Recreational catch per unit effort of hogfish (*Lachnolaimus maximus*) in the Southeast US using MRFSS-MRIP intercept data, 1991-2012.

Wade Cooper
Florida Fish and Wildlife Conservation Commission
Fish and Wildlife Research Institute
100 8th Avenue SE, St. Petersburg, FL 33701

April 11, 2014

Introduction

Indices of relative abundance were developed from MRFSS and MRIP recreational intercepts for the two primary hogfish stocks for which suitable data were available: Western Florida (WFL) and Southeast Florida including the Florida Keys (FLK/SEFL). Because few intercepts were available from Georgia through North Carolina (106 intercepts across states, years, and gear types; Table 1), a separate index of abundance was not developed for this stock. These stock delineations are based on genetic analyses conducted on sampling from the WFL through North Carolina, suggesting little if any contemporaneous exchange between three distinct geographic groupings (WFL, FLK/SEFL, and N. Carolina; Seyoum et al. 2014).

Methods

Spatial and Temporal Extent

Given the distribution and stock structure of hogfish, two separate indices were constructed: one for the WFL and FLK/SEFL. For the two Florida-centered stocks, only those intercepts from counties in the core distribution area were used: Franklin to Collier for the WFL stock, and Monroe to Indian River for the FLK/SEFL stock (Figure 1). MRFSS intercepts were used for the period of 1981-2003, while MRIP was used for 2004-2012. Only those years from 1991-2012 were used for the construction of the index because the party code only began to be recorded in 1991. Prior to 1991, interviews done on multiple individuals from the same trip could not be distinguished, and therefore should not be included in the standardization procedure.

Identification of Appropriate Surveys

While catching hogfish on hook and line is rare, the total number of trips catching hogfish using spear versus hook and line are similar (Table 1). The use of hook and line data during the previous hogfish assessment (SEDAR 6) was a contentious issue with the reviewers due to the rarity of catch (Kingsley 2004). In the previous assessment, any trips designated as reef fish trips were used as appropriate surveys. For this updated analysis, we separate spear versus hook-and-line as two separate fisheries and subsequent indices of abundance, and additionally filter the total reef fish trips by identifying those species caught in association with hogfish using a cluster analysis. By identifying those trips that caught associated species but failed to catch hogfish, one can infer zero-catch trips that were appropriate to include in the analysis (Stephens and MacCall

2004). Different approaches exist to identify the associated species and subsequent trips, including logistic regression techniques (Stephens and MacCall 2004) and multivariate clustering (Shertzer and Williams 2008; Muller 2009, O'Hop et al. 2012). Affinity propagation clustering (APC) was chosen to determine associated species, because it has been shown to perform well relative to other cluster techniques and does not require that the number of cluster be pre-specified (Frey and Dueck 2007). APC automatically chooses an optimal number of clusters in the dataset, thereby providing an objective criterion for which to group associated species. To conduct the cluster analysis, the data were first filtered to remove all uncommon species that occurred on only a small proportion of the total trips. For hook and line, species caught on less than 0.1% of the trips were removed, while species caught on less than 1% of the trips were removed for spear fishing trips. The lower cutoff was used for hook and line due to hogfish being rare on hook and line trips (<1% of trips). The APC procedure was then applied using the Morisita measure of similarity, since this measure is recommended for count data and is insensitive to sample sizes (Krebs 1999). Once the associated species within the hogfish cluster were identified for each of the stocks and gear types, all trips on which these species were caught for a specific gear type were used as suitable trips in the subsequent analyses. The APC technique was done in R 3.0.1 (R Core Team 2013) using the apcluster package (Bodenhofer et al. 2011).

Standardization Model

Standardized indices of abundance were calculated using a generalized linear modeling procedure that combined the analysis of the binomial information on presence/absence with the lognormal-distributed positive catch data (also known as two-part, hurdle, or zero-adjusted models, Zuur et al. 2009) as:

$$I_y = c_y p_y \quad [1]$$

where c_y are estimated annual mean CPUEs of non-zero catches modeled as lognormal distributions and p_y are estimated annual mean probabilities of capture modeled as binomial distributions. The lognormal submodel considers only trips in which a hogfish was caught (*i.e.*, non-zero catches). The binomial model considers all trips in which hogfish or associated species were caught. While other approaches exist to model zero-inflated data (i.e., Poisson and negative binomial distributions; zero-inflated models; Zuur et al. 2009), the two-part model used here is advantageous in that it provides inference on both the presence-absence and abundance processes occurring within a population, and can easily accommodate different predictor variables for each sub-model in the statistical analysis.

To determine the most appropriate models, predictor variables were selected using a forward step-wise approach where each predictor was added to each submodel individually and the resulting reduction in deviance per degree of freedom (Dev/DF) analyzed. The factor causing the greatest reduction in Dev/DF was then added to the base model. Year was retained in all models to obtain an index of abundance over time. Other potential predictors included wave, mode, area, hours fished, number of anglers, avidity, and time fished. We assume that there are no significant interaction terms with year in this model and consider only the main effects. Criteria for model inclusion also include a reduction in Dev/DF≥0.05%. This process was then repeated until no factor met criteria for model inclusion. Final year-specific marginal means estimates and standard errors of the two sub-models were used to generate distributions of estimates for each sub-model from a Monte Carlo simulation (5000 Student's t distributed realizations). The product of these distributions (eq. 1) provided an estimate of the median catch rate with year-specific variability. All analyses were done using R 3.0.1 (R Core Team 2013).

Results and Discussion

Identification of Appropriate Surveys

The APC technique was performed separately for the two gear and two stocks (hook-and-line versus spear; WFL versus FLK/SEFL stocks). For the WFL spear, the APC procedure selected 5 clusters from a total of 18 species. The species group in which hogfish clustered comprised the largest cluster with a total of six other species (Table 2). For the WFL hook and line, the APC procedure selected 14 total clusters from a total of 63 species. The species group in which hogfish clustered included five other species (Table 2). For the FLK/SEFL spear, the APC procedure selected 6 clusters from a total of 28 species. The species group in which hogfish clustered included four other species (Table 2). And lastly, for the FLK/SEFL hook and line, the APC procedure selected 23 clusters from a total of 113 species. The species group in which hogfish clustered included four other species (Table 2). Figure 2 presents the frequencies of hogfish caught per trip for the two stocks and gear types after filtering for only those trips expected to encounter a hogfish (i.e., those trips either catching a hogfish or the associated species).

Standardization Model

The results from the forward-stepwise model selection procedure are presented in Tables 3-10. The final predictor variables for each model component (binomial and positives model components, two stocks, and two gear types: eight total models) were those that explained greater than 0.5 % of the residual deviance/DF in the deviance tables (percent.reduction column). Two of the models (binomial component of the WFL spear model, and positives component of the WFL hook-and-line) did not have any factors that improved the model over the null model with just year as a predictor (Tables 3, 6). Figures 3-10 present the diagnostics plots for each of the eight component models. In general, the models for spearfishing gear had relatively good fits to the positives data using a lognormal distribution (QQ plots are approximately normal), while the hook-and-line component models performed poorly. Alternative modeling approaches (e.g., zero-inflated) or distributions (e.g., negative binomial) could be attempted for the hook-and-line data in the future to see if they improve the fit; however, these adjustments would likely have a minimal impact on the final index of abundance, particularly given the relatively high variability in the final indices.

The indices of abundance are presented in Tables 11-14 and Figures 11-14. Overall the indices were highly variable with average coefficients of variation (CVs) of 36, 49, 22, and 27 % for the WFL spear, WFL hook-and-line, FLK/SEFL spear, and FLK/SEFL hook-and-line, respectively. The WFL spear index was generally stable from 1992-2005, but has increased steadily since 2005. This pattern was not evidence in the WFL hook-and-line index, where the hook-and-line index was generally flat but marked with large peaks in abundance for 2003 and 2010. The FLK/SEFL spear and hook-and-line indices were relatively similar in that both suggest a decreasing trend over time.

References

Bodenhofer U, A Kothmeier, and S Hochreiter. 2011. APCluster: an R package for affinity propagation clustering. Bioinformatics 27:2463-2464. DOI: 10.1093/bioinformatics/btr406.

Kingsley, MCS, ed. 2004. The Hogfish in Florida: assessment review and advisory report. Report prepared for the South Atlantic Fishery Management Council, the Gulf of Mexico Fisheries Management Council, and the National Marine Fisheries Service. Southeast Data and Assessment Review. vi+15pp.

R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Seyoum S, Collins AB, Puchulutegue C, McBride RS, Tringali MD. 2014. Genetic population structure of hogfish (Labridae: *Lachnolaimus maximus*) in the southeastern United States. SEDAR 37-01.

Zuur, AF, EN Ieno, NJ Walker, AA Saveliev, GM Smith. 2009. Mixed effects models and extensions in ecology with R. Springer Science+Business Media, New York NY. 574pp.

Tables

Table 1. Total number of trips catching and/or targeting hogfish by state and gear type.

State	Gear	Trips Catching Hogfish	Trips Targeting or Catching Hogfish
AL	Hook & Line	1	1
FL	Cast Net	4	9
FL	Dip Net	1	1
FL	Gill Net	3	5
FL	Hand	1	1
FL	Hook & Line	718	830
FL	Spear	877	1333
GA	Hook & Line	4	4
LA	Hook & Line	4	4
MS	Hook & Line	4	4
NC	Hook & Line	82	82
NC	Spear	2	2
SC	Hook & Line	10	10

Table 2. Species clusters for the two stocks (West Florida, WFL; Florida Keys and Southeast Florida, FLK/SEFL) and gear types (spearfishing, hook-and-line) used to select those trips where a hogfish was likely to occur.

WFL Spear	WFL HL	FLK/SEFL Spear	FLK/SEFL HL
COBIA	HOGFISH	BLACK GROUPER	AFRICAN POMPANO
GAG	KNOBBED PORGY	BLACK MARGATE	BAR JACK
GRAY SNAPPER	LANE SNAPPER	CREVALLE JACK	HOGFISH
GREATER AMBERJACK	LITTLEHEAD PORGY	HOGFISH	KNOBBED PORGY
HOGFISH	TOMTATE	MUTTON SNAPPER	WHITEBONE PORGY
PINFISH	VERMILION SNAPPER		
SPANISH MACKEREL			

Table 3. Deviance table for the binomial component of the WFL spear model. The null model with year as a predictor is listed as step 0, and subsequent steps list the most predictive factors.

Step	Variable	Deviance	Resid. Df	Resid. Dev	AIC	percent.reduction
0	year	NA	388	483.5597	525.5597	0

Table 4. Deviance table for the positives component of the WFL spear model. The null model with year as a predictor is listed as step 0, and subsequent steps list the most predictive factors.

Step	Variable	Deviance	Resid. Df	Resid. Dev	AIC	percent.reduction
0	year	NA	233	185.0968	684.4411	0
1	num_anglers	22.53864	231	162.5581	655.4609	11.4163
2	hr_fished	4.292282	228	158.2658	654.664	1.204219

Table 5. Deviance table for the binomial component of the WFL hook-and-line model. The null model with year as a predictor is listed as step 0, and subsequent steps list the most predictive factors.

Step	Variable	Deviance	Resid. Df	Resid. Dev	AIC	percent.reduction
0	year	NA	2530	1307.467	1351.467	0
1	hr_fished	26.52506	2523	1280.942	1338.942	1.756918

Table 6. Deviance table for the positives component of the WFL hook-and-line model. The null model with year as a predictor is listed as step 0, and subsequent steps list the most predictive factors.

Step	Variable	Deviance	Resid. Df	Resid. Dev	AIC	percent.reduction
0	year	NA	167	84.85086	430.9978	0

Table 7. Deviance table for the binomial component of the FLK/SEFL spear model. The null model with year as a predictor is listed as step 0, and subsequent steps list the most predictive factors.

Step	Variable	Deviance	Resid. Df	Resid. Dev	AIC	percent.reduction
0	year	NA	516	497.9772	541.9772	0
1	num_anglers	8.934278	514	489.0429	537.0429	1.41199
2	avidity	5.283548	512	483.7594	535.7594	0.684182
3	wave	2.203405	511	481.556	535.556	0.255208

Table 8. Deviance table for the positives component of the FLK/SEFL spear model. The null model with year as a predictor is listed as step 0, and subsequent steps list the most predictive factors.

Step	Variable	Deviance	Resid. Df	Resid. Dev	AIC	percent.reduction
0	year	NA	415	257.1642	1054.447	0
1	num_anglers	33.4255	413	223.7387	997.6007	12.57641
2	hr_fished	9.097213	410	214.6415	985.4609	2.940967

Table 9. Deviance table for the binomial component of the FLK/SEFL hook-and-line model. The null model with year as a predictor is listed as step 0, and subsequent steps list the most predictive factors.

Step	Variable	Deviance	Resid. Df	Resid. Dev	AIC	percent.reduction
0	year	NA	1173	1467.058	1511.058	0
1	wave	29.29505	1168	1437.763	1491.763	1.577323
2	avidity	14.28719	1166	1423.476	1481.476	0.810892
3	num_anglers	11.95327	1162	1411.523	1477.523	0.486478

Table 10. Deviance table for the positives component of the FLK/SEFL hook-and-line model. The null model with year as a predictor is listed as step 0, and subsequent steps list the most predictive factors.

Step	Variable	Deviance	Resid. Df	Resid. Dev	AIC	percent.reduction
0	year	NA	385	188.1038	886.8854	0
1	num_anglers	6.095999	381	182.0078	881.4771	2.224919

Table 11. Standardized index of abundance from the WFL spear model. Note: year 1991 was removed due to convergence issues.

year	Total.num.trips	Num.pos	Mean	std.dev	CV
1992	13	5	1.443826	0.788723	0.546273
1993	14	11	4.000411	1.265951	0.316455
1994	21	10	1.874579	0.707339	0.377332
1995	16	8	2.079268	0.843176	0.405516
1996	28	9	0.798433	0.337858	0.423151
1997	23	10	1.38286	0.524321	0.379157
1998	21	14	2.616298	0.769566	0.294143
1999	34	26	1.846531	0.375926	0.203585
2000	10	3	1.278246	0.982112	0.768328
2001	20	13	1.638807	0.504285	0.307715
2002	19	7	1.317326	0.612578	0.465016
2003	26	17	2.171707	0.5654	0.260348
2004	18	10	2.093937	0.755937	0.361012
2005	17	8	1.109907	0.46188	0.416143
2006	14	11	1.692072	0.527033	0.311472
2007	10	8	2.745495	1.011081	0.368269
2008	24	17	2.808413	0.725105	0.25819
2009	22	16	2.755354	0.728718	0.264473
2010	20	16	3.543814	0.945166	0.266709
2011	17	15	3.098297	0.816493	0.26353
2012	22	20	5.577909	1.269787	0.227646

Table 12. Standardized ndex of abundance for the WFL hook-and-line model.

year	Total.num.trips	Num.pos	Mean	std.dev	CV
1991	44	1	0.32336	0.55143	1.705316
1992	130	8	0.103127	0.045541	0.441597
1993	110	6	0.153604	0.078233	0.509313
1994	82	8	0.205855	0.08737	0.424427
1995	122	11	0.173212	0.064167	0.370452
1996	75	6	0.211736	0.108214	0.511079
1997	85	9	0.186695	0.075301	0.403337
1998	126	12	0.200083	0.07028	0.351252
1999	163	10	0.102653	0.040046	0.390111
2000	137	9	0.136437	0.055661	0.407963
2001	148	8	0.118422	0.052163	0.440482
2002	165	10	0.09643	0.037164	0.385404
2003	148	16	0.342653	0.10195	0.297532
2004	160	4	0.044853	0.028919	0.644763
2005	163	7	0.129531	0.060803	0.469412
2006	66	4	0.127265	0.079955	0.628251
2007	80	10	0.225016	0.085377	0.379427
2008	121	12	0.172345	0.060593	0.351579
2009	144	15	0.177461	0.055574	0.313163
2010	49	9	0.675596	0.255226	0.377778
2011	72	9	0.424696	0.167936	0.395428
2012	162	5	0.168238	0.098392	0.584842

Table 13. Standardized ndex of abundance for the FLK/SEFL spear model.

year	Total.num.trips	Num.pos	Mean	std.dev	CV
1991	7	4	1.994887	0.995293	0.498922
1992	32	24	2.865121	0.537329	0.187541
1993	19	14	2.56872	0.625195	0.243388
1994	31	30	2.548989	0.381095	0.149508
1995	14	12	2.064031	0.522786	0.253284
1996	20	15	1.868412	0.461086	0.24678
1997	14	10	1.927575	0.580389	0.301098
1998	20	15	1.928215	0.456421	0.236707
1999	24	21	2.728113	0.499895	0.183238
2000	13	9	1.991561	0.599747	0.301144
2001	19	14	2.369064	0.556748	0.235008
2002	22	16	2.516233	0.561271	0.22306
2003	26	21	2.677963	0.509953	0.190426
2004	36	30	2.098498	0.334603	0.159449
2005	24	21	2.177644	0.410965	0.18872
2006	18	16	1.611543	0.345773	0.21456
2007	38	35	2.551408	0.359466	0.140889
2008	32	27	2.383003	0.398424	0.167194
2009	34	27	2.150687	0.365185	0.169799
2010	21	16	1.663643	0.380359	0.22863
2011	25	20	1.710873	0.335661	0.196193
2012	49	40	1.607146	0.222916	0.138703

Table 14. Standardized index of abundance for the FLK/SEFL hook-and-line model.

year	Total.num.trips	Num.pos	Mean	std.dev	CV
1991	17	5	1.00428	0.490139	0.48805
1992	49	28	0.965804	0.185406	0.191971
1993	60	26	1.066079	0.225792	0.211796
1994	56	25	0.852024	0.181044	0.212487
1995	39	25	1.204035	0.242562	0.201458
1996	34	17	1.303322	0.320114	0.245614
1997	38	16	0.598802	0.157751	0.263444
1998	47	21	0.649418	0.145902	0.224666
1999	63	23	0.637007	0.149179	0.234188
2000	43	8	0.261647	0.109044	0.41676
2001	62	16	0.605467	0.171665	0.283524
2002	64	10	0.338851	0.128719	0.379869
2003	75	26	0.719481	0.152771	0.212335
2004	70	25	0.630087	0.141239	0.224158
2005	59	18	0.502286	0.132921	0.264631
2006	46	15	0.54304	0.158956	0.292715
2007	64	21	0.632424	0.150061	0.237279
2008	77	22	0.524849	0.125146	0.238443
2009	49	14	0.468333	0.141143	0.301374
2010	58	12	0.411565	0.139703	0.339443
2011	48	12	0.476065	0.157186	0.330177
2012	77	22	0.559993	0.134359	0.239929

Figures

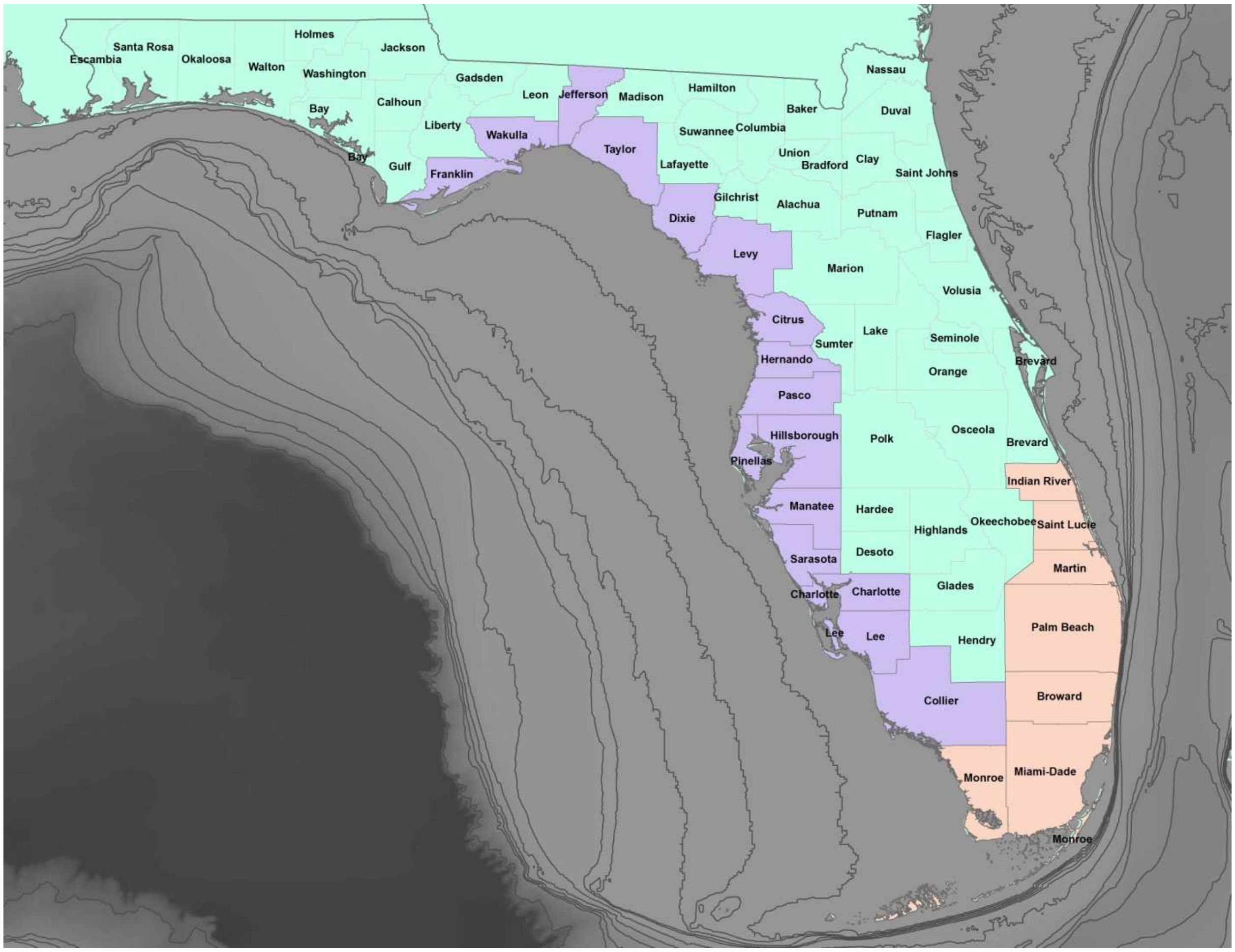

Figure 1. Florida county delineations used to represent the core distributions of the two hogfish stocks: West Florida (WFL; purple) and Southeast Florida including the Keys (FLK/SEFL; peach).

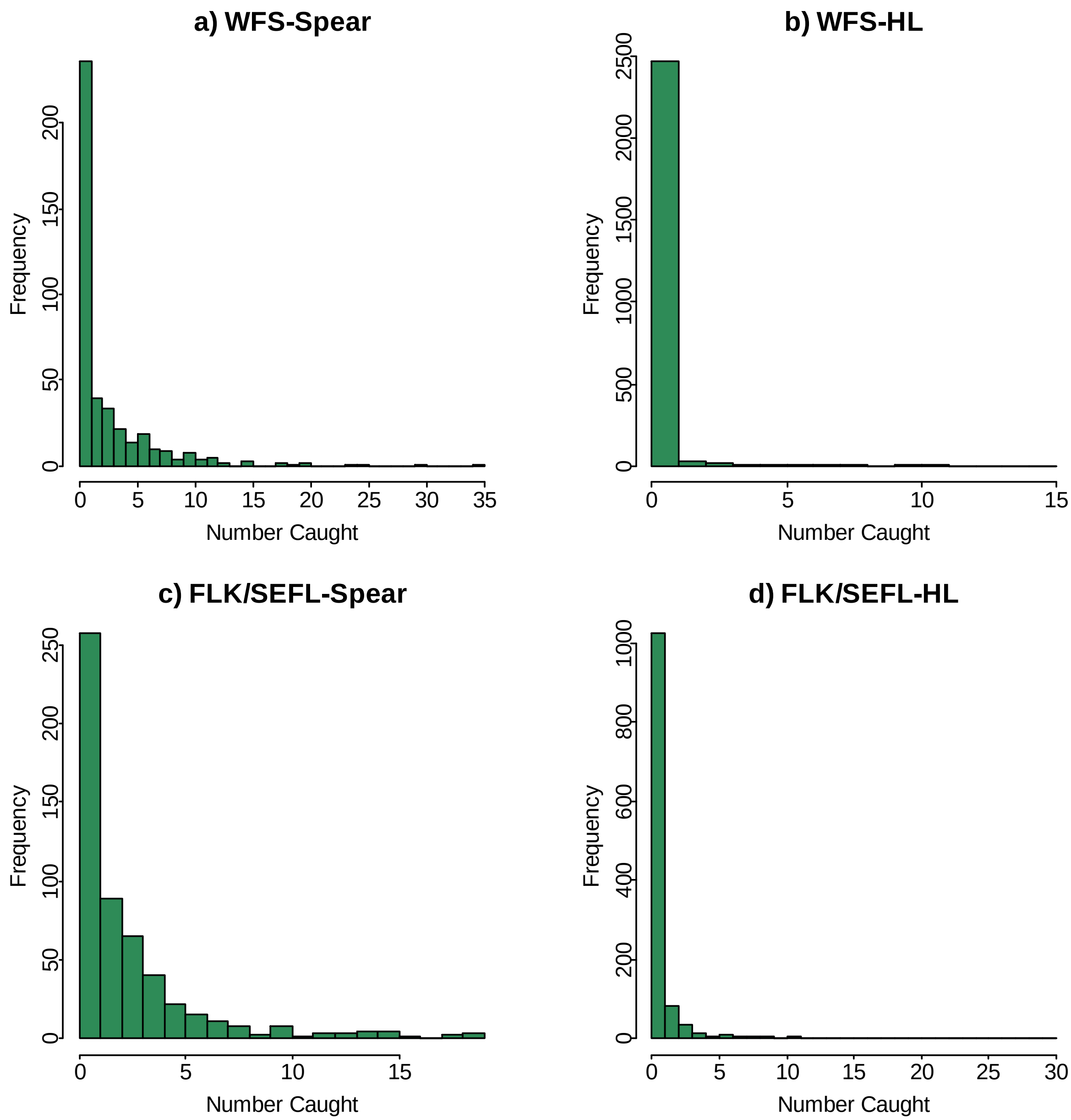

Figure 2. Frequencies for the number of hogfish caught per trip using spear fishing (a, c) and hook and line (b, d) for the WFL stock (a, b) and the FLK/SEFL stock (c, d).

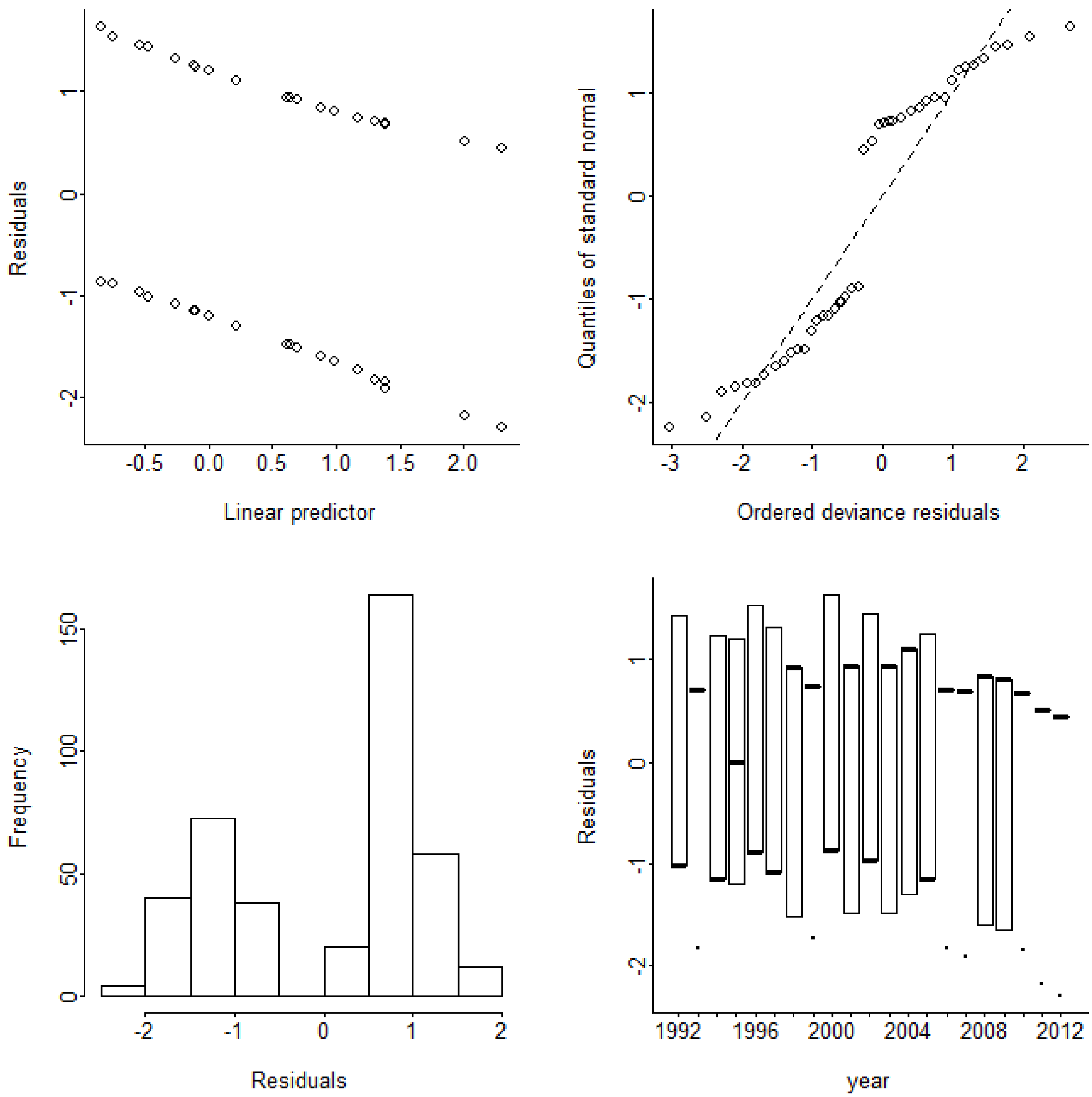

Figure 3. Diagnostic plots from the binomial component of the WFL spear model.

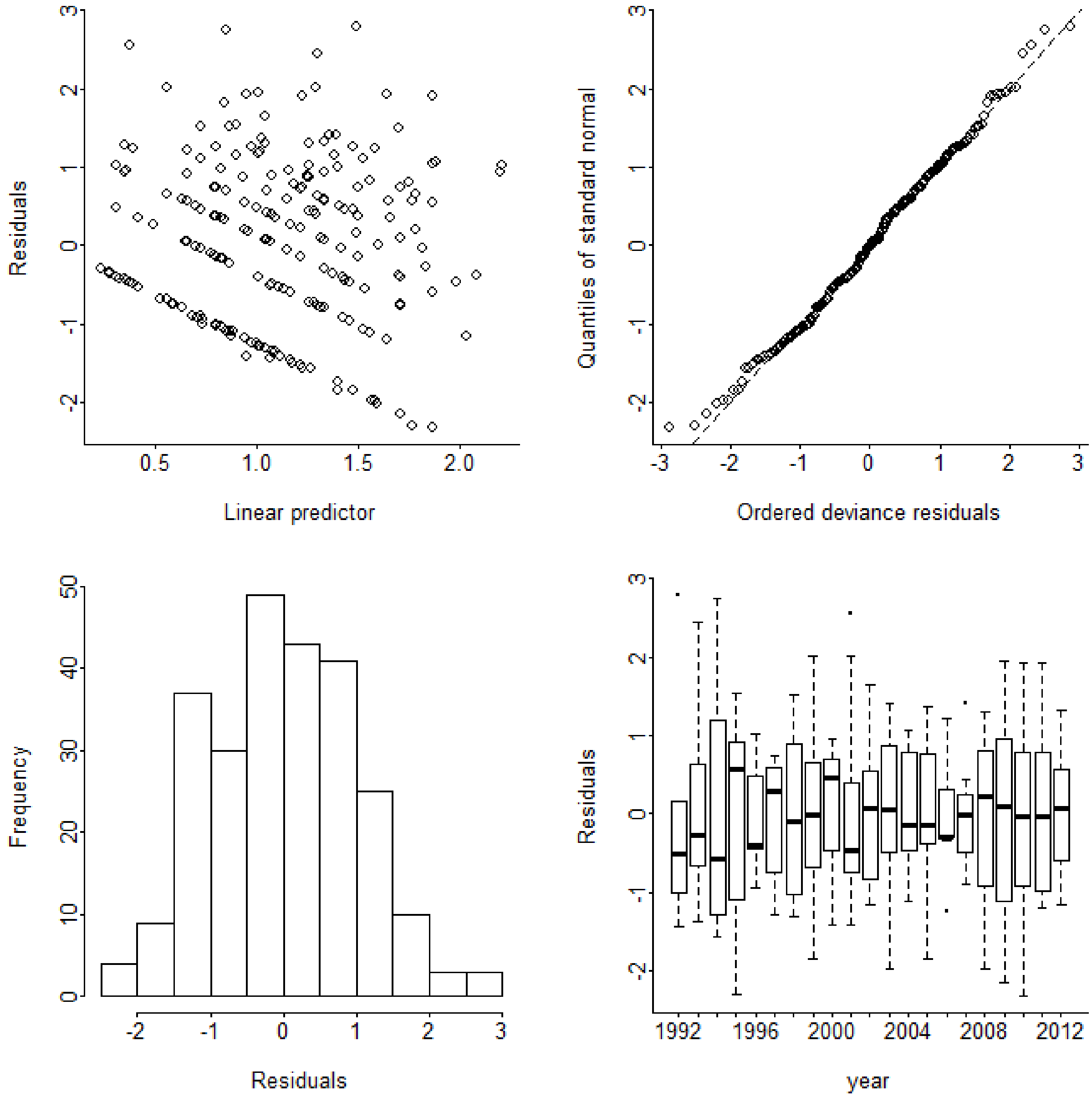

Figure 4. Diagnostic plots from the positives component of the WFL spear model.

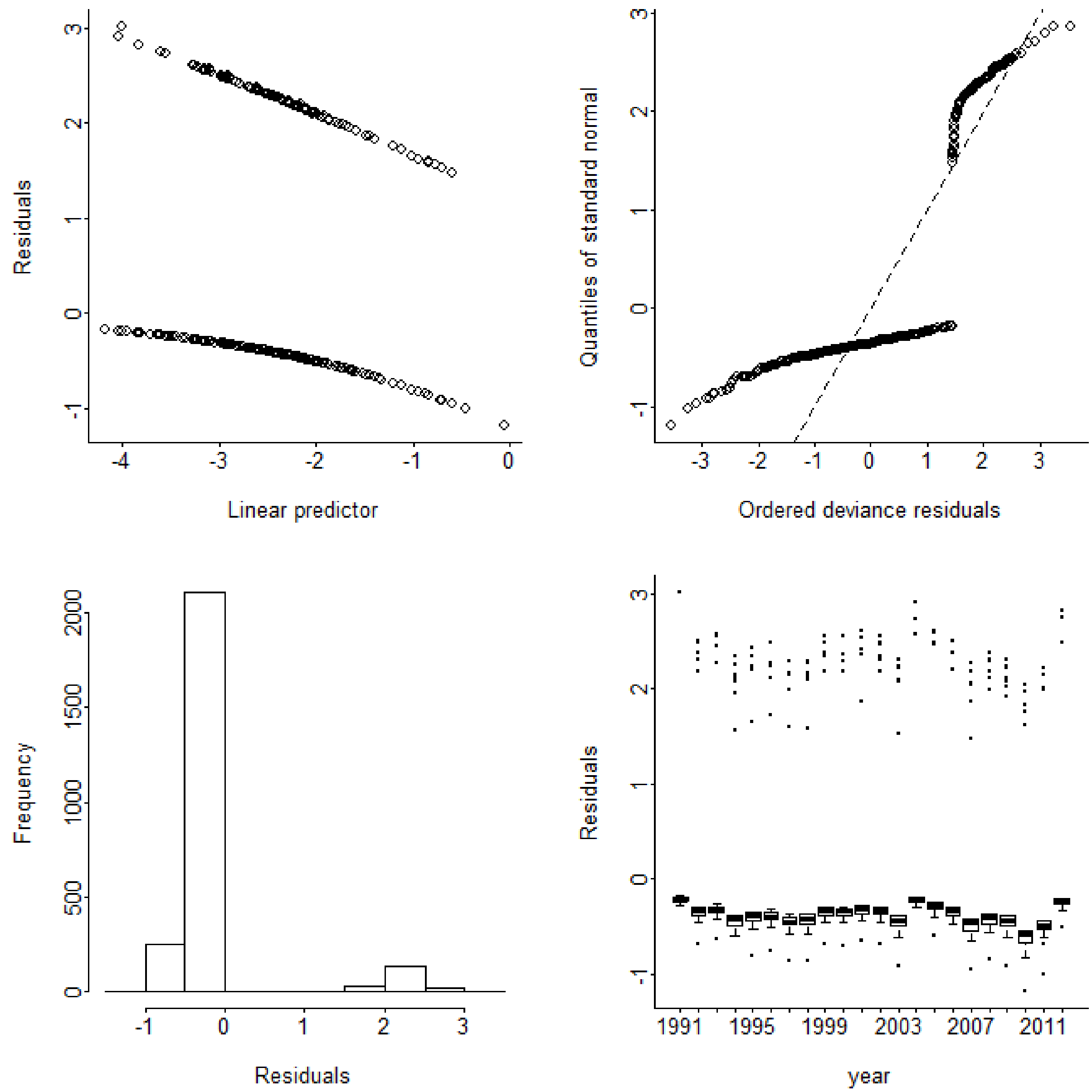

Figure 5. Diagnostic plots from the binomial component of the WFL hook-and-line model.

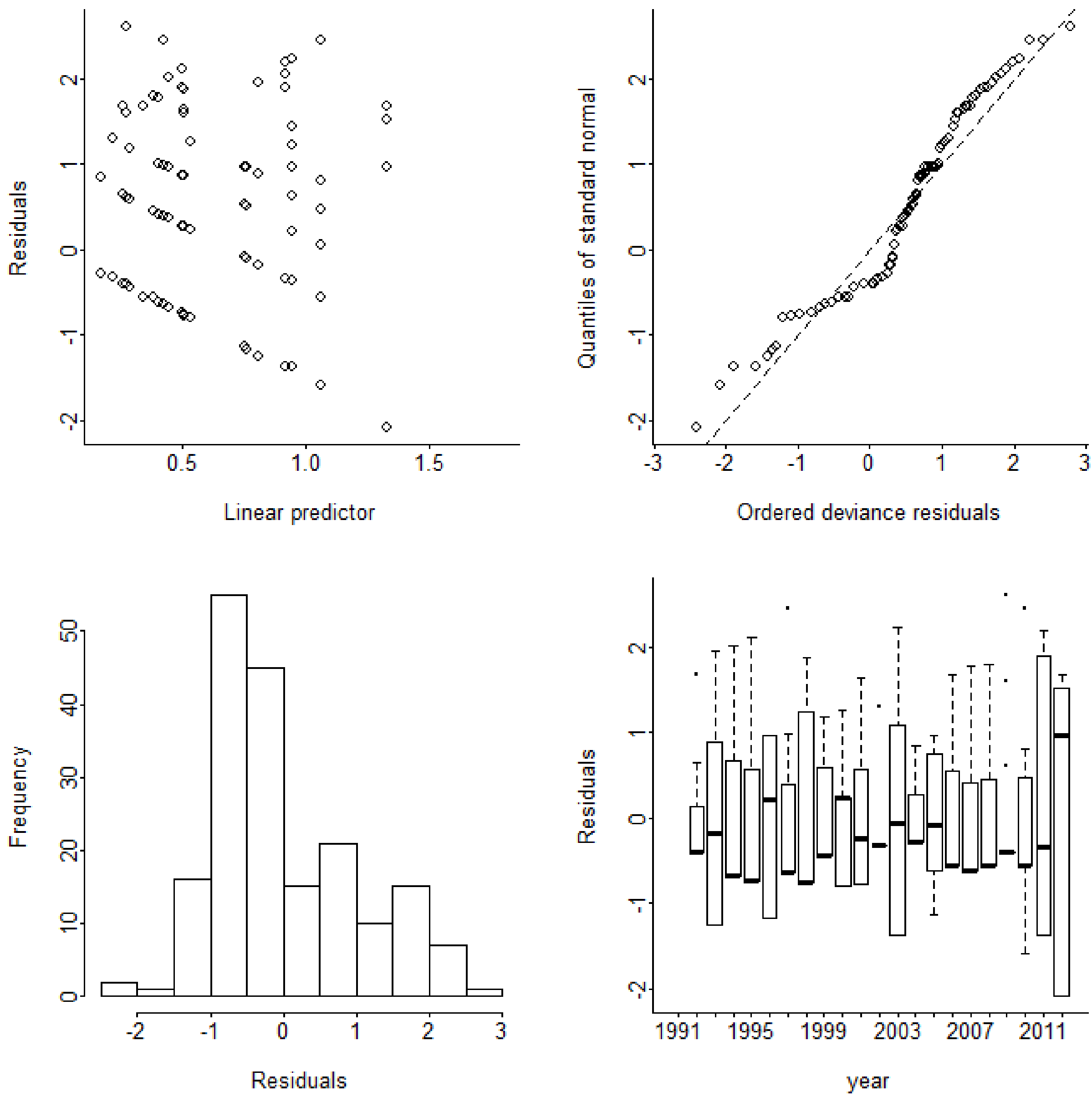

Figure 6. Diagnostic plots from the positives component of the WFL hook-and-line model.

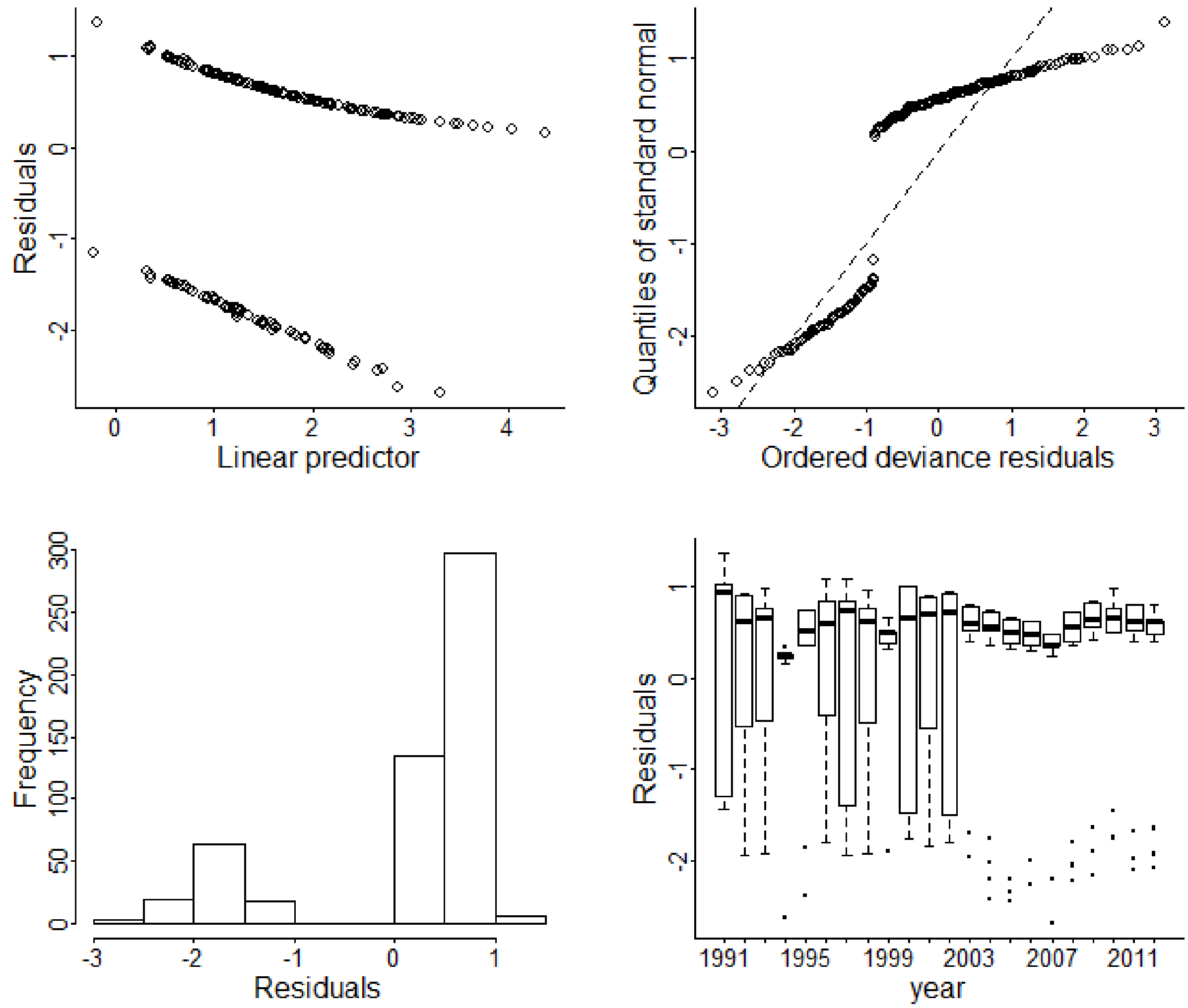

Figure 7. Diagnostic plots from the binomial component of the FLK/SEFL spear model.

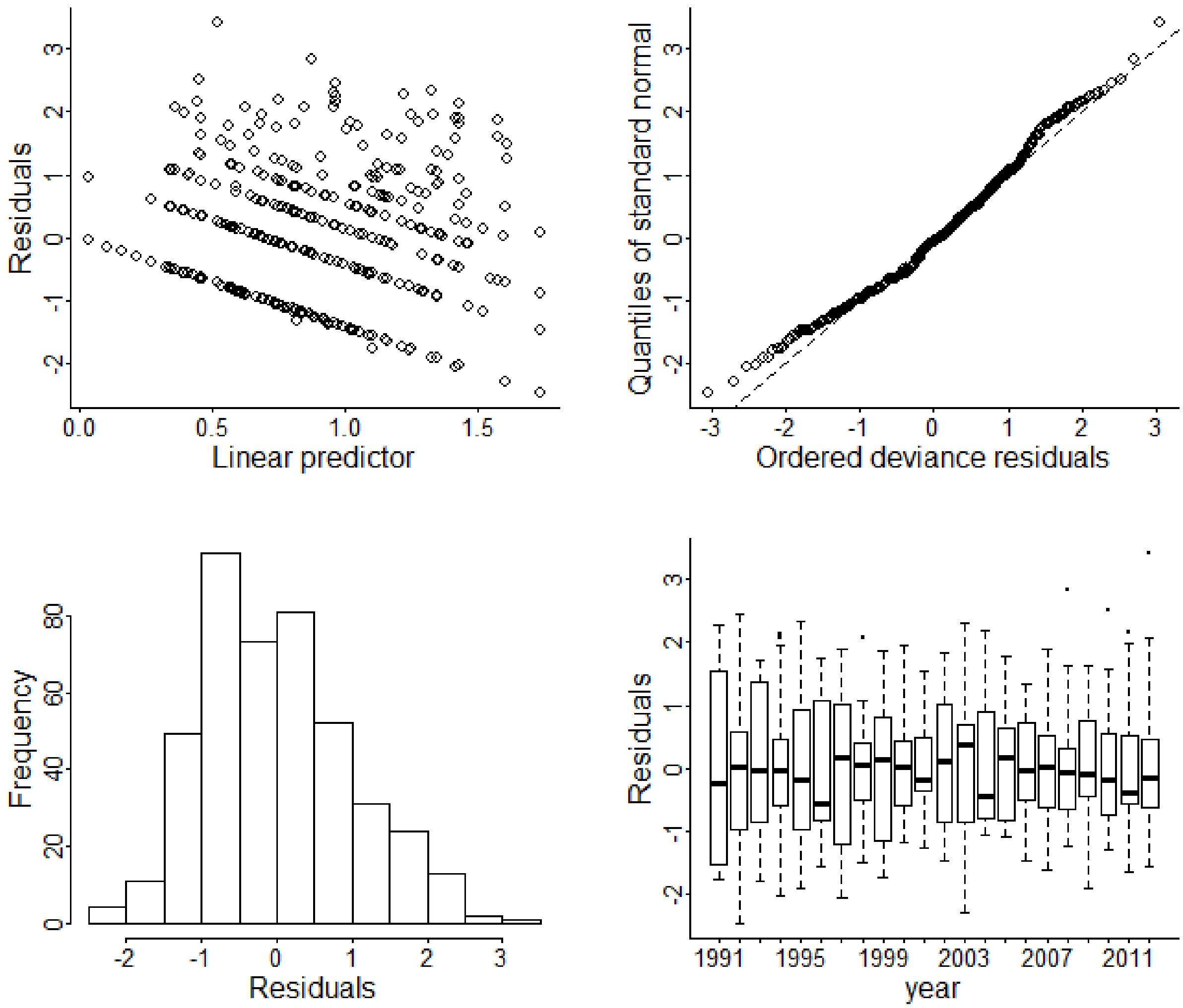

Figure 8. Diagnostic plots from the positives component of the FLK/SEFL spear model.

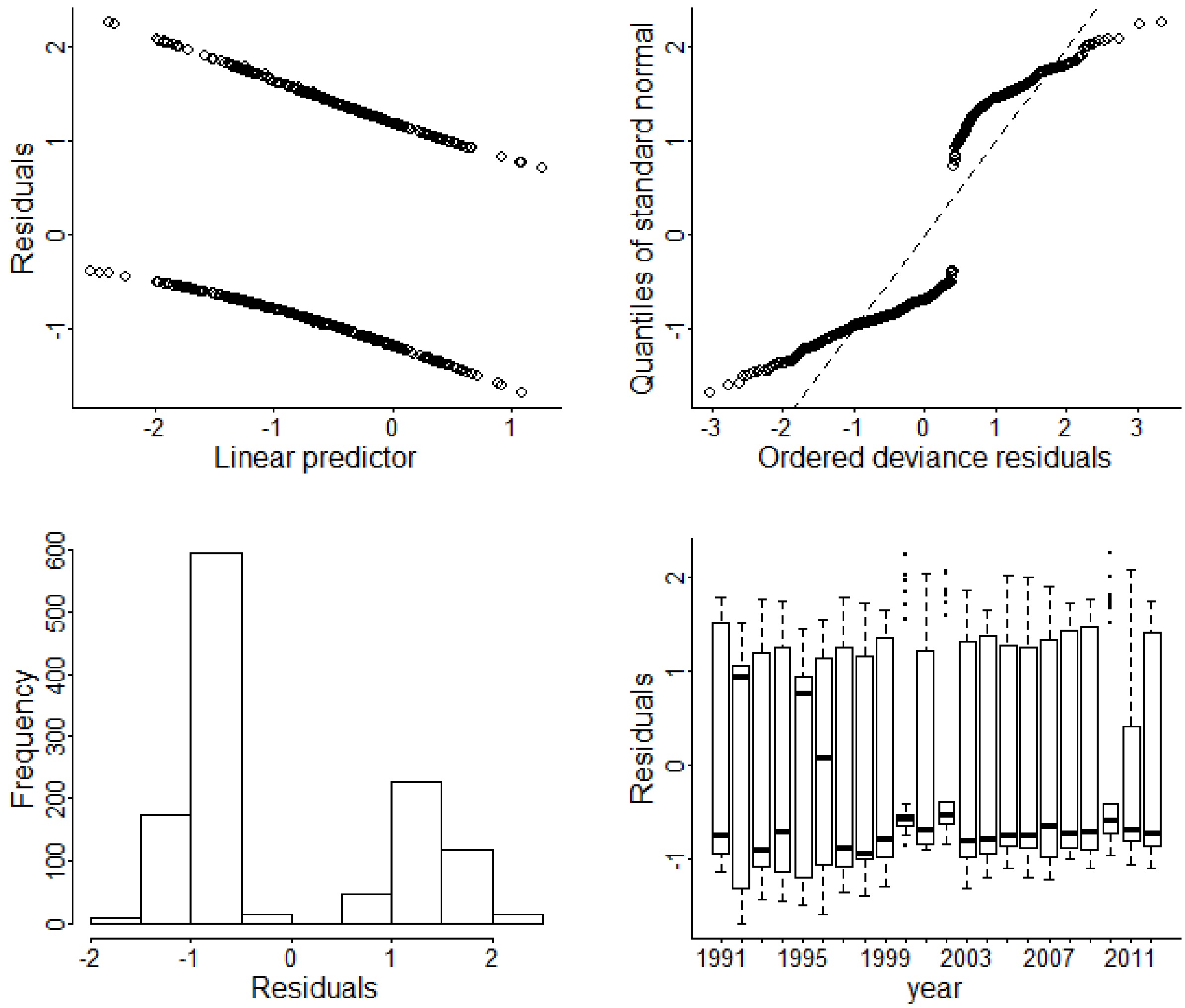

Figure 9. Diagnostic plots from the binomial component of the FLK/SEFL hook-and-line model.

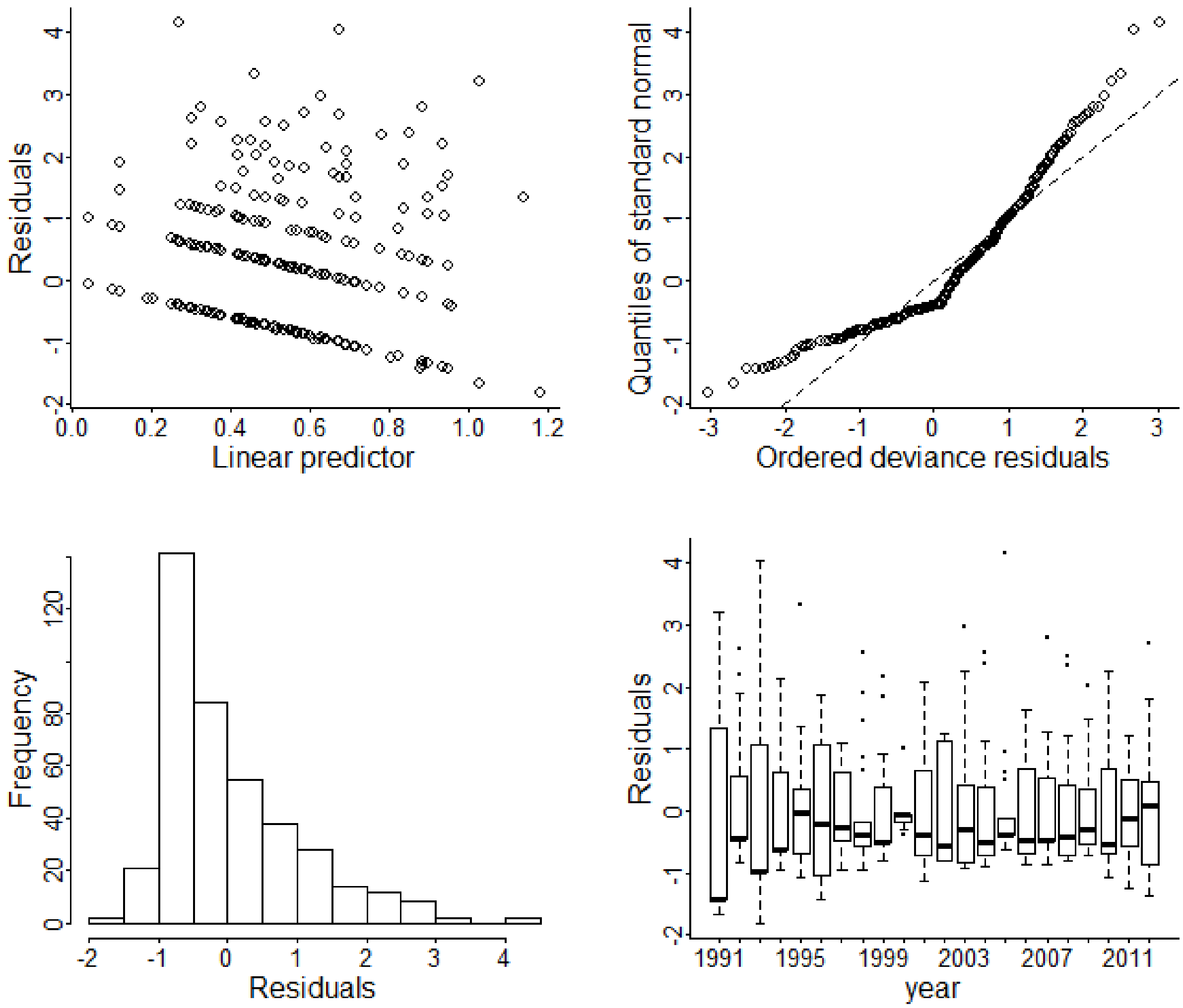

Figure 10. Diagnostic plots from the positives component of the FLK/SEFL hook-and-line model.

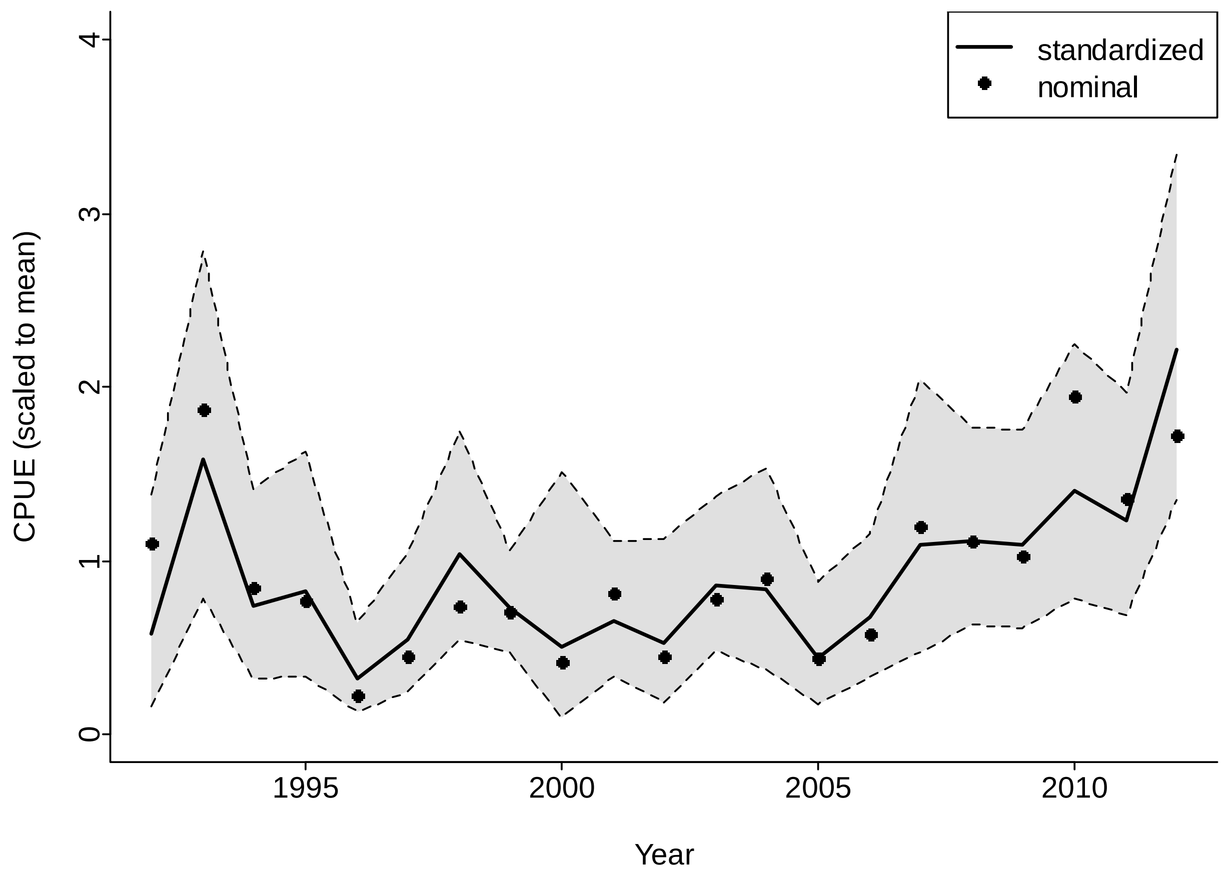

Figure 11. Standardized index of abundance for the WFL spear model.

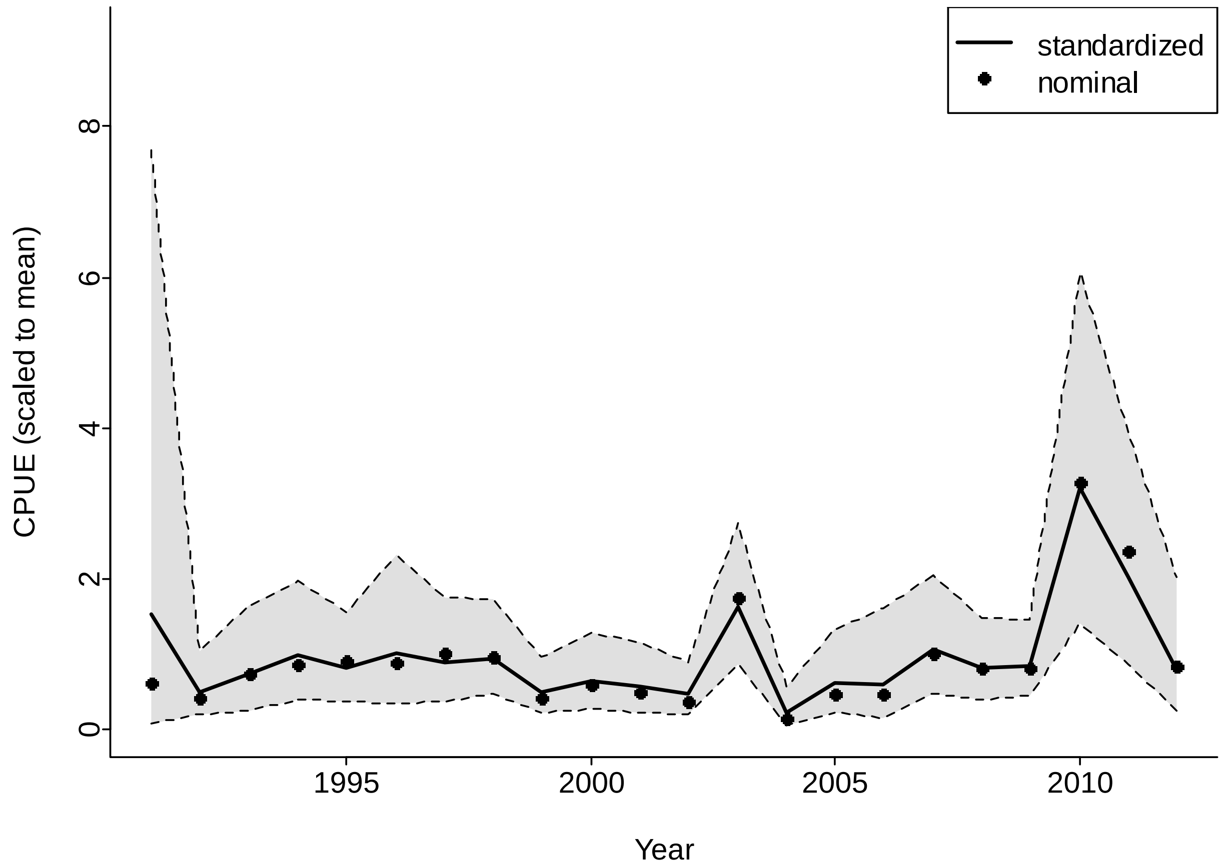

Figure 12. Standardized index of abundance for the WFL hook-and-line model.

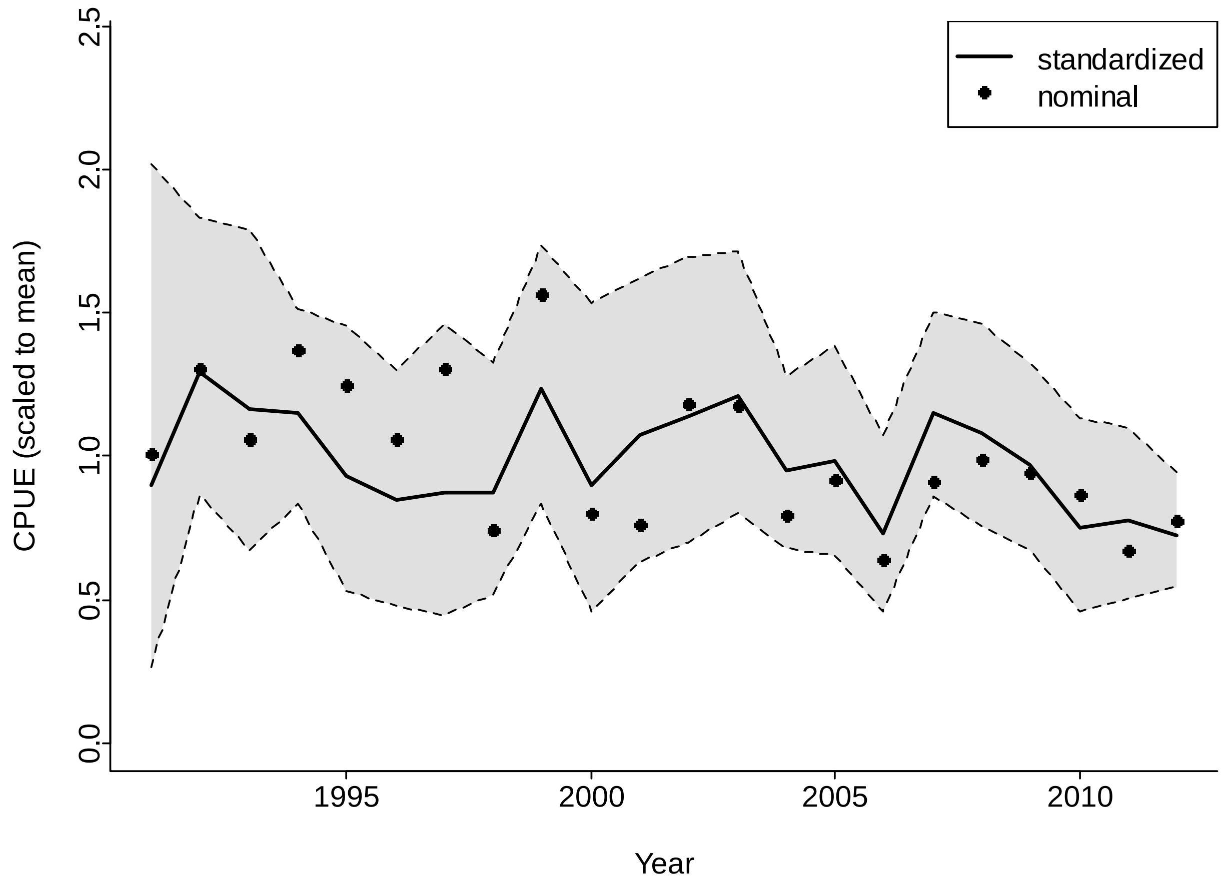

Figure 13. Standardized index of abundance for the FLK/SEFL spear model.

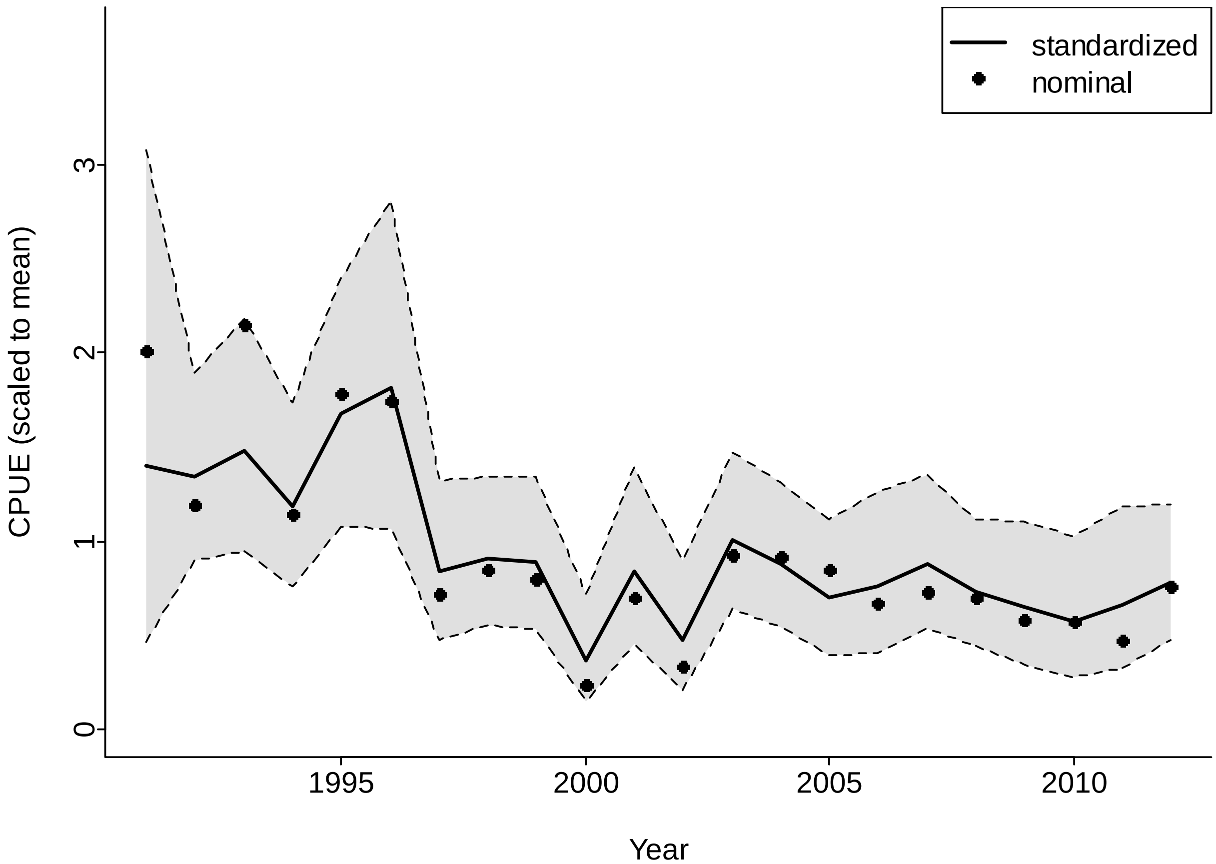

Figure 14. Standardized index of abundance for the FLK/SEFL hook-and-line model.