SEDAR10-DW19

Standardized catch rates of gag from the commercial handline fishery off the Southeastern United States

January 2006
Revised: February 2006 to include DW recommendations (Appendix 2)

Contact: Kyle Shertzer
Kyle.Shertzer@noaa.gov

Overview

Using catch-per-unit-effort (CPUE) data from the commercial handline fishery off the Southeastern U.S., we computed standardized catch rates of gag for possible use as an index of abundance. The time series spans 1992–2004. Trips used in the analysis were selected based on the probability of catching gag, computed according to the method of Stephens and MacCall (Stephens and MacCall, 2004, Fish. Res. 70:299-3210). Standardized catch rates were estimated using a generalized linear model assuming delta-lognormal error structure (Lo et al., 1992, Can. J. Fish. Aquat. Sci., 49:2515-2526). Explanatory variables were year, month, and geographic area.

Data

Variables reported in commercial logbooks are described in Appendix 1. The duration of the data set is short (1992-present), with only partial reporting in 1992. Spatial coverage was constrained to areas off the Southeastern United States (24-35 degrees latitude). Each record describes a single species caught on a single trip.

Of trips that caught gag, over 95% used handline gear, defined here as gear with code H or E (Table 1). Thus, the analysis included handline gear (H and E) only. Excluded were records clearly misreported or misrecorded: The variable effort (hooks/line) was constrained to between 1 and 40, the variable numgear (number of lines) to between 1 and 10; the variable crew (number on boat) to less than 12, and hours fished to positive values. These constraints removed less than 1% of handline records. Also excluded were records that did not report area fished, number of lines, number of hooks, time fished, or days at sea. The resulting data set contained 564,634 records with 165,722 trips.

Methods

Standardized catch rates were estimated using a generalized linear model assuming delta-lognormal error structure (Lo et al., 1992, Can. J. Fish. Aquat. Sci., 49:2515-2526), in which the binomial distribution describes positive versus zero CPUE, and the normal distribution describes the log of positive CPUE. Explanatory variables were year, month, and geographic area; and the response variable (CPUE) was in units gag per angler. Variability of estimates was estimated via empirical bootstrap (n=200).

Effective effort was based on those trips that caught gag (positive CPUE) and those that *could* have caught gag (zero catch, but positive effort). Positive catches are readily available from the data, but without information on targeting by fishermen, zero catches must be inferred. To do so, we applied the method of Stephens and MacCall (Stephens and MacCall, 2004, Fish. Res. 70:299-3210). In essence, the method uses multiple logistic regression to estimate a probability for each trip that gag was caught, given other species caught in that trip. Species used as factors in the regression were selected as those caught in at least 5% of trips. This cutoff simplifies the regression, by excluding rarely caught species; however, preliminary analyses indicated results were insensitive to the value of the cutoff (examined over a range of 0% to 10%). Trips were included if their associated probability was higher than a threshold probability. The threshold's value was defined as that which results in the same number of predicted and observed positive trips, as in Stephens and MacCall (2004). Our method differed slightly from that of Stephens and MacCall, in that we included all positive trips, not just those with probability higher than the threshold.

Results

Trips selected for the analysis (n=41,525 of a possible 165,722 trips) included all positive trips (n=31,086) for gag and zero trips (n=10,439) identified via multiple logistic regression. The regression used 20 species as explanatory variables (Figure 1). A trip was identified as a zero trip if its probability of catching gag exceeded the threshold probability P=0.3583 (Figure 2). The assumption of log-normal error in positive trips appeared adequate (Figure 3).

The estimated index of abundance showed little trend, with perhaps an increase in the last two years (Figure 4, Table 2). Annual coefficients of variation, as estimated by empirical bootstrap, were below 10%.

Table 1. Gears used to catch gag, as reported in commercial logbooks. H=hook-and-line, E=electric reel, TR=trolling, L=longline, T=trap.

gear	Frequency	Percent	Cumulative Frequency	Cumulative Percent
H	26999	82.90	26999	82.90
E	4144	12.72	31143	95.62
TR	768	2.36	31911	97.98
L	381	1.17	32292	99.15
T	278	0.85	32570	100.00

Table 2. Estimates of gag CPUE and variability from commercial handline.

Year	CPUE (pounds gag/hook-hr)	standard deviation	CV
1992	1.33	0.11	0.09
1993	1.56	0.11	0.07
1994	1.33	0.08	0.06
1995	1.56	0.10	0.07
1996	1.64	0.10	0.06
1997	1.38	0.09	0.07
1998	1.65	0.11	0.06
1999	1.52	0.12	0.08
2000	1.47	0.11	0.07
2001	1.34	0.10	0.08
2002	1.52	0.12	0.08
2003	1.83	0.12	0.07
2004	2.00	0.14	0.07

Figure 1. Estimates of species-specific regression coefficients used to estimate a trip's probability of catching gag.

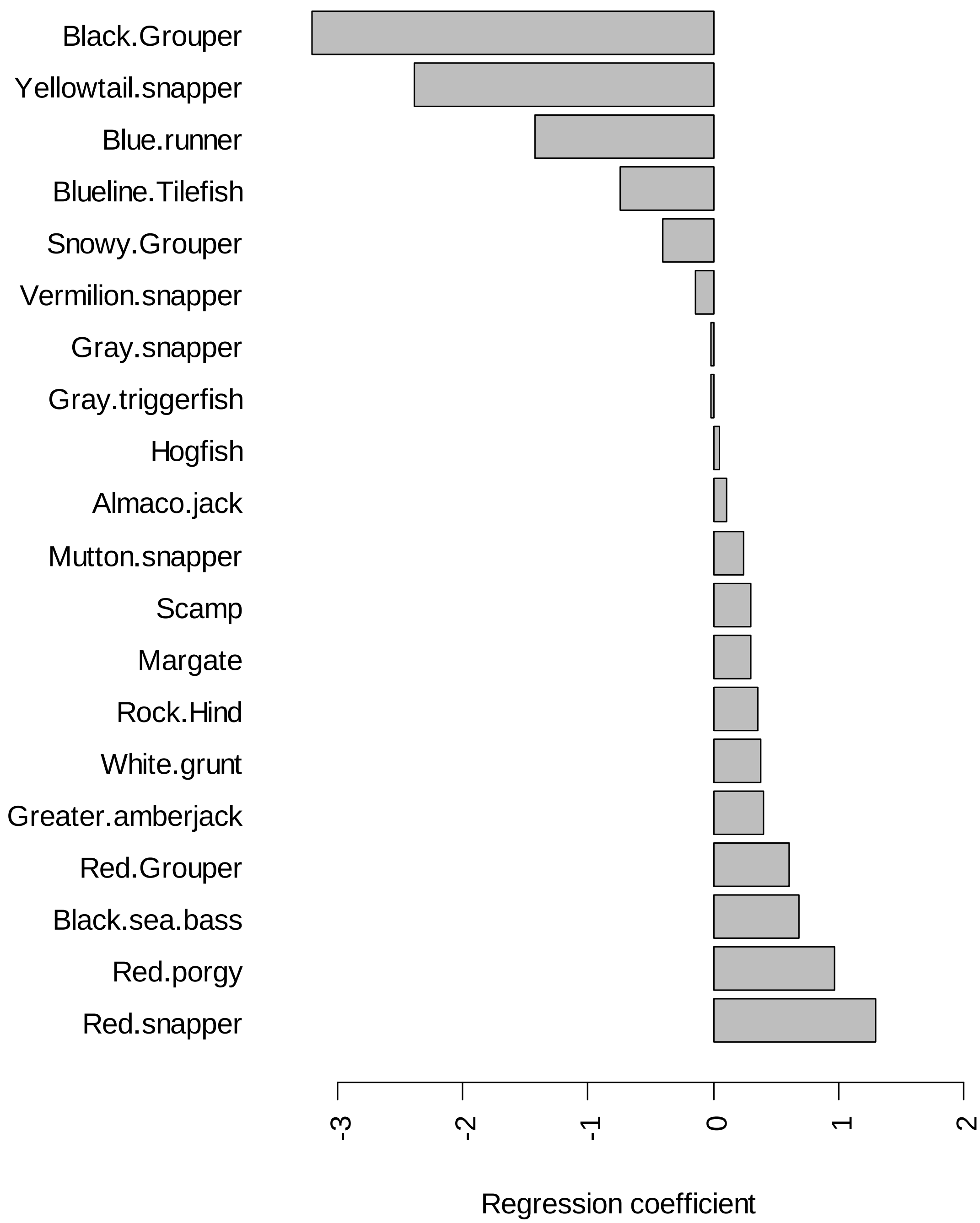


Figure 2. Absolute difference between observed and predicted number of positive gag trips. Top and bottom panels differ only in the range of probabilities shown.

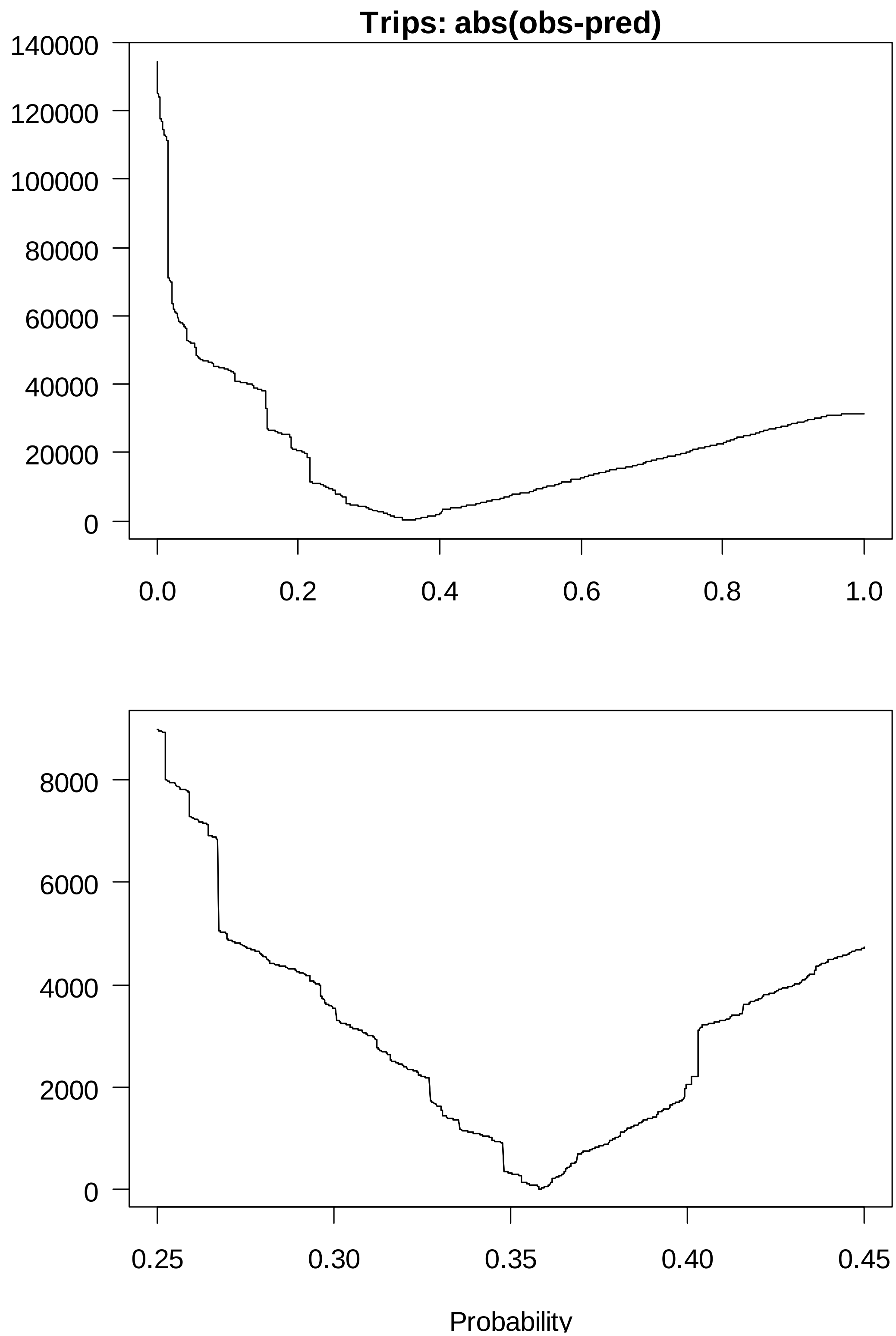


Figure 3. QQ plots of residuals from positive commercial handline trips.

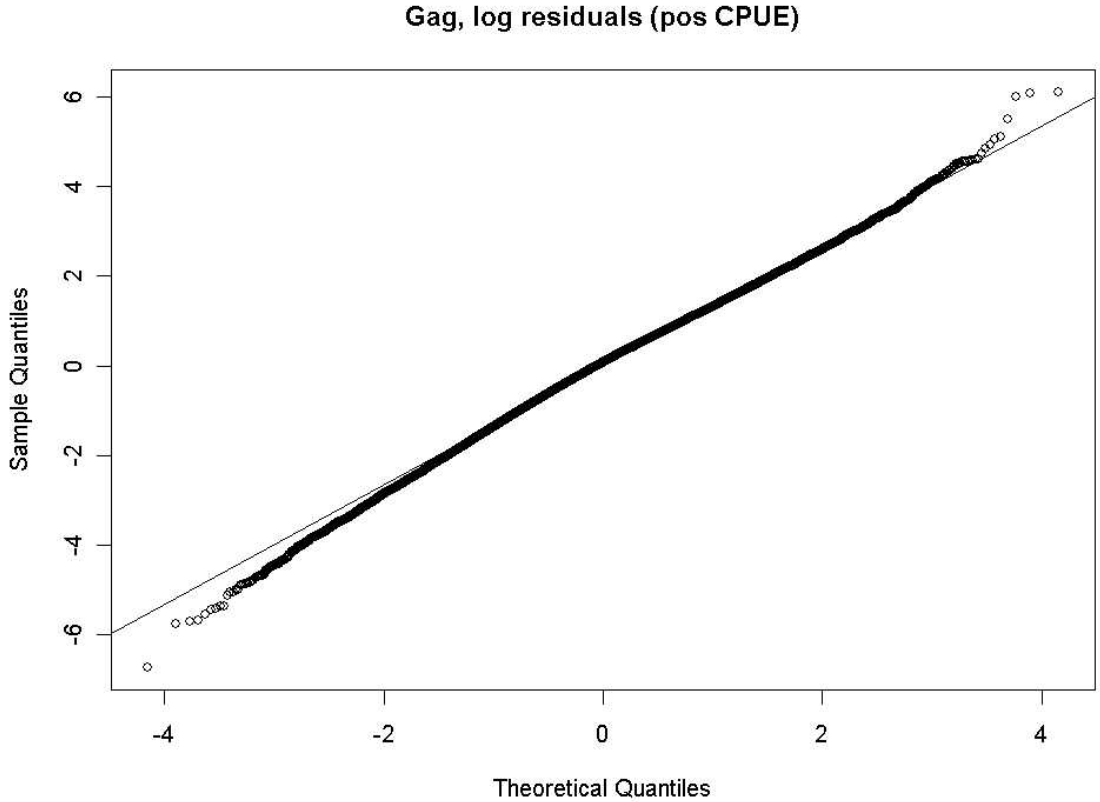
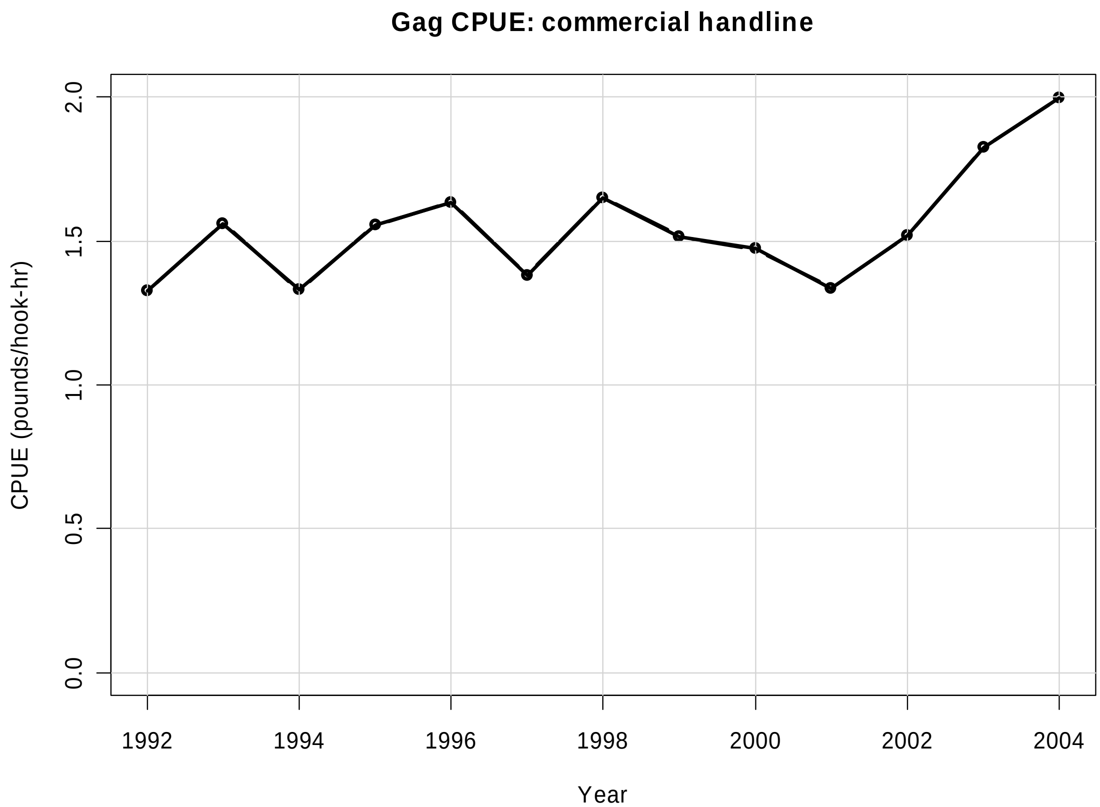


Figure 4. Index of abundance of gag from commercial handline data collected off the Southeastern U.S.



Appendix 1 The commercial logbook data set contains the following variables (all are numeric unless otherwise noted):

schedule: this is a unique identifier for each fishing trip and is a character variable
species: four-digit character variable to define species. Gag = 1423.
gear: a character variable, the gear type, multiple gear types may be used in a single trip, L = longline, H = handline, E = electric reels, B = bouy gear, GN = gill net, P = diver using power head gear, S = diver using spear gun, T = trap, TR = trolling
area: area fished, in the south Atlantic these codes have four digits- the first two are degrees of latitude and the second two are the degrees of longitude
conversion: conversion factor for calculating total pounds (totlbs) from gutted weight
gutted: gutted weight of catch for a particular species, trip, gear, and area
whole: whole weight of catch for a particular species, trip, gear, and area
totlbs: a derived variable that sums the gutted (with conversion factor) and whole weights, this is the total weight in pounds of the catch for a particular species, trip, gear, and area
length: length of longline (in miles) or gill net (in yards)
mesh1 – mesh4: mesh size of traps or nets
numgear: the amount of a gear used, number of lines (handlines, electric reels), number of sets (longlines), number of divers, number of traps, number of gill nets
fished: hours fished on a trip, this is problematic for longline data as discussed later
effort: like numgear, the data contained in this field depends upon gear type; number of hooks/line for handlines, electric reels, and trolling; number of hooks per longline for longlines; number of traps pulled for traps; depth of the net for gill nets, this field is blank for divers
source: a character variable, this identifies the database that the record was extracted from, sg = snapper grouper, grf = gulf reef fish, all records should have this source code
tif_no: a character variable, trip identifier, not all records will have a tif_no
vesid: a character variable, a unique identifier for each vessel
started: numeric (mmddyy8) variable, date the trip started
landed: numeric (mmddyy8) variable, date the vessel returned to port
unload: numeric (mmddyy8) variable, date the catch was unloaded
received: numeric (mmddyy8) variable, date the logbook form was received from the fisherman
opened: numeric (mmddyy8) variable, date the logbook form was opened and given a schedule number
away: number of days at sea, this value should equal (landed-started+1)
crew: number of crew members, including the captain
dealer: character variable, identifier for the dealer who bought the catch, in some cases there may be multiple dealers for a trip
state: character variable, the state in which the catch was sold

county: character variable, the county in which the catch was sold
area1 – area3: areas fished, if the trip included catch from multiple areas, those areas will be listed here
trip_ticke: character variable, trip ticket number, a unique identifier for each trip not all trips have this identifier.

Appendix 2: Updates since the DW

Issues discussed at the DW:

Issue 1: Trip selection using method of Stephens and MacCall (2004)
Option 1: Include all positive trips and use Stephens and MacCall method to identify zero trips only.
Option 2: Include only those trips with associated probability of catching gag above the threshold probability, as in Stephens and MacCall (2004).
Decision: Option 2, to be consistent with the published method and to exclude trips with incidental catches of gag.

Estimated coefficients are in Figure A2.1, and threshold probability in Figure A2.2. After changes in data described below, the Stevens and MacCall method selected 31553 trips, of which 24899 (78.9%) were positive.

Issue 2: Misidentification of gag as black grouper
Option 1: Take data as reported
Option 2: Devise a correction method to achieve landings consistent with proportions of species as indicated by TIP data. The method would need to be applied on a trip by trip basis.
Option 3: Exclude problematic areas. For other areas where black grouper are known to be rare, convert all landings reported as black grouper to gag.
Decision: Option 3. Much effort was devoted to achieving option 2, however, an acceptable method for correcting the landings could not be developed during the DW given available data. Option 3 was chosen because it corrects many records believed to be errors, with little chance of introducing new errors (i.e., converting black grouper to gag incorrectly).

The ratio of gag to gag plus black grouper was examined for possible misidentification. The logbook data indicated a substantial proportion of gag misidentified as black grouper (Figure A2.3). These fish are believed to be misidentified, because black grouper are rare north of Cape Canaveral, FL, a notion consistent with the General Canvass data (commercial landings reported by dealers). To correct misidentification, option 3 above was implemented, by excluding areas south of 29 degrees latitude (near Cape Canaveral) and converting all reported black grouper to gag in areas equal to and north of 29 degrees. A map with logbook areas is shown in Figure A2.4.

Issue 3: Interaction terms in the delta-GLM
Option 1: Include only main effects
Option 2: Investigate interaction terms
Decision: Option 2. Investigate interaction terms. The group decided not to include interactions with year effects, because such effects may be inseparable from changes in abundance.

Miscellaneous decisions

•Exclude months of March and April from all years in the analysis, because of bag limits that started in 1999.

•Include areas 2482 and 2382 in the Atlantic, because of council boundaries. Due to the decision on issue 2 (above), however, these areas were not used in the analysis, because they are south of 29 degrees latitude.

Updated analyses

A forward stepwise approach was used to construct each GLM (binomial and lognormal). First a GLM was fit on year. These results reflect the distribution of the nominal data. Next, each main effect (area and month) was examined for its reduction in deviance per degree of freedom. The factor that caused the greatest reduction was added to the base model if it was significant based on a Chi-Square test ($\chi^2 \leq 0.05$) and if the reduction in deviance was greater than 1%. This model then became the base model. The process was repeated, adding main effects first and then two-way interaction terms, until no factor or interaction met the criteria for inclusion.

The iterative method above requires adequate sample sizes per year per factor, and an approximately balanced design. To achieve these requirements, areas were combined as follows:

2900≤area<3000	3300≤area≤3377
3000≤area<3100	3378≤area<3400
3100≤area<3200	3400≤area≤3476
3200≤area≤3278	3477≤area<3500
3279≤area≤3300	3500≤area<3700

The forward stepwise approach identified area as the only factor other than year to be used in the binomial GLM (Table A2.1), and it identified area, month, and area*month interaction as factors to be used in the lognormal GLM (Table A2.2). Estimates of CPUE and CV are presented in Table A2.3 and in Figure A2.5. Diagnostics plots are in Figure A2.6.

Table A2.1 Linear regression statistics for the final GLM model on proportion positive trips.

Source	%RED DEV/DF	CHISQ	Pr>ChiSq
YEAR	NA	118.32	<0.0001
AREA	5.48	1783.28	<0.0001

Table A2.2 Linear regression statistics for the final GLM model on catch rates of positive trips.

Source	%RED DEV/DF	CHISQ	Pr>ChiSq
YEAR	NA	192.13	<0.0001
AREA*MONTH	22.38	6408.32	<0.0001
AREA	20.29	5655.87	<0.0001
MONTH	2.4436	270.06	<0.0001

Table A2.3. Estimated CPUE (pounds/hook-hr) of gag off the Southeastern U.S., including lower (LCI) and upper (UCI) 95% confidence intervals and CV. Estimates based on handline gear reported in commercial logbooks.

YEAR	CPUE (lb/hook-hr)	Relative CPUE	LCI	UCI	CV
1992	1.505	0.908	0.797	1.034	0.065
1993	1.566	0.944	0.868	1.027	0.042
1994	1.505	0.907	0.835	0.986	0.041
1995	1.553	0.937	0.862	1.017	0.041
1996	1.660	1.001	0.924	1.085	0.040
1997	1.274	0.768	0.703	0.839	0.044
1998	1.577	0.951	0.872	1.037	0.043
1999	1.686	1.017	0.926	1.116	0.047
2000	1.512	0.912	0.823	1.009	0.051
2001	1.438	0.867	0.791	0.951	0.046
2002	1.668	1.006	0.917	1.103	0.046
2003	2.226	1.342	1.223	1.473	0.046
2004	2.388	1.440	1.313	1.579	0.046

Figure A2.1. Estimates of species-specific regression coefficients used to estimate a trip's probability of catching gag.

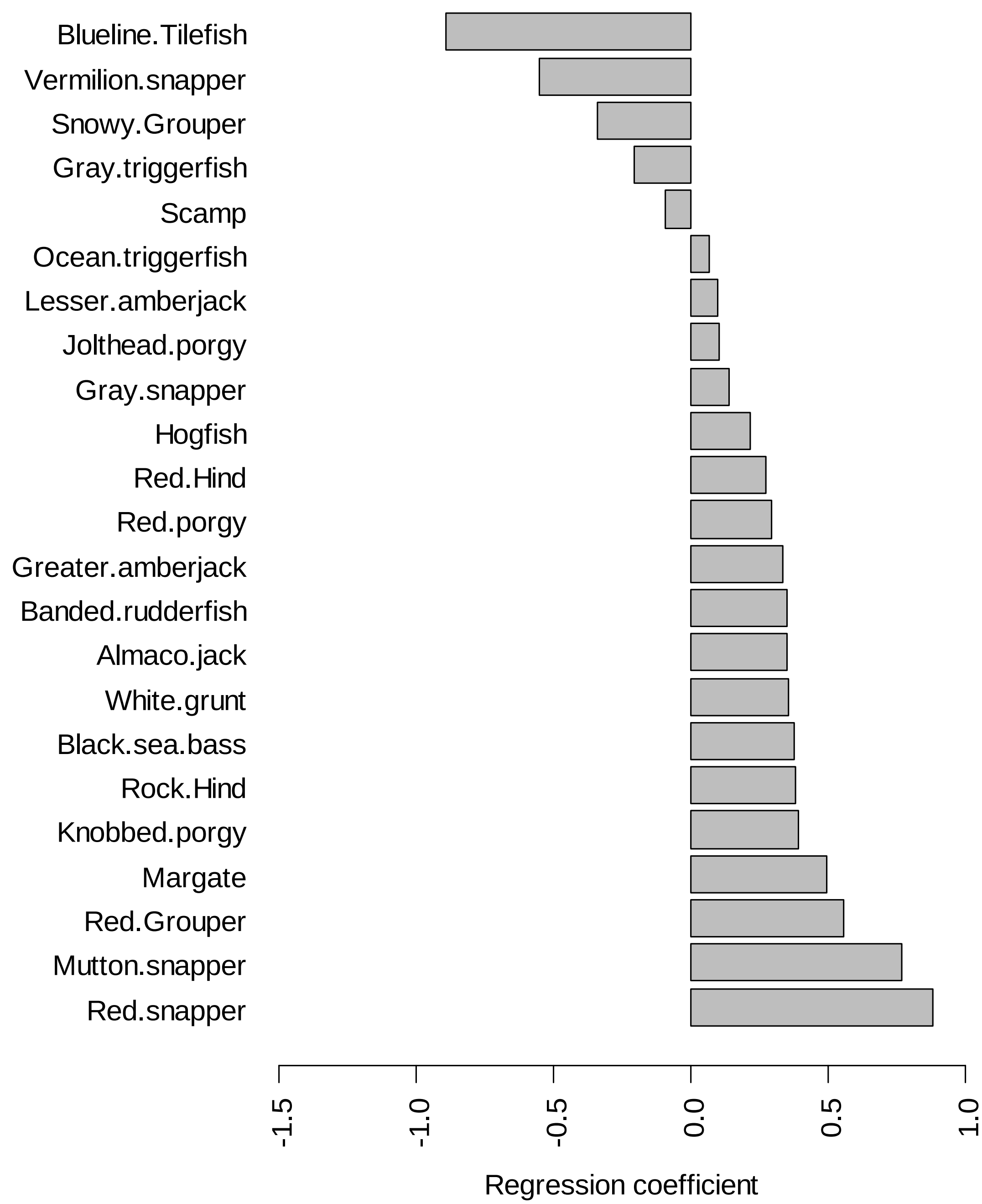


Figure A2.2 Absolute difference between observed and predicted number of positive gag trips. Top and bottom panels differ only in the range of probabilities shown.

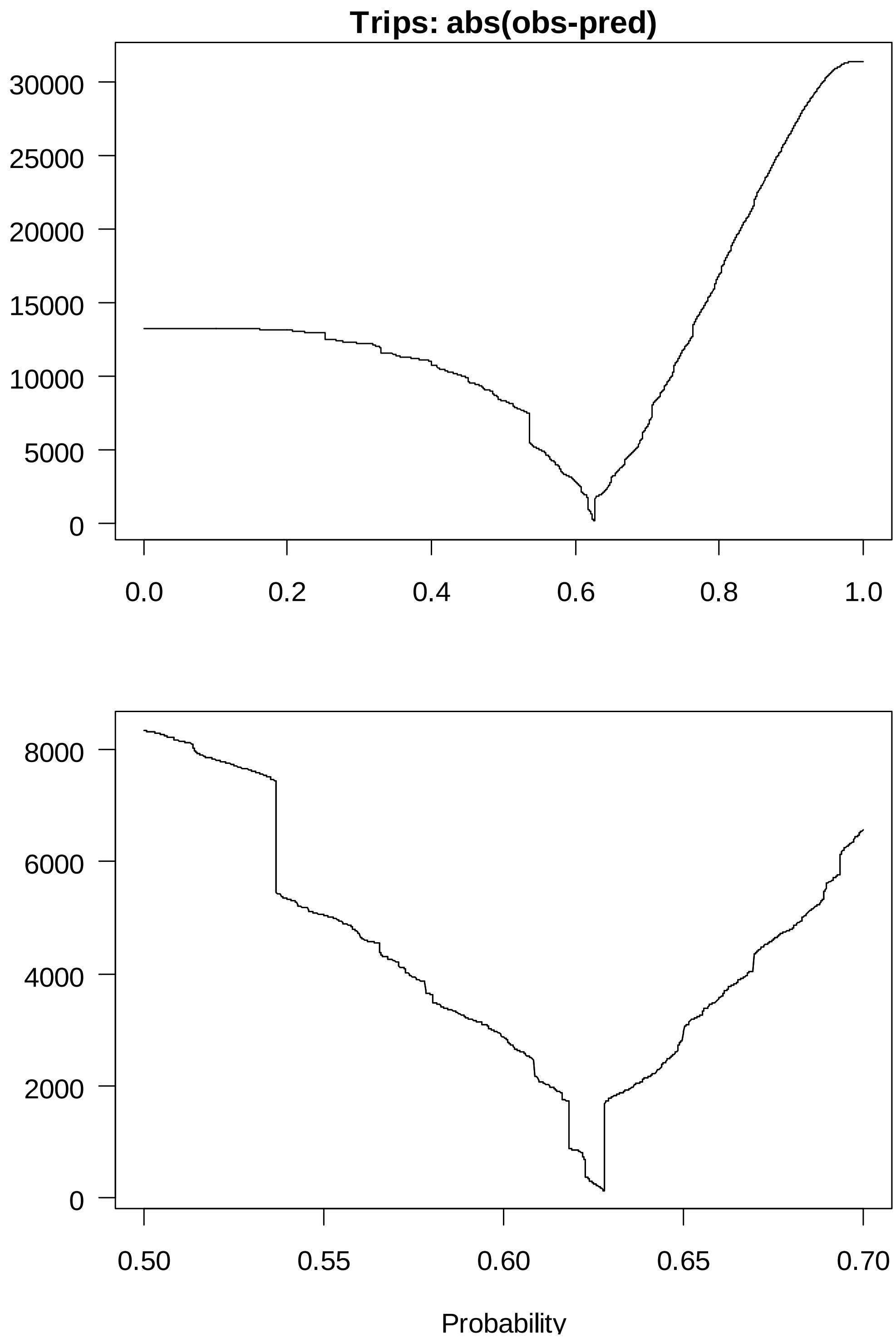


Figure A2.3. Misidentification of gag as black grouper by area in commercial logbooks (handline gear).

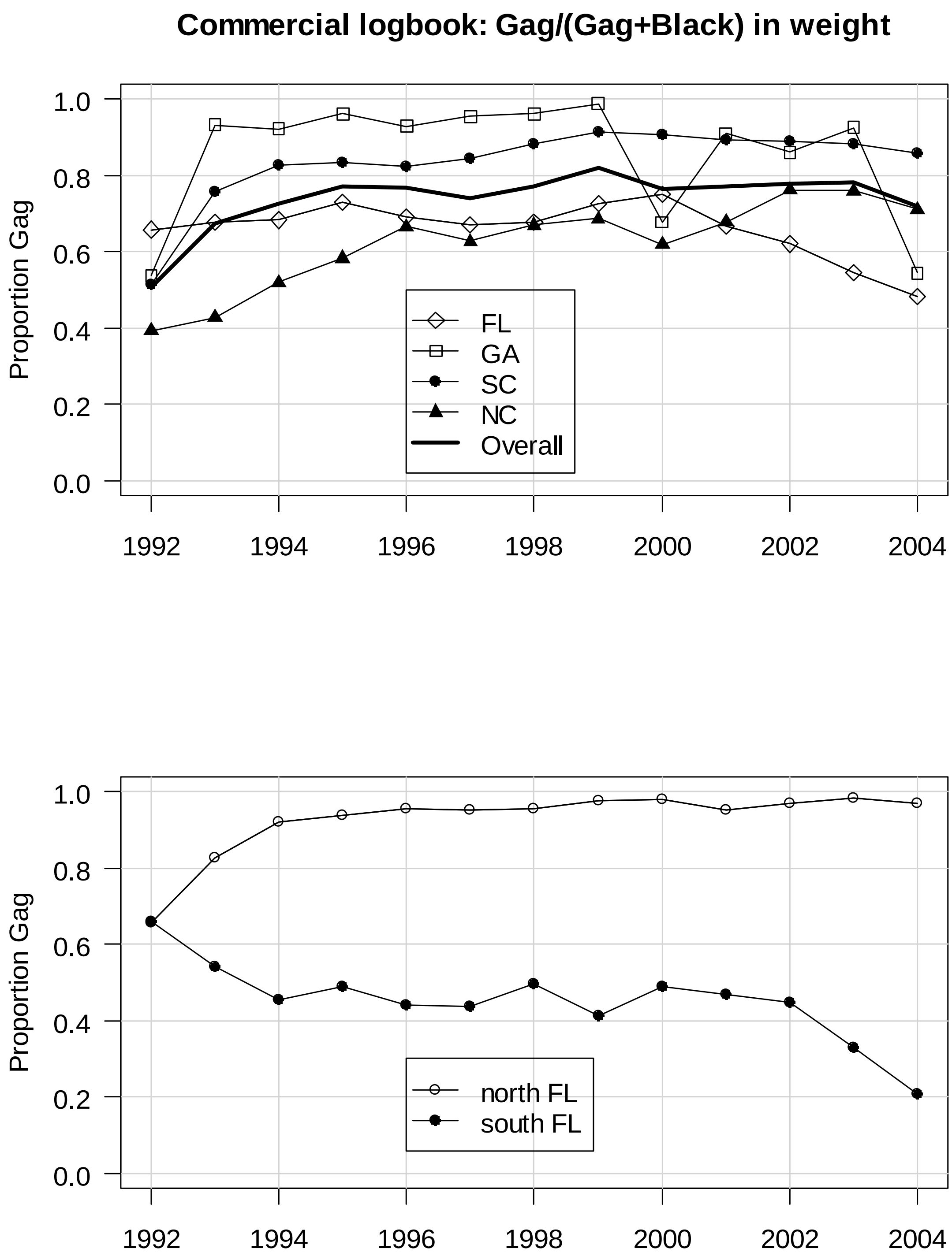


Figure A2.4. Logbook areas.

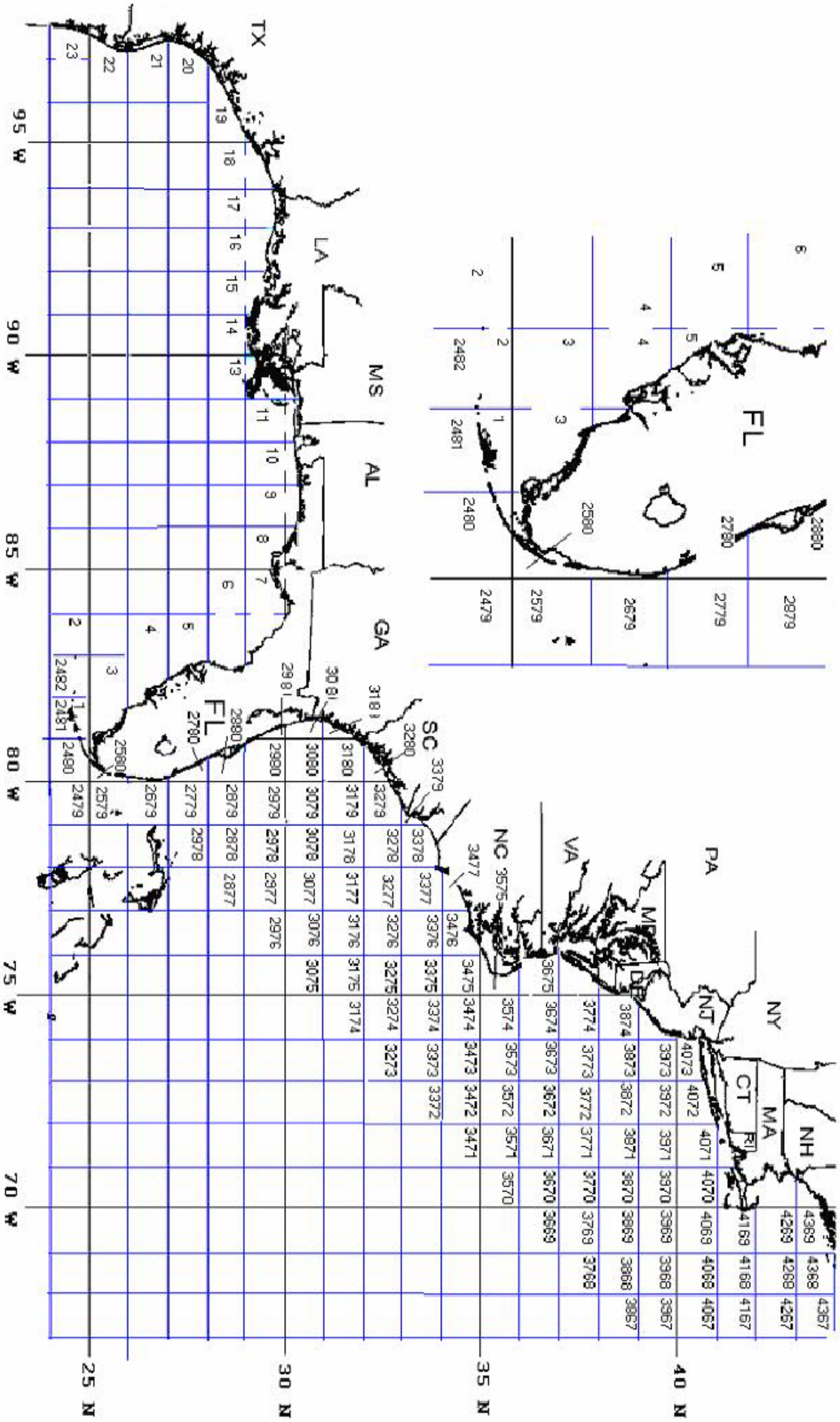


Figure A2.5. Index of abundance for gag off the Southeastern U.S. Estimates based on handline gear reported in commercial logbooks.

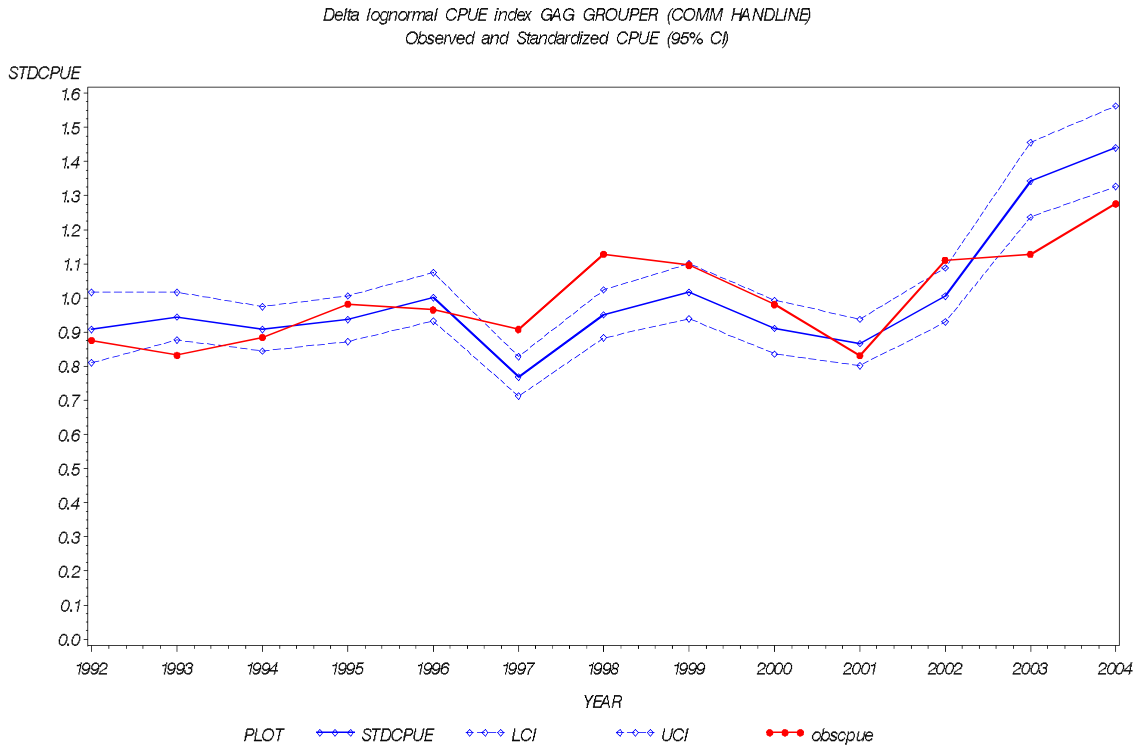
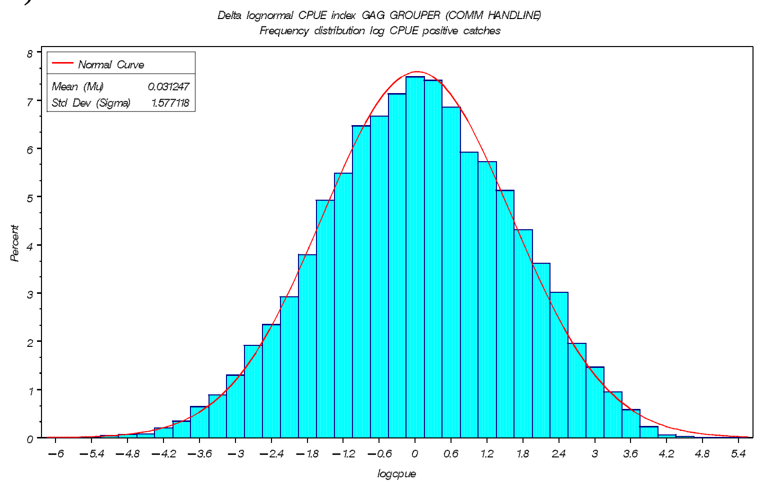
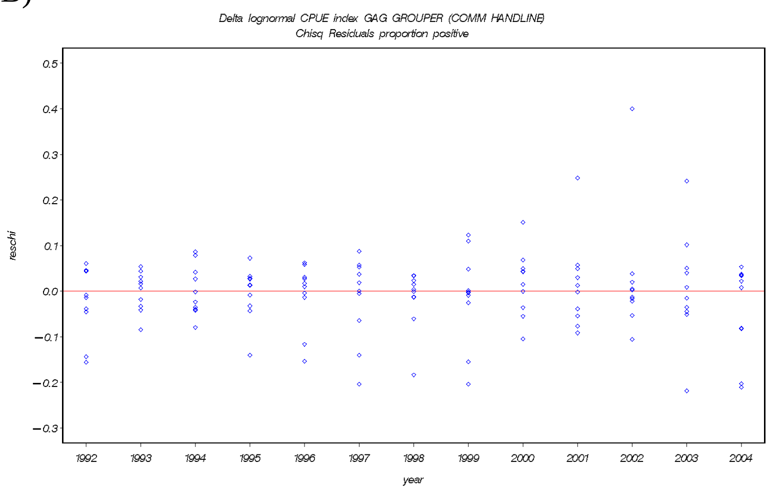


Figure A2.6 Diagnostics of model fit.

A)



B)



C)

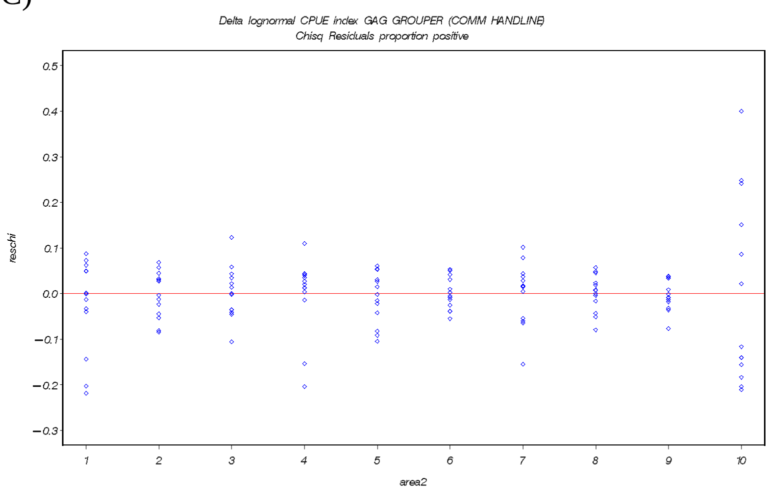
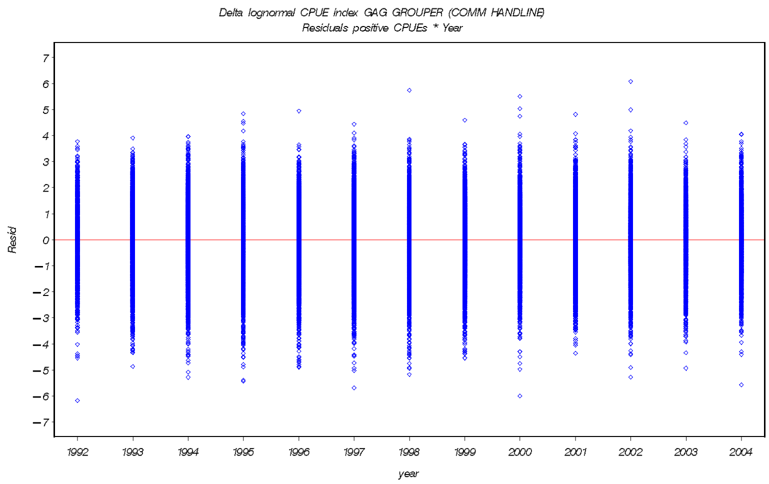
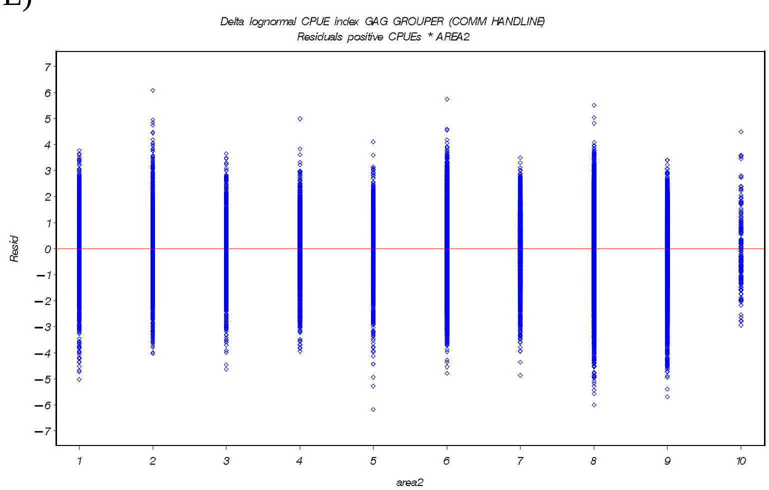


Figure A2.6 (cont.)
D)



E)



F)

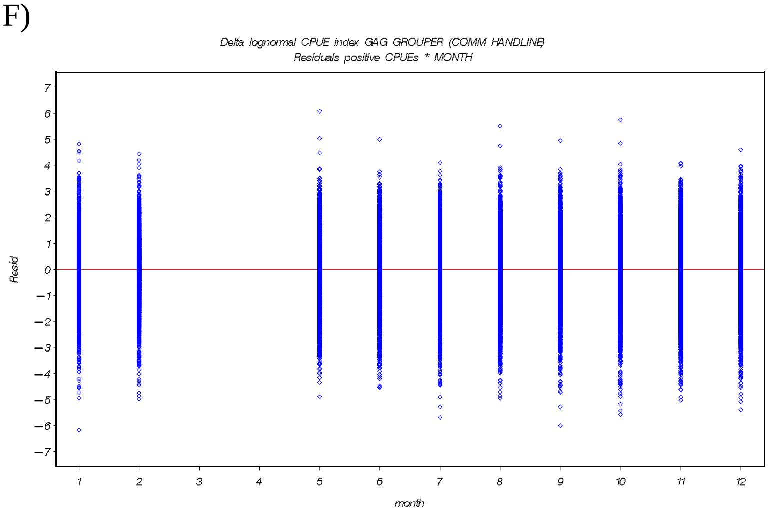


Figure A2.6 (cont.)
G)