling were set by number of sea days available
Nominally 10 minute tows; some set to 15 min in deeper areas, or where low catches expected
Time of day ignored -- stations worked upon arrival along shortest path
Speed nominally 2.5 knots, but based on expected relationship with propeller pitch in early years
Oregon II only, except 1986
1972: conducted as 2 cruises; some repeats of grids between cruises. Sampling began 9/26.
1976: reduction of tows per grid allowed if first catch was low
1978: automated station selection program begun
1980: some tows of 5-9 minutes duration reported (retained)
1985: dropped to single tow per grid, attempt day/night repeats, expanded coverage, bad weather
1986: single tow per grid, expanded coverage, dropped day/night repeats, multiple vessels

FF ('First Fall' SEAMAP)

Single year, 1987
SEAMAP design in counterclockwise pattern
Nominal speed 3 knots
Weather disrupted, poor coverage
Multiple vessels

FS ('Fall SEAMAP'). Cruise dates, and % of strata covered are collected in Table 2.

1988- on (currently, data through 2002 are available)
SEAMAP design in clockwise pattern

One station per stratum
Nominal speed 3 knots
Multiple vessels
1989 on: no sampling in shrimp stat zone 10.

Summer Season

TC ('Texas Closure')

Single year, 1981
Texas coast only
Night sampling only
Depth stratified random design
Variable number of stations per stratum
Cross stratum trawling
Nominal speed 3 knots
Oregon II only
Only 1 fathom of deeper, wider strata covered, due to operator misunderstanding

ES ('Early SEAMAP'). Cruise dates and number of strata sampled are collected in Table 3.

1982-1986
SEAMAP design in counterclockwise pattern
Night sampling only
Variable number of stations per stratum
Nominal speed 3 knots
Multiple vessels
1982: one vessel failed to complete catch work-ups; those data dropped, with loss of coverage

C2 ('Collection 2')

1984 only
SEAMAP design in clockwise pattern
Night only
Variable number stations per stratum
Nominal speed 3 knots
Multiple vessels
Some ES stations picked up on the C2 sweep (retained in ES)

SS ('Summer SEAMAP'). Cruise dates and % of strata covered are collected in Table 4.

1987 – on (data currently available through 2003)
SEAMAP design in clockwise pattern
Day and night, as separate strata
Single station per stratum (some exceptions)
Nominal speed 3 knots
Multiple vessels
1989: poorer coverage than most other years
1990 on: no sampling in stat zone 10

There were numerous other trawling projects conducted aboard the cruises that produced these time series. Closely related projects are summarized in Appendix 1. Discussion is sequestered in the appendix, because information about the individual variables (next section) is required to explain these other projects. I did not bring data from these other projects to the SEDAR. I identified these projects here only because some participants may be aware of them, and wonder if they were included in the time series or not. They are not.

Description of Variables:

This section was extracted from the documentation we are developing for our overall database, and explains the variables in some detail, perhaps more than actually needed by the SEDAR. The summary data files listed in survey design sections (the xxCATCH files) do not have all of these variables, but their parent files do, and these parents will also be available should anyone want to explore something needing less aggregated data. The main differences between the xxCATCH files and their parents are that those files have been limited to red snapper, summed over all catches during a single 'event' as described below, and have explicit entries of zero for events where no red snapper were caught. The parent files available at the SEDAR have all species data from all successful trawls relevant to each survey design, with separate files for the station and biological data, and no explicit zeros when there was no catch of a particular species.

There is one point of 'usage' to be made first. The database now distinguishes between STATIONs and EVENTs in a formal manner described immediately below. The individual sampling tows are now formally referred to as EVENTs, but the term 'station' has been in common usage in this region for many years as the general term for individual hauls. Users should be alert to whether the term 'station' is being used formally, referring to a specific variable in the data base that links to all sampling in a small spatial area over a short time interval, or more casually, referring to individual trawl hauls in general. Usually, the variable STATION is referred to in caps.

Trawl Event Variables ('Station' Data)

VESSEL CRUISE STATION EVENT PASC

The first 4 are the identifier variables for each trawl haul in the files. In combination, the values are unique. PASC is short for 'Pascagoula station number,' which is discussed below. All are character variables, but at present the only characters used are numbers. VESSEL values are two digit codes. There are about 70 vessel codes in the entire data base. In the trawl survey files, only 7 vessels contributed data:

04	Oregon II	NOAA
16	Western Gulf	Texas Parks & Wildlife
17	Tommy Munro	Gulf Coast Research Laboratory (MS)
23	A. E. Verrill	Alabama DNR
24	Florence May	Texas Parks & Wildlife
26	Suncoaster	Contract to NMFS from FIO
35	Pelican	Louisiana Wildlife and Fisheries

(The agency affiliation listed was the SEAMAP participant. Some of the vessels are owned and/or operated by other organizations.) CRUISE is a 3-digit code. Some vessels build the CRUISE code out of a 2-digit value for year, and a single digit for cruise in year. Other vessels number cruises sequentially, ignoring year. CRUISE should just be used as an identifier. For cases where there is a gap in time between trips to sea, there is no consistency in the base as to what is a new cruise vs what is a continuation of an existing cruise. Cruise identifiers may not be the same as those used by the SEAMAP state partners in there own databases.

At the establishment of SEAMAP, the Pascagoula method of identifying station operations was adopted by all partners. Unfortunately, for identifying operations taken together in space or time – the usual conception of a 'station,' the Pascagoula method was ambiguous, and in fact differed among sampling programs. For example, each trawl put over the side received a unique station number (PASC in the current base; previous users of these data may have named that variable 'Station,' which is not the same as the variable STATION used here.) All pieces of plankton gear put over at a single location received the same PASC. If both plankton and trawl gear were used a site, sometimes the plankton hauls got their own PASC, sometimes they got the same PASC as the trawl gear. Environmental casts (i.e. CTD, XBT, etc) did not receive a unique PASC unless no fishing gear was deployed at the site (fairly rare in the data base). Far more commonly, data from environmental casts were assigned the same PASC as the fishing gear. If multiple gears, or multiple hauls with the same gear, were used at a station, there was no simple way to link the environmental data to the catch data except for one haul. Indeed, there was no simple way to recognize that the multiple fishing hauls came from the same conceptual 'station,' either.

We have attempted to eliminate these ambiguities by adding STATION and EVENT identifiers after the fact. The data base then becomes much easier to use for matching data from separate operations at the same site at the same time. Conceptually, each deployment of a piece of gear over the side is a unique EVENT. (We have also extended the concept of EVENT to other types of data collected – you look outside and write down cloud, precipitation, and sea state data – that's an event. You write down data from the vessel's weather instruments – that's an event.) A STATION is a collection of EVENTs with a close, meaningful proximity in time and space. The requirement of 'meaningful' implies a level of subjectivity, in that the scale of 'meaningful' might be different for different projects. Our convention here: the replicate tows in a single groundfish block/grid are considered to be 3 EVENTS at 1 STATION. Multiple SEAMAP tows to complete a single transit in cross stratum trawling are multiple EVENTs at a single STATION. Trawl hauls taken in the same spatial sampling unit, but well separated in time (e.g. separated by intervening sampling) are separate STATIONs. For a localized project like sampling at a dredge disposal site, different sampling locations around the site are assigned separate STATIONs, even though the spatial and temporal separation between STATIONs in this type of project may be smaller than the separation between EVENTs at a STATION in the large scale surveys. Sampling by paired tows (common), and identifiers for environmental data in these local projects still appear as separate EVENTs at each single STATION. What constitutes a single STATION thus depends on the project, usually in an obvious manner.

At present, STATION and EVENT have been assigned uniquely within each CRUISE, in the sequence of the original PASC values. As errors may be discovered and corrected, the sequential nature could be lost, so users should treat these variables as unique identifiers only, and not rely in any way on a sequential relationship. With both STATION and EVENT assigned uniquely over each cruise, there is some intentional redundance when using both STATION and EVENT in an identifier – often, EVENT could be used alone. However, we recommend using the complete identifier VESSEL CRUISE STATION EVENT in all logic – it will make programming more consistent, and the collection to the station level will be there automatically, if needed. For NOAA vessels, PASC was almost always assigned sequentially in time, although there are exceptions that have been examined and found to be correct. For state vessels, PASC assignments were generally not time sequences. In the earlier cruises aboard NOAA vessels, the PASC sequence was continued over cruises. In later years, assignment switched to making PASC unique only within cruises. In state collected data, the value for PASC was usually constructed in the form of the 2-digit vessel code and 3 digits for 'station' in some sequence (sometimes a spatial sequence, sometimes temporal, sometimes unknown). For many state collections, values for PASC appear to have been added after the fact, and thus may not be available on original state data sheets, or in separate data bases kept by the states.

STATION and EVENT values were forced to fill their 4 and 5 character fields by starting each cruise with 1001 and 10001, respectively. Most but not all PASC values in the data base fill the 5 character field.

The point of this extended discussion is that the additions of STATION and EVENT are major changes for users familiar with previous versions of these data. For users that have existing programs that use the identifier set VESSEL CRUISE PASC, that identifier will still work the same way it did with previous versions of this data base. However, users should recheck their old logic to make sure the VESSEL CRUISE PASC identifier actually works the way intended. Logic that worked for some analyses is vulnerable when the data subset under consideration is expanded to include data not using the same conventions that a user's original programming logic is counting on. We strongly recommend converting to logic based on a VESSEL CRUISE STATION EVENT identifier.

PURPOSE

PURPOSE is a 2-character code that identifies membership of a data record to some particular project. In the underlying tables of the data base, PURPOSES exists as a separate table, with unique values for VESSEL CRUISE STATION EVENT PURPOSE. Thus an EVENT may have more than one PURPOSE (i.e. belong to more than one project). PURPOSE values use the same codes as the PROJECT variable, which identifies projects at the cruise level in the Projects table in the underlying base. Within the context of trawl survey views, only 8 codes are relevant (the same as the 2-character abbreviations in the view names). A few more relevant codes are explained in the 'Closely Related Trawling Projects' appendix to this document. A complete list of codes is collected in the Oracle table Project_Descriptions, which currently includes only the project code and a brief project name.

Addition of the PURPOSE variable is also a major change for users familiar with previous versions of these data. There was nothing similar in past versions, and often no consistent logic using other variables that would allow easy extraction of data by project. PURPOSE has been added only with considerable labor. Some assignments could be automated, for subsets of the data where consistent conventions were used among other variables; but for many cruises, project determination meant checking cruise reports, other written documents, data sheets, etc, and assigning project membership 'by hand.' In reality, once all the information was examined one cruise at a time, intent was cleanly resolvable in most cases, with very little ambiguity left over. (There are a few records scattered in the data base coded as unknown, undocumented, or unusable, that probably should not be included in <u>any</u> analysis.) We are now very confident in the accuracy of the assignments of data records to projects. We will now be able work at a much higher degree of rigor than was possible in the past.

Use of the PURPOSE variable is absolutely crucial for any analysis that needs to make use of the probability-based sampling of the designed surveys. Oracle views give users a very simple way of making use of the PURPOSE designation. For projects for which views have been constructed, all data in the view should be included, all other data in the base should be excluded. Data in the time series files for this SEDAR were selected by (and limited to) the appropriate PURPOSE designations, and were in fact constructed from Oracle views.

YR BLKGRD STRATUM WATE

These variables are the structural variables for the survey designs. The variable BLKGRD is not included in any of the files based on the SEAMAP design (it has no application, and thus is always null in the underlying tables for SEAMAP cruise data). STRATUM was the primary structural variable in the SEAMAP designs. However, it became clear that another structural designator would be useful in groundfish design cruises, so for convenience , the variable STRATUM was given a second duty for those cruises, and the variable does appear in the groundfish views.

YR is a character variable made of a 4-digit year value. The only anomaly in the base is one survey (part of the LW project, not part of the times series of this paper; see Appendix), customarily sampled in December, that happened to be conducted in January. By convention, YR was assigned a value for the year preceding the January sampling. (The same convention was used in the SEAMAP stratum designation for that cruise. Year is immediately extractable from STRATUM, or (with allowance of the January / season convention) from STARTEV, but because year is used so often as a class variable in analysis, we added the YR variable for convenience.

Under the groundfish design, the sampling units are the 2.5 minute grids, which are identified in the data base by the character version of an 8 digit number – two digits for degrees latitude, one for 10s of minutes latitude, the analagous 3 digit identifier for longitude, and a 2 digit identifier for 'grid within 10 minute block,' running from 1 to 16, across then down, in a chart with north at the top. The 10 minute 'blocks' are identified by the convention of the [degree -- tens of minutes] location at the lowest lat lon of the block (*i.e.* lower right). This 8 character identifier is a change from block/grid identifiers in previous versions of these data, where different conventions were followed at different times and places. The new convention is used for all stations selected under the block/grid strategy, even those outside the Primary Area. The variable STRATUM is used in the groundfish views of the Primary Area, with values of EASTERN or WESTERN, dividing the grids at 89°30′. Grid areas in hectares are attached, based on calculated areas incorporating the change in km per minute longitude with changing latitude. These spatial areas are recorded under the variable WATE, spelled that way to avoid any chance of confusion with biomass measurements.

SEAMAP-based data use all 15 characters of the variable stratum as a stratum identifier. The first character is S or F for summer or fall. The next 4 are the (4-digit) year. The next single character is S C E or L, referring to the (now) 'standard' clockwise sweep, a counterclockwise sweep, the 'early' summer phase east of the river, and the 'late' phase east of the river, respectively. Four characters then describe the alongshore portion of the stratification – the beginning and ending shrimp statistical zones of the stratum. The next character is D or N for membership in the day or night stratum. The last 4 digits are the depth limits of the stratum. Spatial areas (hectares) for each stratum are attached as the value of the WATE variable. For SEAMAP, these are based on the 5-fm spatial area estimates of Patella 1975, further subdivided by similar method into the finer depth increments (courtesy of M Brassfield, NMFS-Miami). Users may note some instances of equal areas for different strata – these were used when chart resolution and/or size of area were too limiting to measure individual stratum areas by polar planimetry. In those cases, WATEs are Patella's areas divided by 5, one for each fathom.

Users may detect some samples where the recorded location data place a sample nominally outside a stratum or grid to which it has been assigned in the data base. We recommend accepting the strata as assigned, without adjustment. There can be several reasons for an anomaly between stratum and location. In many cases, interchanging the closed or open side of the spatial interval will put a station back in the correct box. On other occasions, the vessel may have drifted outside the intended boundary before setting. In the early days, locations were not measured or recorded as precisely as is done now. Field Party Chiefs have been authorized to move stations 'slightly' from their random selection location if faced with an obstruction, rather than taking time to move to a new random selection (which could be 60 miles away or more). Occasionally, avoiding obstructions has moved a station outside its target stratum area. SEAMAP tows have a minimum tow time, so sometimes depths may extend outside the target stratum. Sometimes, locations could be recorded incorrectly, but with error not sufficient to trigger the automated error checks based on location or distance vs time anomalies. In such cases, intent was usually clear, even if exact location was in doubt. And finally, sets have sometimes simply been made in the wrong location. Within limits, we have placed stations within their <u>intended</u> stratum or block/grid. For SEAMAP, we have been forgiving with alongshore anomalies, but strict with depth, in accord with the expected rate of change in the two directions. For groundfish block/grids, we tolerated anomalies of about 2 km if we could otherwise discern intended location. For mapping purposes, we recommend using starting latitude and longitude as given. For analysis taking means over space, we recommend using the stratum or grid designation as given in the data base, accepting the few anomalies as a consequence of the realities of working at sea. The alternatives are to set aside an additional fraction of the data (we believe unnecessarily) or to modify the stratum weights after the fact (which we consider an unneeded over-elaboration). Stations for which an anomaly was just too large for us to accept, or for which the location in general was in serious question, have already been set aside, labelling their purpose as 'UE' (unrecoverable error). These stations are not included in the Oracle views or in the SEDAR data files, but remain in the primary data tables.

Users should not sum the values of the WATE variable over trawl hauls when calculating stratum or grid means. The WATE variable is a property of the stratum or grid, not the individual sample.

Including the WATE variable in the data base has proved to be a very useful convenience, but the uses of the WATE values as they stand are linked to the specific designs. Users may wish to override the values for some purposes. For example, setting WATEs=0 is an easy way of obtaining averages over spatial areas less than the total survey area. WATEs might be increased for some treatments of missing values. For analyses not matching the SEAMAP stratum boundaries (e.g. catches in state waters), WATEs can be reset to estimates of areas within some spatial target. WATE can even be disconnected from its underlying definition as a spatial area, and used more abstractly for defining relative 'values' of different samples. Values of WATE for all the stratum definitions used in SEAMAP are available in the Oracle table SEAMAP_Wates.

STARTEV ENDEV TIMEZONE

These are the variables that place an event in time. STARTEV ('start event') and ENDEV ('end event') are date/time variables in Oracle, and transfer as date/time variables in SAS. Their actual values are thus the number of seconds from some arbitrary reference time. Both Oracle and SAS have formatting conventions that change the numbers into meaningful dates and times automatically. Recording ENDEV is a relatively recent addition to our shipboard procedures – many early records will have a null ORACLE value for ENDEV, which convert automatically to the SAS missing value '.'. The earliest cruises have STARTEV recorded only to the nearest hour (back when one tried to jam all data onto an 80-column punch card). We had an additional source of time (and location) information available for many cruises aboard the Oregon II, other than the actual data sheets. The vessel bridge kept a handwritten log of locations, times, and operations that was not keypunched as part of the original data. We have since obtained copies of these logs, and are working through entering, and then reconciling any differences. This is why there a sometimes uneven levels of resolution evident among cruises, or even within a cruise. It has been slow work, but fortunately, the nature of the disagreements between the alternate sources usually either point up an obvious error, or causes are non-obvious because the disagreements are trivial. We make the best decision we can about which is the more accurate information, and go with it in the data base.

TIMEZONE is a 3-character abbreviation of the local timezone name (e.g. CDT, EST). For the cruises in the views, CDT is the customary zone used, but the time does change from CDT to CST during the fall survey. We have not been consistent on exactly when we change our clocks at sea. Some cruises changed at midnight on the date used by the nation. Some cruises changed at 2 a.m. Some cruises were completed retaining CDT. Some cruises changed at the end of a leg. Some changed at the end of a watch. In some cases, a new watch came on duty, and went back to CDT, so users may note some flipping back and forth within a cruise. We have examined the anomalies (primarily by looking at the succession of times and timezones over groups of stations), and believe the correct timezone is attached to each STARTEV. There were only at handful of cases where the timezone recorded appeared to be in error. We have changed these to the code that appears to be correct. The latest Oracle could accommodate timezone into its date/time variables, but most other software cannot, so we keep timezone separate.

STARTLAT STARTLON ENDLAT ENDLON

These variables locate each sampling event in space. Units are decimal degrees. By convention, west longitudes are recorded as negative numbers. ENDLAT and ENDLON are relatively recent additions to our data records.

Precision of values recorded over our survey program has varied, but most commonly the original record was [degrees, whole minutes] in early surveys, improving to [degrees, minutes to tenths] later on. We retain variables with the original entries elsewhere in the base to make checking and editing easier, but just present the decimal degree versions in the views. 'Precision' in views is nominally to the number of digits the machine carries without special instruction, but most presentations are formatted to only three decimals, roughly similar to a tenths of minutes resolution common in the original data.

Accuracy has varied over the time series as well. NOAA vessels have steadily upgraded navigational equipment, but the exact dates of past changes are not known. We have been through LORAN A, LORAN C, and improving GPS. Even if we could recover the dates of upgrade, we have no mechanism for reporting the spatial variation in accuracy associated with the old LORANs. Accuracy was also limited by how the data were actually recorded. On the bridge, the main concern is safety, including avoiding other boats, and making sure gear operations are secure. Location readings were frequently delayed until the vessel operator could turn his attention to hand-recording the readings from the available equipment. Nowadays, readings are recorded electronically for the Oregon II, but this is very recent, and has not occurred on the partner vessels. Fortunately, these levels of inaccuracy are not major issues for most applications of trawling data. However, these inaccuracies are the source of some of the anomalies we accept between location and stratum or grid membership. Probably more importantly, these inaccuracies do limit our ability to add editing checks. For example, we check whether it is 'possible' for a vessel to have moved between two locations it was reported at in the nominal time between location records. We have had to be very soft in the tightness of those checks.

A more important issue in the accuracy of locations of the historical data has been the tendency to make 1 digit errors in hand recording, or subsequent keypunching, of locations. These were perhaps the most frequent errors encountered when examining the data base. Most candidate errors are obvious when one plots cruise tracks, and for many, the correction was obvious, and often verifiable by information written down elsewhere. We have made changes wherever we decided change was warranted. For those with any question, we have kept the original entries elsewhere in the base, but we don't really recommend anyone try again to reconcile any of these anomalies – there is not much to go on. If we could not reach a clear choice for error correction, we left the data as they were, but usually flagged the event with an Unrecoverable Error code if the data should have been part of a survey project. These flagged events will not appear in the trawl events views.

If one needs to characterize a sampling event with a single location, we recommend using the starting values. The ending values are more likely to be missing, and those that are present have never received as much analytical examination as the starting values.

STARTWD ENDWD DEPUNITS

The depth variables. Including both STARTWD and ENDWD ('Starting water depth' and 'ending water depth') began with the start of cross-stratum trawling under SEAMAP. For groundfish, samples were characterized by a single depth at the start of a station, placed here in the STARTWD variable. As groundfish trawls probably paralleled depth contours much of the time, using a single depth probably did not conceal much depth variation. In SEAMAP trawling, depth is one of the 'lines of position' used to establish station location – station locations are defined by the intersection of a line or either latitude or longitude and a depth as determined at sea. Initially, trawling always went from the inner depth boundary to the outer, but that proved to be wasteful of time, so trawling in either direction was allowed. Implementing cross-stratum trawling proved difficult at first, in large part because, unknown to the cruise designers, that depth sounding equipment on the Oregon II in 1981 only read out to 1 fm resolution (this impacts project TC). Equipment was upgraded for succeeding years, and depths are recorded to 0.1 fm. The variable DEPUNITS is included because some of our programs are metric. All trawling depths are fathoms (coded F).

SPEED

Speeds are listed in knots. Our ability to measure speed has changed considerably over years. The Oregon II originally had no direct capability to measure speed. Speed is changed by changing the pitch of a variable propeller. Pitch was the only information recorded for the first several years. There was a presumed relation between pitch and speed, but we have never seen documentation on any calibration trials, nor do we have information on variations to be expected under different conditions. As navigational equipment improved, so has our ability to estimate speed, but we have no documentation of exactly when changes were made, or what the specific impact of any change would have been. Currently, the Oregon II can produce an electronic readout of speed using GPS over very short time/distance series. Other vessels participating in SEAMAP probably cover a wide range in speed estimation capability, but specifics are unknown.

We have variation in speed in the data from 3 sources: design difference, variable success in attaining the nominal speed, and speed variations during single tows. The groundfish program has an intended speed of 2.5 knots. A target speed of 3 knots was set independently for the Texas Closure survey based on the presumed speed used by the fishing industry, and the SEAMAP surveys that evolved from the Texas Closure work retained the 3 knot target. Local conditions and operator practices cause variations in a characteristic ('average') speed among hauls. And finally, speed is never constant over an entire tow, both due to small scale variation, and because it is necessary to change speeds to deploy and retrieve nets successfully.

As we lacked the capability to measure speed precisely until very recently, we have never studied the effect of speed on gear performance in any detail. Informal analysis of small sets of data do not show much indication of efficiency change between 2 and 3 knots, but towing at much over 3 knots can pick the gear up off bottom. We have found it prudent to remind the vessel operators frequently to keep speed at 3 knots or below, to counteract the natural tendency to hurry up when behind schedule. For our routine analyses, we retain stations with recorded speeds above 3 knots, assuming as an approximation that catch rates per area covered were unaffected.

Because there is known variation in speed in the data base, we presume most users will want to adjust for it, putting catch rates more on an 'area covered' basis. That makes it necessary to have estimates of speed for every haul, even when not measured directly. We have assigned SPEED the value of the intended speed for all cases where speed was not directly estimated or recorded. These assignments are obvious, as records for an entire cruise will all have identical entries for speed.

DURATION

DURATION is the name now used for time fished by a piece of gear. Values are in minutes of time.

Because of the coarse resolution of time recorded in the early surveys, we have not tried to use DURATION and STARTEV to add values for ENDEV where it is missing. For similar reason, we have also not tried to guarantee agreement between ENDEV minus STARTEV and DURATION in the groundfish data. A check between the starting, ending, and fishing times was built into previous versions of SEAMAP entry software , and thus disagreements should have already been addressed at entry time for the later data. If a user does uncover a disagreement, we recommend using STARTEV and DURATION, and distrusting ENDEV, as ENDEV has never received much analytical scrutiny in the past. But please alert us to any discrepancies you find.

DISTANCE

All values of DISTANCE (km) in the data base are estimated from change between starting and ending locations, assuming a great circle path between. This variable is primarily for editing, in the sense of spotting implausible values. Given the limitations of the location data, we strongly recommend against trying to use the DISTANCE variable as a measure of area covered. Using time, adjusted by speed if desired, will be far more accurate.

SCS

Effectively a boolean variable (codes Y or N), indicating whether the automated electronic system currently being made operational aboard the Oregon II was used to pipe electronically sensed data directly into the data base.

PV

A variable to indicate that another vessel(s) was conducting a paired tow alongside the vessel in the record. This is a 6-character variable, concatenating up to 3 vessel codes for the paired vessels. (Currently, only 3 vessels max have ever towed in close parallel, so the last 2 characters have always been blank).

We may consider modifying this variable to identify the vessel, cruise, and event pairs or triplets directly, but as of now, the paired data sets are small enough that is fairly easy to recognize which stations were paired from the space and time information. The PV variable just allows quick isolation of stations where pairing took place.

OPCODE

If the fishing gear fails to operate properly, an explanatory OPCODE is placed on the data sheet, sometimes with more extensive written comment. Sometimes a haul is then repeated, but if time is short, or the cause of failure was damage to the gear from the bottom, the vessel will move on without repeating.

No sampling event with an opcode other than null is accepted in the survey data. Catch data, even if any were recorded, are not considered reliable. Records with non-null OPCODEs are not included in the data files at the SEDAR, but do remain in our primary data tables.

ALDIST OFFLDIST

The coast of the Gulf of Mexico curves in lat/lon space, so there is no single 'natural' variable to characterize location in the alongshore direction. We have constructed one in ALDIST ('alongshore distance'). A smoothed curve was developed from the 10 fm contour locations used in D. Koi's MAPP software (Galveston lab). This smoothed curve was then digitized into points about 1 km apart by great circle distance. Each point is given a value (km from an arbitrary start at the Texas / Mexico border) that becomes the scale for ALDIST. Brute force was used to project every sampling event onto the digitized curve. The point on the curve with minimum great circle distance to a given sampling event is that event's ALDIST. OFFLDIST ('offline distance') is that minimum distance (km) to the digitized curve (negative if shoreward).

SEAMAP TOW WHICHNET

These are SEAMAP precursors of the station / event structure of the present data base.

The variable SEAMAP (short for 'SEAMAP Station Number') was added as a station-level designator at the start of the program, but unfortunately, we never maintained consistent coding conventions over more than a few cruises. Users who deduce the coding conventions used in any subset of cruises should never rely on those conventions in their programming logic. The STATION or STRATUM variables in the data base now serve the functions that users might have expected of the SEAMAP variable. SEAMAP is a 5-character 'station number' mixing alpha and numeric characters. Where membership in a SEAMAP survey design is indicated by the PURPOSE variable, it is valid (but unnecessary) to use SEAMAP as a unique identifier of a single, cross-stratum station, but there are other, inconsistent uses of the variable for non-member events. There may also be entries in the SEAMAP variable on non-SEAMAP cruises. In these cases, someone probably recorded information in the otherwise unused SEAMAP field for editing purposes. Any such entries should be generally ignored, but we have retained them where they exist, just in case.

TOW is a single digit, stored as a character, assigned to separate trawl hauls in sequence of occurrence within a SEAMAP station. WHICHNET is also a single digit, again stored as a character, coding the net location (1 – port, 2 – starboard, 3 – astern). Sometimes a paired towing project was conducted in conjunction with a SEAMAP survey. In those cases, users consulting the underlying tables will find two events with different WHICHNETS for each value of TOW, with possible multiple TOW values within the same SEAMAP. In the survey views, the tows intended to be used as standard events are included, the but the paired events, usually using experimental gear, will not appear. More often than not, the port side was used as the standard, with the starboard used for other experiments. This was not always true – for example, we would switch if any problems arose with the primary winch. We are not aware of any side bias for the Oregon II, but as we tend to use one side as the standard, we have not done extensive testing.

Information on net location was recorded on many of the cruises in the groundfish program, including non-survey projects. If that information existed, it has been brought forward into this data base in the WHICHNET variable. We did uncover a good many errors in the older data base entries when comparing the data with the sheets when identifying purposes. We have corrected any errors so located to the values on the data sheets, if the data sheets appeared to be internally consistent. Entries for TOW have also been made for some of groundfish data, well after the fact. We added tow sequence numbers when trying to determine block/grid and project membership for editing purposes. We have kept these values, but there is probably no reason to use them.

We do not recommend rejecting a station because of a null value for TOW or WHICHNET, or (on rare occasion) even SEAMAP. TOW and WHICHET appear to have been omitted fairly frequently in recording or entering data in the past. We believe we have project membership and stratum assigment correct for those few stations where SEAMAP was not recorded for whatever reason.

COMTRAWL

This is a character variable that brings forward any written remarks in the Comments section of previous versions of this database. There has been no effort to make the field searchable, and certainly no conventions of any consistency, other than might have been used by individual watch leaders.

Catch Variables ('Biological' Data)

VESSEL CRUISE STATION EVENT ENTRY

In combination, the unique identifiers for catch data records. VESSEL, CRUISE, STATION, and EVENT are as described in the trawl event views section above. Entry is 3-digit number assigned in sequence at entry for each catch record within a trawling event. There is no significance to the ordering in the sequence -- it is just a way of generating a unique identifier. We use ENTRY because the other potential identifiers may be changed over the life of the data base.

BIOCODE YOY TAXON

BIOCODE, a 9-digit numeric, is a hierarchical code for extracting catch data at the species, genus, family, ... levels. TAXON is a 13-character abbreviation for the scientific name for a species (or higher level name). There is a one to one match between BIOCODE and TAXON. The match between TAXON and BIOCODE is documented in the table BIOCODES in the Oracle database.

For names of individual species, TAXON is a concatenation of the first 7 letters of the genus name, and the first 6 letters of the species name. Mid-field blanks are used if the genus name is shorter than 7 letters. For higher taxonomic levels, the taxon name is spelled out, up to 13 characters.

YOY ('young of the year') is a single character viable, coded either 'Y' ('yes') or blank. If two distinct size classes are easily evident in the catch of one species in a trawl haul, the catch of each size classes may be reported separately, with YOY coded 'Y' for the smaller size group. This convention was developed at the start of the groundfish program to retain some obvious information, and to avoid otherwise inexplicable calculations of mean weights for individuals in a catch. YOY cannot be used to separate out a size or year class across stations – if only one class was present in any single haul, then the YOY variable was not used for that haul in most cases. There are no objective standards for determining when YOY should be used, so some inconsistency should be expected. Most users will want to sum all entries for a BIOCODE in a sample, and ignore the YOY variable.

Identifications are sometimes changed upon expert, shoreside examination of specimens. Accepted scientific names may change over time, as may position of a taxon in the systematic hierarchy. In the current database, all catch records have been brought up to date – i.e. taxonomic information on the catch records has been modified such that a single BIOCODES table applies to all cruises, but there is no way to track previously made changes. This database does have two variables to help track these changes in the future: ORIGTAX and ORIGBIOC (not present in the views). These will retain the original shipboard decision under the biocode table current at the time. At present, these values all match the BIOCODE and TAXON entries, as we have not attempted to reconstruct past changes in the biocoding system (and probably never will for data present at the time this base was first created.)

We are developing major modifications to the biocoding system for a number of reasons. The highest levels of our biocode hierarchy are not convenient to use. The numerical ordering of the biocoding is breaking down, as we are 'running out of room' between existing numeric codes. Particularly in the higher level plankton identifications, we are starting to encounter cases where 13 letters are not unique identifiers. There is no structural support at present for making and tracking changes to the system over time. When development is complete, large scale changes will be made at one time. Impact on users should not be great. User's will have to modify their programs to accommodate any biocode changes, but that has to be done even today, without benefit of an formal system for changes.

The resources surveys program has always worked hard at taxonomic expertise and discipline, making every effort to take identifications to the lowest level possible, without being reckless by assigning to a category beyond the level of knowledge available. An extensive library of keys is maintained on the vessels, and specimens in question are brought back for consultation with appropriate experts. Most users are unlikely to find any controversy in fish identifications, at least beyond any already under active discussion by experts in the region. Invertebrates (other than those of commercial importance or highly abundant) had not historically been taken to as high a level of taxonomic resolution as the fishes, but we have been upgrading our invertebrate taxonomic resolutions over time. This 'upgrading' (which extends into the fishes as well) presents a potential analytical problem in constructing time series, in that taxonomic resolution can change over time. For example, the fraction of *Cynoscion* sp. has declined over the years as progressively smaller specimens were identified to species. However, given the high fraction of *Cynoscion* sp, on average, and the variation in that fraction over time, any time series analysis of *C. nothus* or *C. arenarius* individually could be meaningless. Users should generally check for entries at higher levels in the taxonomic hierarchy than they are seeking, to consider if there could be a substantial fraction of their target taxon tied up in identifications at a lower taxonomic resolution.

Because both identifications and the systematic hierarchy can change, VESSEL CRUISE STATION EVENT BIOCODE YOY or VESSEL CRUISE STATION EVENT TAXON YOY are not necessarily unique identifiers.

N W

These are the total number and weight (kg) values for a catch record entry from each trawl haul. Values of zero are not stored in the catch records. Because of the use of the YOY variable, and the possibility that identifications were changed upon examination of specimens after first entry, the variables TAXON and BIOCODE are not necessarily unique within a trawl haul. With this structure, user's must sum over all records of the desired value of BIOCODE or TAXON for each sampling EVENT to get total catches for a particular taxon. Programming logic should be set up such that if no entries for a taxon are encountered, a value of zero will be returned for the catch of that taxon in that sampling event.

For those hauls where subsampling was conducted, N and W are the 'expanded' values for the total catch. Subsampling details, where known in the historical data, are included in the underlying tables. Only single subsamples have been taken for any catch. Most historical records exist only as 'expanded' catches, with no information on subsampling that might have taken place.

Resolution for W has varied over the years. Initially, weights were measured in pounds, and recorded to the nearest whole pound. (It is quite likely that the scales available then were accurate only to the nearest pound.) There may have been some variation in rounding convention for weights estimated at less than one half pound – either 1 or 0 may have been recorded. Equipment improved, and so did the recorded resolution, but the dates of improvement are not documented. SEAMAP switched to metric weights, recorded to gram resolution. All weights throughout the trawl data base are now in kilograms.

All data in the SEDAR files listed in the design sections of this paper have been limited to red snapper, and summed over all entries for an event, which leads to replacing the N and W variables by new variables named NUMBER and WEIGHT.

PASC

The PASC variable in the catch records matches the PASC in the event records. It is not necessary to use PASC in the matching, but we have retained it both for editing, and for those users who wish to retain previous logic in analyses.

Simple Indexes

Primarily because multiplication is distributive over addition, there is often a myriad of possible ways to calculate indexes from even the simplest design. As design complexity increases, so does the number of possible estimators. Extending into model based estimators, the options grow further still. To propose a 'base index' to be associated with each time series, I have decided use go with indexes that are quite simple in structure, computation, and traditional form, at the cost of adding some assumptions about gear performance and variance structure. For example, not all tows are equal length, sometimes by design, sometimes due to operational problems. I assume that catch rates can be normalized to unit time by proportion. We have some data on variation of speed from the nominal target for some cruises. I again assume that catch rates can be normalized to the target speed by proportion. In combination, we are normalizing to area swept, but I will express each observation in the more customary 'number per hour' (at the target speed). Where different vessels were used, or different sides on the same vessel, I assume that catch efficiencies were identical. (We have some analyses examining that assumption under development.) By assuming rather than estimating equal efficiencies, I also assume no contribution to index variance due to imperfectly known among-vessel calibration factors. For the simple indexes, I will make no adjustments for any missing coverage.

For groundfish surveys of the Primary Area (FG), I have elected to use mean catch rate in each grid, calculated as the sum all catches divided by the sum of all efforts (as fishing times, at 2.5 knots; adjusted for speed, when known) as an observation. Two strata were imposed: east and west of 89° 30′. Within strata, the simple mean of catch per hour over the sampled grids was calculated. The index for the entire primary area is thus:

$$\sum W_h \cdot \overline{Catchrate}_h \ ,$$

where the W_h is the fraction of the number of grids present in each of the 2 strata (roughly 0.43 East to 0.57 West). Variance is estimated as:

$$\sqrt{\frac{\sum W_h^2 \cdot s_h^2}{n_h}},$$

where s_h^2 is the variance among grid means with stratum h, and n_h is the number of grids sampled in each stratum. The finite population correction is ignored, despite sampling fractions of about 17%. As there are no formal hypothesis tests intended – the variances are merely a guide – there is no motivation to argue for lower values for the sake of lower values. As the actual fraction of area covered within each grid is small, a finite population correction seems rather superfluous.

An obvious alternative structure would be a general model with each of the replicate tows as an observation, and analysis of variance structure containing among and within grid variation. We did not select this structure because 1) there is no virtue in having variance estimates among tows at the single spatial scale imposed by the grids, and 2) unequal numbers make the mean squares messy functions of underlying variance components. We also decided against the alternative of using years as effects in a single model, as there is no reasonable expectation of additivity or equal variances.

For SEAMAP-design time series (ES, SS, FS, TC, C2, FF), I again use sum of catch over sum of effort (here, adjusted to 3 knots), but this time summed over all hauls within each SEAMAP stratum, as an observation. (For most SS and FS strata, there is just one trawl haul.) The index is then the weighted mean of these observations, with the weights being the spatial area for each stratum:

$$\frac{\sum W_i \cdot \overline{Catchrate_i}}{\sum W_i}$$

The index variance estimate is similarly weighted:

$$\frac{\sum Wi \cdot (\overline{Catchrate_i} - Wtdmean)^2}{\sum W_i}$$

I treat the weights as if they were known constants, not variables measured with error. Note the W term in the SEAMAP design, vs the W^2 terms in the groundfish variance expression. This is correct. The groundfish grids are repeated draws of equal value within a stratum. The SEAMAP stratum totals are single samples of unequal value in the single, population-wide estimate.

It is possible to calculate indexes base on weight per hour, and on presence / absence. A weight index is probably not very useful for population level analyses, and the changes in resolution of the weight measurements may limit utility in any case. A presence / absence index might be more useful, but is in some ways more problematic. For some species, often those of low abundance and patchy spatial distributions, indexes based on presence / absence have appeal, in that they often lack the unusually large fluctuations seen in number per hour indexes. On the other hand, it seems reasonable to expect number and weight indexes will be proportional to abundance over a wide range of abundances. Proportionality is difficult to envision for presence / absence. It also seems reasonable to make straightforward adjustments for varying effort (time or length of trawls, number of trawls, and expansions from subsampling) for number indexes. Because red snapper are abundant in the trawls, and do not show the kinds of fluctuations that make presence / absence attractive for other species, I have not included any presence /absence indexes in this document.

Results for Base Indexes:

Base index results are collect in 3 figures: the Fall Groundfish (FG) series in Figure 1; the First Fall (FF) and Fall SEAMAP series (FS) in Figure 2; and the Texas Closure (TC), Early SEAMAP (ES), and Summer Seamap (SS) series in Figure 3. Points within individual series may be considered comparable, but even for those depicted on the same graphs, points should not be considered comparable across the separate time series. All results are number per hour, but no analysis is included here to attempt to calibrate across the series differences. Work to allow calibration across series is the subject of a separate paper (Nichols 2004) at this SEDAR.

Some analytical considerations

I close with some comments on the prospects of combining indexes across time series. There is a range of options, trading off simplicity against rigor, and varying the explicitness of assumptions:

1) assume a simple conversion function on the means of the time series indexes
2) estimate a conversion function separately from the index model, and apply it to the means
3) make estimation of effects appropriate to time series differences part of a single model,

Analysts can also consider including external covariates (in space, for each event; or for each annual mean), estimating effects of missing stations, evaluating use of multiple vessels, and making assumptions about error distributions.

Almost everyone who has encountered these data begins to think immediately in terms of elaborate model-based estimation, probably because so many potential terms are available. That approach can be very useful for searching for patterns, but I recommend caution in using this approach for stock assessment issues. Most approaches assume additivity and homogeneity, assumptions almost certain to be wrong. Additivity can be a real killer. We should usually expect most effects to be multiplicative. However, the analytical options for multiplicative structure have not been satistfying. Log-transformed models are plagued by the frequent presence of zeros in the catches. Delta distribution approaches tend to be led astray by numerical analysis problems, and the practical properties of the presence / absence portion of delta models are not well known. I explore an approach using the negative binomial in Nichols 2004.

References:

Kemmerer A J, Gutherz E J, and Rohr B A. 1982. Summary of fishery independent surveys for groundfish in the northern Gulf of Mexico (1972-1981). Report to the SEFC 1982 Stock Assessment Workshop.

Leming T D & Holley H J. 1978. A computer software system for optimizing survey cruise tracks. Fish. Bull. 76(3):706-714.

Nichols S. 2004. Calibration among the separate trawl survey programs to extend the time seires for juvenile snapper indexes. Report to the Gulf of Mexico Fishery Management Council. Red snapper data SEDAR, April 2004.

NMFS and GSMFC. 2001. SEAMAP Field Operations Manual for the Collection of Data. Available from NMFS Pascagoula lab, or Gulf States Marine Fisheries Commission.

Patella F. 1975. Water surface area within statistical subareas used in reporting Gulf coast shrimp data. Mar. Fish. Rev. 37(12):22-24.

Other descriptive papers:

Anonymous. 1976. Groundfish survey cruise track optimization system. Project Report. SEFC contrib. No. 457.

Juhl, R, Roe, R B, Gutherz, E J, Drummond, S B, Roithmayr, C M, & Benigno, J A. 1974. Oceanic resource surveys and assessment task status report. Pascagoula laboratory internal report.

Nichols, S. 1983. Development of the June-July Texas Closure / stock assessment sampling program for brown shrimp. Report to the Gulf States Marine Fisheries Commission.

Nichols, S. 1984. Goals and future directions of the shrimp/bottomfish research cruise program. Report to the Gulf States Marine Fisheries Commission.

Nichols, S & Pellegrin Jr, G P. 1989. Trends in catch per unit effort for 157 taxa caught in the Gulf of Mexico fall groundfish survey, 1972-1988. Report to the Gulf of Mexico Fishery Management Council.

Stuntz, W E & Henwood, T A. 1995. Marine biodiversity data in the southeastern US. Pascagoula Laboratory internal report.

Table 1. Cruise dates, and number of grids occupied with at least one successful trawl for the Fall Groundfish Survey.

1972	27-Sep-72	30-Nov-72	102
1973	13-Nov-73	9-Dec-73	225
1974	5-Nov-74	27-Nov-74	229
1975	28-Oct-75	17-Nov-75	232
1976	2-Nov-76	23-Nov-76	265
1977	5-Oct-77	23-Oct-77	237
1978	10-Oct-78	1-Nov-78	235
1979	25-Oct-79	19-Nov-79	255
1980	29-Oct-80	24-Nov-80	200
1981	14-Oct-81	15-Nov-81	211
1982	12-Oct-82	21-Nov-82	249
1983	20-Oct-83	14-Nov-83	202
1984	9-Oct-84	9-Nov-84	205
1985	15-Oct-85	7-Nov-85	102
1986	29-Oct-86	9-Nov-86	41

Table 2. Cruise dates, and percent of strata successfully sampled in the First Fall and Fall SEAMAP surveys. There were 220 strata scheduled each year.

1987	23-Oct-87	22-Nov-87	71%
1988	20-Oct-88	21-Nov-88	95%
1989	20-Oct-89	20-Nov-89	95%
1990	16-Oct-90	18-Nov-90	95%
1991	14-Oct-91	18-Nov-91	98%
1992	18-Oct-92	19-Nov-92	91%
1993	15-Oct-93	18-Nov-93	97%
1994	14-Oct-94	20-Nov-94	97%
1995	16-Oct-95	4-Dec-95	98%
1996	11-Oct-96	22-Nov-96	98%
1997	11-Oct-97	20-Nov-97	97%
1998	14-Oct-98	18-Nov-98	97%
1999	16-Oct-99	20-Nov-99	98%
2000	14-Oct-00	19-Nov-00	98%
2001	15-Oct-01	15-Nov-01	99%

2002	12-Oct-02	17-Nov-02	94%

Table 3. Cruise dates, and number of strata successfully sampled in the Texas Closure and Early SEAMAP surveys.

1981	7-Jun-81	2-Jul-81	27
1982	1-Jun-82	13-Jul-82	112
1983	1-Jun-83	15-Jul-83	118
1984	6-Jun-84	3-Jul-84	96
1985	10-Jun-85	5-Jul-85	93
1986	10-Jun-86	6-Jul-86	104

Table 4. Cruise dates, and percent of strata successfully sampled in the First Fall and Fall SEAMAP surveys. There were 220 strata scheduled each year.

1987	11-Jun-87	15-Jul-87	98%
1988	11-Jun-88	14-Jul-88	91%
1989	7-Jun-89	16-Jul-89	79%
1990	7-Jun-90	13-Jul-90	90%
1991	3-Jun-91	13-Jul-91	99%
1992	4-Jun-92	13-Jul-92	98%
1993	3-Jun-93	18-Jul-93	96%
1994	2-Jun-94	18-Jul-94	97%
1995	6-Jun-95	19-Jul-95	96%
1996	5-Jun-96	17-Jul-96	95%
1997	4-Jun-97	16-Jul-97	93%
1998	2-Jun-98	16-Jul-98	91%
1999	3-Jun-99	20-Jul-99	97%
2000	9-Jun-00	19-Jul-00	88%
2001	8-Jun-01	22-Jul-01	66%
2002	3-Jun-02	17-Jul-02	97%
2003	17-Jun-03	17-Jul-03	65%

Figure 1. The base index for the Fall Groundfish time series.

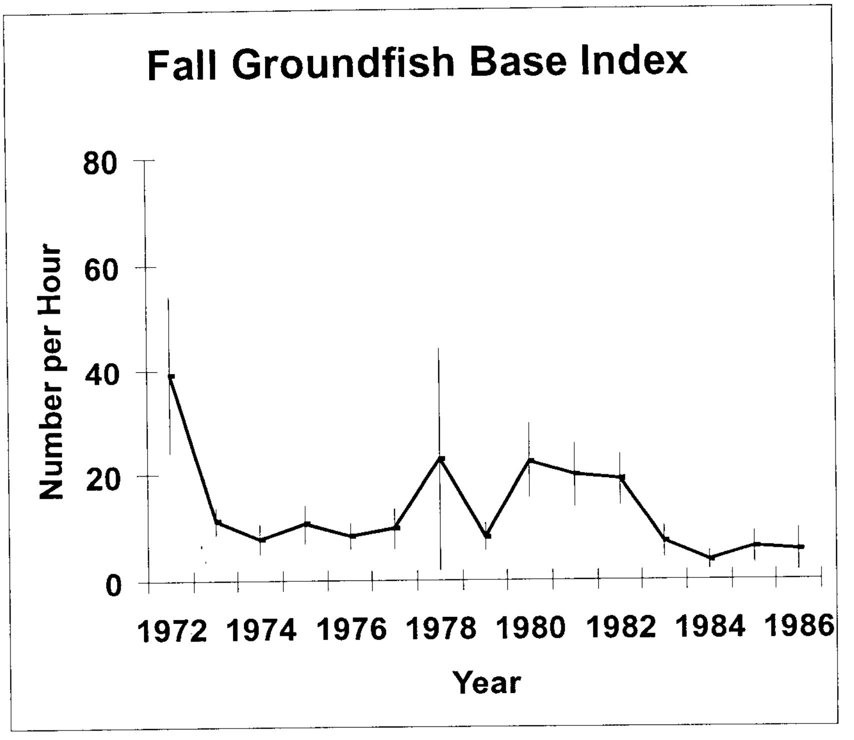


Figure 2. The base indexes for the First Fall and Fall SEAMAP time series.

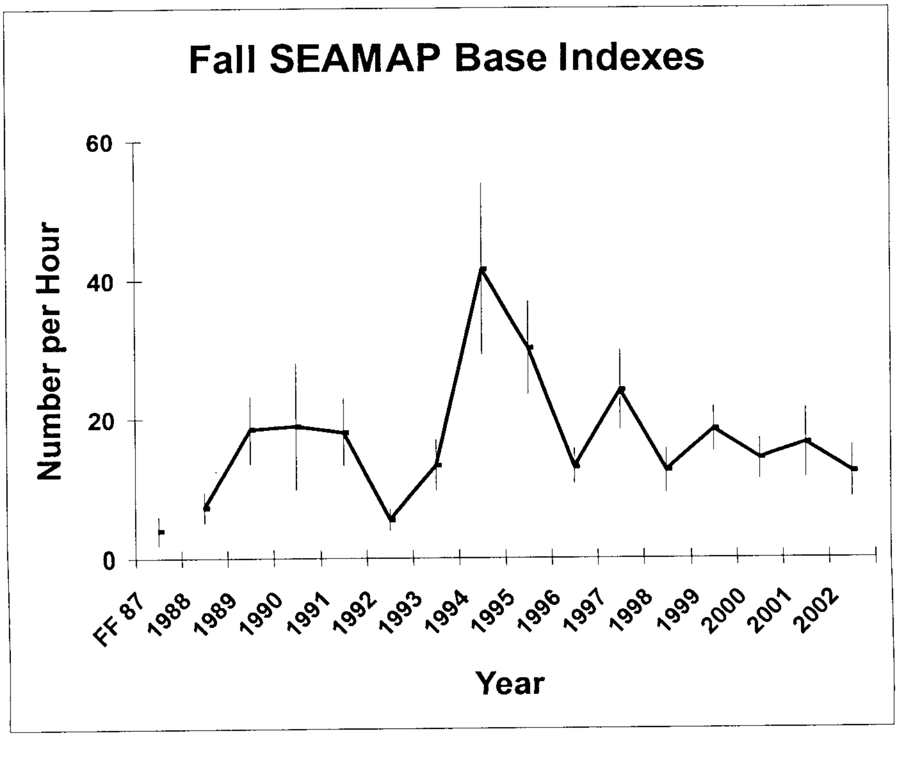


Figure 3. The base indexes for the Texas Closure survey (TC) and the Early SEAMAP time series (ES). These two were plotted on the same graph for convenience. Do not assume that the TC and ES raw results in number per hour are comparable – they do not have the same implications regarding stockwide abundance.

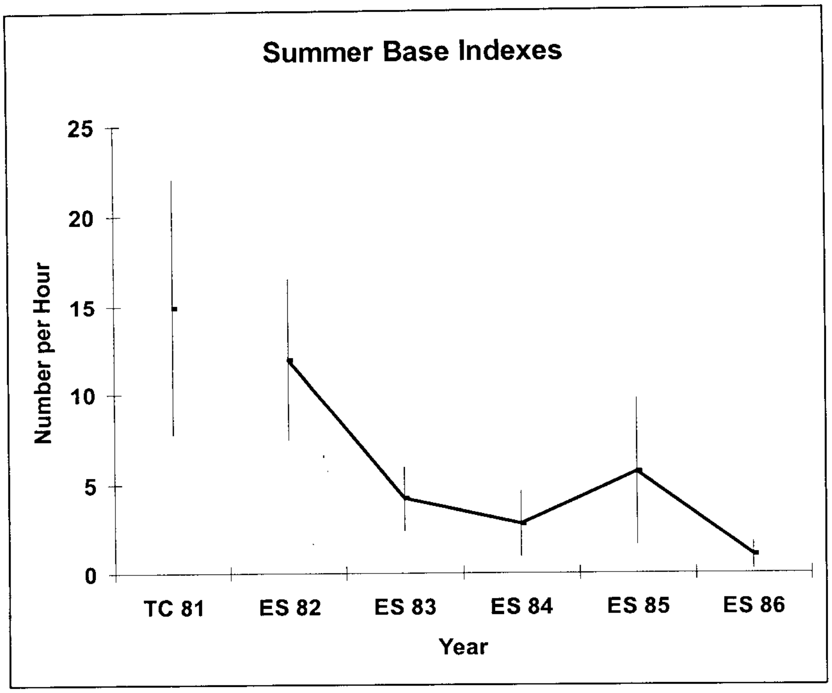
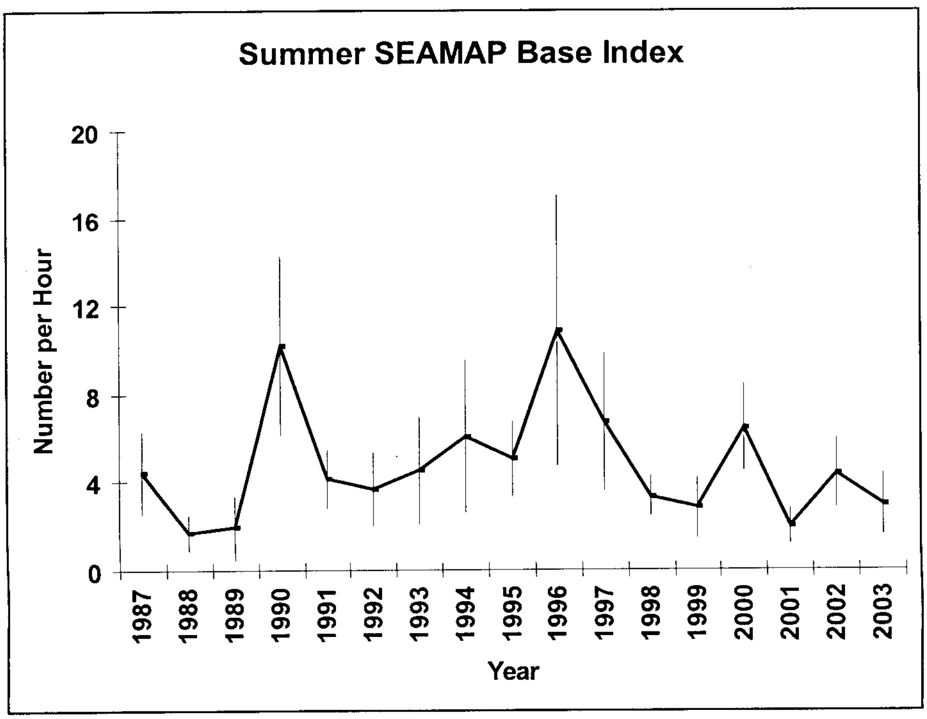


Figure 4. The base index for the Summer SEAMAP time series (SS).



Appendix. Closely Related Trawling Projects

There were numerous other projects conducted aboard the trawl survey cruises. Those unrelated to the main surveys, or those using non-standard gear, are not listed here, but several programs were related to the surveys, usually in the sense of extending their ranges in space or time. However, temporal and spatial coverage were never consistent enough or complete enough to warrant treating these projects as time series unto themselves, or to consider the expanding the formal survey area to include these data. It may be possible to use these data to check the degree that the formal survey areas actually 'contains' a stock, and whether change in abundance of any species are associated with geographic expansion of its range outside the survey area, but the 'incomplete' coverages will be limiting, particularly for the latter use. There will probably be no Oracle views set up for these projects, but they are readily extracted from the primary tables using the variable PURPOSE.

NS, NF: 'Nearshore Summer,' Nearshore Fall'

Sampling by state vessels inside the 5 fm limit of the Oregon II. This project has only been able to cover part of the coast (and a variable part over years) for part of history of SEAMAP sampling. Design is the same as for the main survey area. At present, we do not have spatial area estimates (WATEs) for the nearshore strata.

OS, OF: 'Offshore Summer,' 'Offshore Fall'

There is a 50-60 fm set of strata to the SEAMAP design, added shortly after the programs inception, and usually assigned to the Oregon II. However, if operational problems arise on a cruise and time becomes short, these stations are usually the first dropped, so coverage is discontinuous in time an space. At present, we do not have spatial area estimates (WATEs) for the 50-60 fm strata.

E2, S2: 'Late' (2^{nd}) sampling east of river, associated with ES and SS series

The original summer SEAMAP design called for covering the area east of the Mississippi river twice, at the beginning and end of the survey window, to assist some state management objectives. The two legs are near the outer limit of what we would consider synoptic. Coverage on the 'early' leg has been nearly complete most years. Coverage on the 'late' leg has been spotty, as any delays due to operational problems come home to roost at the end of a cruise. To establish the time series, data from the 'early' leg are included in the ES and SS series. All 'late' stations have been assigned Purpose= E2 or S2 on the purpose records. Coverage was fairly complete in ES series, so no E2 samples were included. For SS, early coverage was less complete in some years, so trawl samples from the 'late' sampling were also given an SS designation, but only if the stratum was missed in the early sampling. This convention improves coverage at some cost to synopticity.

LP, LS, LF, LW 'Louisiana sPring, Summer, Fall, Winter'

SEAMAP partner Louisiana Department of Wildlife and Fisheries uses a slightly different design from the rest of SEAMAP, covering the area out to 20 fm in statistical zones 13-15. Originally, the LDWF coverage was intended as an enhancement of sampling density in an important area for state management, but as concerns grew about synopticity, the Louisiana vessel was assigned the sole coverage of this area much of the time, with overall SEAMAP accepting the variations inherent in using Louisiana's protocol. The SEAMAP design calls for independent draws of day and night samples; Louisiana's calls for repeating the location for a portion of day/night pairs. SEAMAP trawls nominally at 3 knots; Louisiana nominally uses 2.5 knots. From the data, it appears Louisiana sometimes applies a maximum trawling time, breaking off cross stratum trawling if covering the full width of a stratum would put them behind schedule for completing other sites. Louisiana has in some years sampled 4 seasons; the main SEAMAP survey covers two. Louisiana summer sampling coincides with the main SEAMAP survey. Louisiana fall sampling is usually conducted well in advance of the main survey; Louisiana winter sampling is usually in December, well after the main survey, and in some years sampling was as late as January. All Louisiana stations are assigned the appropriate 'Lx' season code for the Purpose variable on the purpose records. Stations from LS are also assigned to purpose SS (or ES in 1986), and SEAMAP stratum identifiers are attached. Louisiana fall and winter samples are not assigned membership in time series FS, but SEAMAP stratum identifiers are attached to the station records, with the 'sweep' character assigned 'E' or 'L' for Louisiana fall and winter, respectively. Thus the FS series provides a potential 5th season of coverage for zones 13-15 inside 20 fm, although at this time, stratum designation have not been attached to LP records.

FZ 'Florida Zone'

We attempted to extend SEAMAP summer sampling eastward into the Florida panhandle in 1983. This proved impractical due to lack of trawlable bottom, and synopticity concerns. We abandoned the project without establishing a time series.

SE, SW, GF, GW 'Secondary East, West'; 'Groundfish Florida,' 'Groundfish Western Gulf'

Groundfish program sampling outside the 'Primary Area' coverage. The 'Secondary' areas to the east and west were of indefinite extent. Sampling was very variable among years and seasons. Most sampling used extensions of the groundfish random station selection protocol, but whether a selected station was occupied often depended on time available, so spatial coverage could be decidely non-random. In one year, transect sampling was obviously used in SW. There was one cruise each off Florida and off Texas using the groundfish station selection pattern. At present only the SE and SW sampling associated with the FG programs have been identified in the data base through the PURPOSE variable. Other SE and SW sampling will be identified as groundfish trawling from seasons other than fall are addressed.

ID, RR 'Intermediate Depths' 'Royal Red grounds'

Seaward extensions of groundfish program sampling. Sampling under ID covered 50-100 fm intermittently, sometimes for only a small segment of the survey area. Station selection appeared to be an extension of the groundfish design in most cases, although sometimes selections and assignments appeared to be at the 10'x10' 'Block' level, without selection of a 2.5'x2.5' 'Grid.' (For those cases, the BLKGRD variable is assigned a value, with '00' in the grid spaces.) RR sampling usually took place beyond 100 fm, often in a small area southeast of the Mississippi river. Station selection protocols were not well documented.