Model-based size composition of vermilion snapper obtained from three visual surveys

John Walter, Kevin Thompson and Ted Switzer

SEDAR67-WP-16

5 February 2020

Please cite this document as:

Walter, J., Kevin Thompson and Ted Switzer. 2020. Model-based size composition of vermilion snapper obtained from three visual surveys. SEDAR67-WP-16. SEDAR, North Charleston, SC. 15 pp.

Model-based size composition of vermilion snapper obtained from three visual surveys

John Walter, Kevin Thompson and Ted Switzer

Southeast Fisheries Science Center
75 Virginia Beach Drive
Miami FL 33149

Introduction

This paper describes a model-based size composition derived from Remotely operated vehicle camera surveys for vermilion snapper. It follows a similar methodology conducted for red snapper (Walter et al. SEDAR) and uses a multinomial predictive model.

Materials and Methods

Data sources

ROV and stationary camera generated length composition data was available for the years 2005-2016 from four different surveys conducted between three labs, Panama City, FWRI and Mississippi Labs (**Figure 1, Table 1**). Sample data covers much of the West Florida shelf to Alabama (**Figure 2**).

Data exclusions

Habitat type 'poor' had only 4 observations of vermilion snapper and so was removed (**Figure 2**). Years 1995, 1996 and 1997 had very few samples in only 2-3 length bins and were also removed from modeling.

Modeling

Modeling was conducted with a multinomial regression model was run with the R function multinom() from the neural network (nnet) library for R (Venables and Ripley 2012)

Model factors included year as a categorical factor, habitat type (fair and good), note that poor habitat samples were excluded. Depth in meters and station latitude was modeled as a linear functions. Source, as one of the three labs: Panama City, FWRI and Mississippi Labs was modeled as a categorical factor. Lengths were in fork length and use bin spacing from 5 to 55 cm by 5 cm intervals, similar to the assessment model.

Four models were tested starting with stepwise addition of factors with model selection conducted by Akaike's information criterion (AIC). The final model selected incorporated reef type, depth as linear factor, station latitude s FLcat~ YEAR + Reef.Type + DEPTH.M + sta_lat +SOURCE.

model	AIC	deltaAIC
FLcat~ YEAR	19319.92	1605.46
FLcat~ YEAR + Reef.Type	18830.54	1116.08

FLcat~ YEAR + Reef.Type + DEPTH.M	18551.28	836.82
FLcat~ YEAR + Reef.Type + DEPTH.M + sta_lat	17874.08	159.62
FLcat~ YEAR + Reef.Type + DEPTH.M + sta_lat +SOURCE	17714.46	0

Prediction and data weighting

We predicted the model across a balanced grid across all model factors, then averaged the predictions to obtain an annual length composition. This was done by using the R function expand.grid() across all model factors. The predictions were not weighted other than by sample size which achieves a *de facto* weighting. Alternative weightings based on habitat area covered by each survey could be derived but seem unnecessary as the number of observed fish provided an implicit population-level weighting for design based surveys.

Results and discussion

There appears to be an increase in mean size with depth (Figure 5) as well as with latitude. As the Mississippi labs sample further offshore in deeper depths (Figures 1 and 4) the larger mean size appears

The model coefficients were usually well determined, except several size bins in a few years had very high CVs indicative of poor model convergence (**Table 2**). Nonetheless the model had quite good performance in cross-validation (**Figure 6**) indicating robust predictive performance.

Overall the observed and predicted size frequencies (**Table 3**) were not that different though they show some divergence in early years with the model predicting smaller sizes due to the influence of the model factor for depth where the early samples were exclusively from the Mississippi Labs samples which samples at deeper depths (**Figure 7**). Individual length frequencies by lab are shown in **figures 8-10** and indicate that each lab had different mean sizes apparently due to different sampling locations and depths.

Literature cited

Venables, W. N. & Ripley, B. D. 2002. Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0.

Walter, J. D. Devries, M. Drymon, W. Patterson, S. Powers, J. Williams. 2012. A proposed methodology to incorporate ROV length data into red snapper stock assessments SEDAR31-AW08

Table 1. Counts of vermilion snapper by year, habitat, survey.

	Panama City			Mississippi labs			FWRI		
year	Fair	Good	Poor	Fair	Good	Poor	Fair	Good	Poor
1995	0	0	0	9	0	0	0	0	0
1996	0	0	0	14	23	0	0	0	0
1997	0	0	0	60	23	0	0	0	0
2002	0	0	0	208	301	0	0	0	0
2004	0	0	0	280	256	0	0	0	0
2005	0	0	0	520	360	0	0	0	0
2006	0	0	0	214	162	0	0	0	0
2007	0	0	0	153	152	0	0	0	0
2008	0	0	0	37	4	0	0	0	0
2009	0	57	0	43	58	0	0	0	0
2010	0	27	0	0	24	0	13	0	0
2011	2	107	0	213	88	0	73	0	0
2012	3	80	0	31	49	0	70	0	0
2013	0	0	0	29	17	0	157	0	0
2014	0	0	0	33	7	0	427	0	0
2015	0	174	0	81	55	0	528	0	0
2016	0	14	0	23	65	0	707	0	4
2017	9	72	0	90	40	0	0	0	0

Table 2. Model coefficients and coefficients of variation.

Coefficients

	(10,15]	(15,20]	(20,25]	(25,30]	(30,35]	(35,40]	(40,45]	(45,50]	(50,55]
(Intercept)	-141.10	-152.73	-170.85	-172.76	-175.09	-188.15	-184.35	-238.99	-229.62
YEAR2004	-90.83	-4.04	-1.80	-1.33	-1.45	-1.40	-1.04	40.50	33.78
YEAR2005	-14.77	-15.80	-15.12	-14.65	-14.53	-13.42	-11.97	29.76	21.15
YEAR2006	4.77	3.43	3.81	3.59	3.36	2.90	-45.81	-4.17	-6.51
YEAR2007	-8.19	-10.35	-9.18	-8.38	-9.43	-10.23	-6.23	33.98	-13.55
YEAR2008	-42.71	1.61	4.01	4.55	5.78	5.17	-14.45	49.08	-4.67
YEAR2009	-49.86	-4.20	-2.94	-2.61	-2.05	-2.21	-22.63	40.01	35.11
YEAR2010	2.22	-0.57	-0.43	0.36	1.62	3.35	4.13	-0.38	-6.83
YEAR2011	-47.97	-51.58	-50.75	-50.03	-49.43	-48.48	-47.14	-6.29	-12.50
YEAR2012	-50.68	-53.63	-53.40	-53.88	-53.78	-53.32	-51.53	-40.95	-46.22
YEAR2013	-0.08	-3.00	-2.63	-2.78	-2.63	-1.18	-1.12	11.60	-1.89
YEAR2014	-6.86	-9.66	-9.19	-9.26	-8.93	-7.48	-5.68	34.60	27.97
YEAR2015	-49.21	-52.51	-52.50	-52.93	-52.45	-51.19	-49.53	-9.91	-15.97
YEAR2016	-49.42	-53.58	-53.17	-53.24	-52.65	-51.95	-51.20	-10.38	-16.83
YEAR2017	2.10	-0.30	0.45	0.64	1.62	2.87	3.33	45.09	9.46
Reef.TypeG	-7.42	-6.97	-6.90	-6.37	-6.27	-6.63	-5.30	-6.25	-7.98
DEPTH.M	0.07	0.03	0.02	0.02	0.04	0.04	0.03	0.02	0.10
sta_lat	7.24	7.92	8.60	8.63	8.62	8.99	8.80	9.27	9.02
SOURCEPasc	4.76	4.68	4.57	5.16	5.56	6.21	5.80	5.90	5.93
SOURCEPC	-12.30	-14.84	-16.22	-15.96	-15.82	-15.53	-18.94	-17.58	-13.05

Coefficient of variation

	(10,15]	(15,20]	(20,25]	(25,30]	(30,35]	(35,40]	(40,45]	(45,50]	(50,55]
(Intercept)	-1.6%	-0.8%	-0.6%	-0.6%	-0.8%	-1.3%	-0.2%	-0.1%	-0.1%
YEAR2004	NA	-9.6%	-11.3%	-15.1%	-16.3%	-26.6%	-60.6%	1.3%	1.9%
YEAR2005	-4.0%	-1.3%	-1.2%	-1.2%	-1.5%	-2.4%	-4.3%	1.1%	2.0%
YEAR2006	11.9%	6.1%	5.0%	5.5%	7.8%	19.4%	NA	NA	NA
YEAR2007	-12.7%	-3.3%	-2.8%	-2.9%	-3.4%	-6.9%	-9.1%	1.6%	0.0%
YEAR2008	NA	55.0%	9.9%	8.5%	7.1%	17.6%	0.0%	1.7%	0.0%
YEAR2009	NA	-9.6%	-8.7%	-9.6%	-14.5%	-24.0%	0.0%	2.1%	1.8%
YEAR2010	38.2%	-68.6%	-83.2%	88.0%	21.8%	13.3%	16.0%	0.0%	NA
YEAR2011	-1.8%	-1.4%	-1.4%	-1.4%	-1.5%	-1.6%	-1.9%	-12.7%	-6.2%
YEAR2012	-1.6%	-1.1%	-1.1%	-1.1%	-1.2%	-1.6%	-1.8%	0.0%	0.0%
YEAR2013	-779.0%	-9.7%	-10.2%	-10.5%	-15.8%	-47.4%	-89.7%	NA	NA
YEAR2014	-8.9%	-2.9%	-2.8%	-3.0%	-4.0%	-7.0%	-11.6%	2.7%	3.1%
YEAR2015	-1.6%	-1.1%	-1.1%	-1.1%	-1.2%	-1.3%	-1.6%	-10.8%	-5.4%
YEAR2016	-1.9%	-1.5%	-1.5%	-1.5%	-1.6%	-1.7%	-2.0%	-11.4%	-7.0%
YEAR2017	28.7%	-90.0%	51.0%	35.5%	16.1%	13.3%	19.1%	1.2%	0.0%
Reef.TypeG	-9.9%	-7.0%	-6.8%	-7.4%	-7.6%	-7.4%	-9.5%	-8.7%	-8.3%
DEPTH.M	54.0%	129.9%	217.6%	203.6%	106.2%	100.2%	110.7%	171.1%	42.0%
sta_lat	1.4%	1.0%	0.9%	0.9%	1.0%	1.2%	0.9%	0.9%	0.9%
SOURCEPasc	12.0%	11.6%	11.8%	10.6%	10.2%	10.7%	11.7%	21.5%	20.8%
SOURCEPC	-6.0%	-3.3%	-3.0%	-3.1%	-3.4%	-4.3%	-6.0%	-9.3%	-11.6%

Table 3. Predicted size frequency input to Stock Synthesis

YEAR	month	fleet	sex	part	Nsamp	X.5.10.	X.10.15.	X.15.20.	X.20.25.	X.25.30.	X.30.35.	X.35.40.	X.40.45.	X.45.50.	X.50.55.
2002	8	10	0	0	509	0.000	0.012	0.355	0.346	0.225	0.056	0.006	0.002	0.000	0.000
2004	8	10	0	0	536	0.000	0.000	0.053	0.419	0.421	0.090	0.011	0.003	0.001	0.003
2005	8	10	0	0	880	0.000	0.020	0.180	0.319	0.319	0.087	0.030	0.030	0.012	0.004
2006	8	10	0	0	376	0.000	0.038	0.289	0.412	0.216	0.043	0.003	0.000	0.000	0.000
2007	8	10	0	0	305	0.000	0.039	0.113	0.327	0.454	0.038	0.002	0.025	0.002	0.000
2008	8	10	0	0	41	0.000	0.000	0.036	0.321	0.339	0.274	0.017	0.000	0.014	0.000
2009	8	10	0	0	158	0.000	0.000	0.120	0.373	0.327	0.139	0.013	0.000	0.002	0.026
2010	8	10	0	0	64	0.000	0.090	0.154	0.170	0.224	0.187	0.116	0.060	0.000	0.000
2011	8	10	0	0	483	0.047	0.115	0.097	0.229	0.291	0.127	0.037	0.036	0.006	0.015

2012	8	10	0	0	233	0.147	0.098	0.218	0.342	0.138	0.038	0.007	0.013	0.000	0.000
2013	8	10	0	0	203	0.000	0.129	0.231	0.351	0.197	0.056	0.028	0.008	0.000	0.000
2014	8	10	0	0	467	0.000	0.109	0.205	0.330	0.196	0.066	0.032	0.049	0.005	0.008
2015	8	10	0	0	838	0.113	0.138	0.231	0.292	0.122	0.049	0.020	0.031	0.002	0.002
2016	8	10	0	0	809	0.130	0.173	0.142	0.282	0.167	0.075	0.018	0.011	0.002	0.002
2017	8	10	0	0	211	0.000	0.064	0.157	0.307	0.233	0.150	0.058	0.022	0.010	0.000

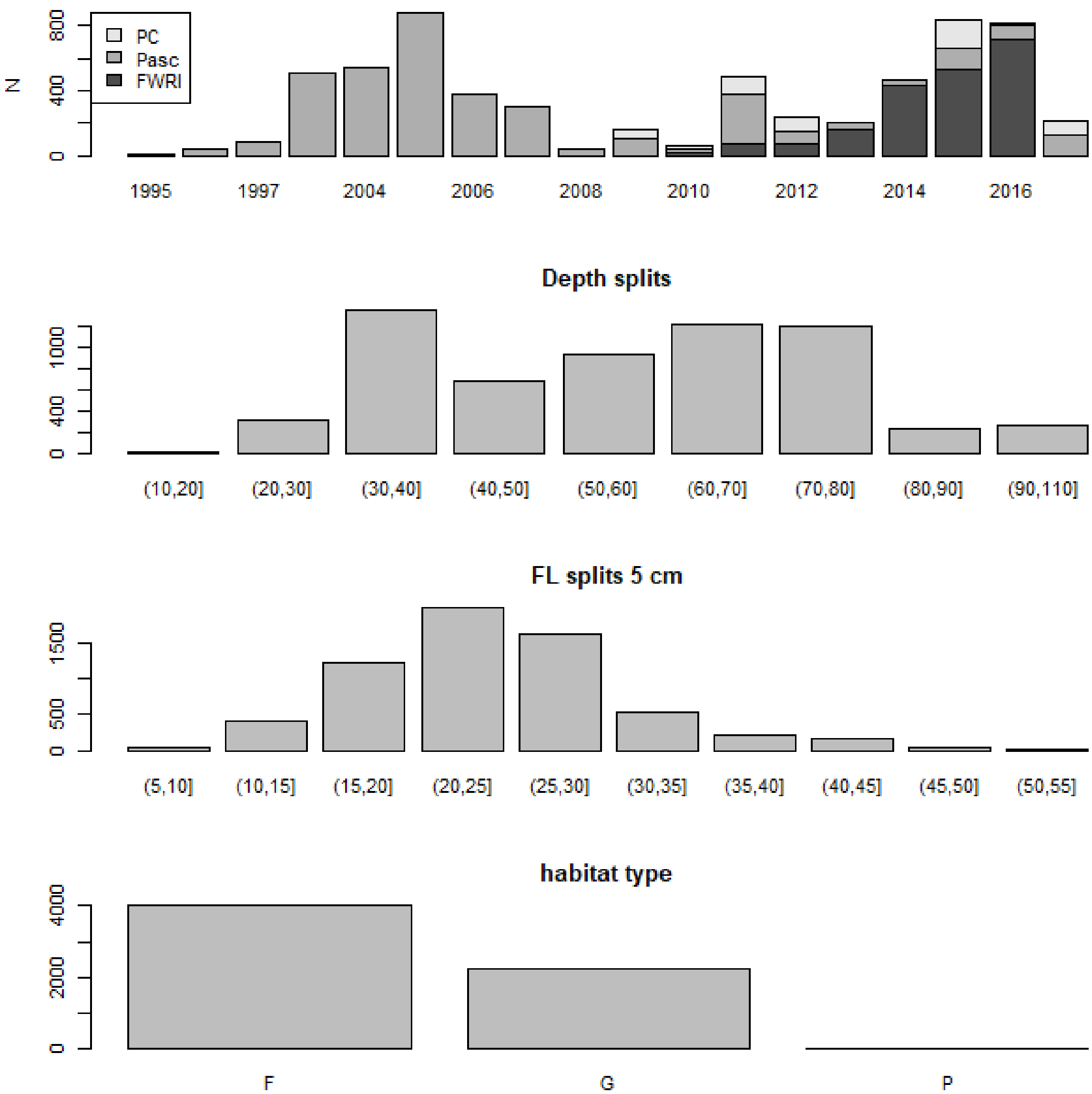

Figure 1. Data by source and year, distribution of samples by depth, distribution by length class and habitat type.

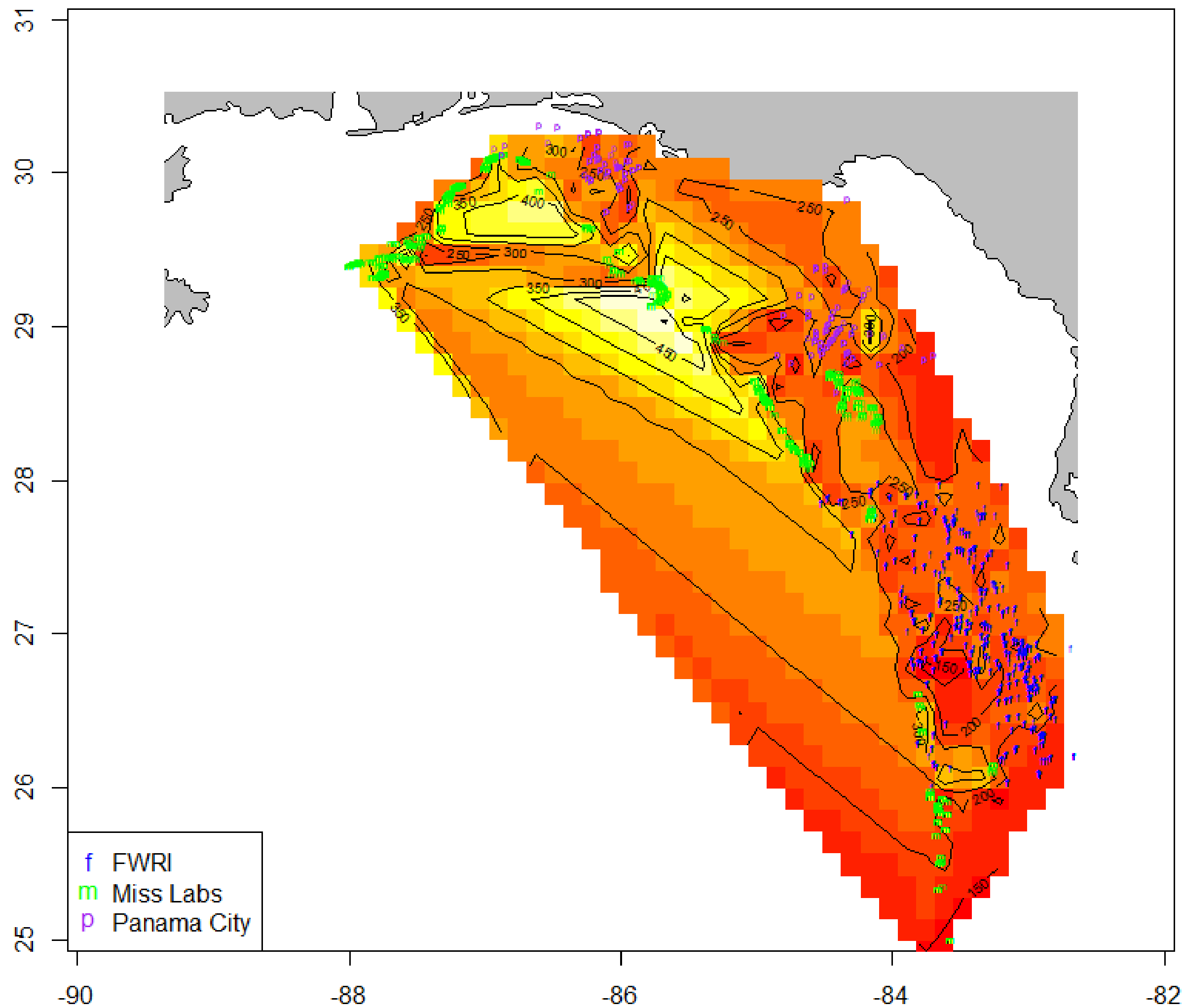

Figure 2. Spatial map and interpolated fork length. Each data source is denoted by a letter (M- Pascagoula, Mississippi lab; P- Panama City Lab; F- FWRI data.

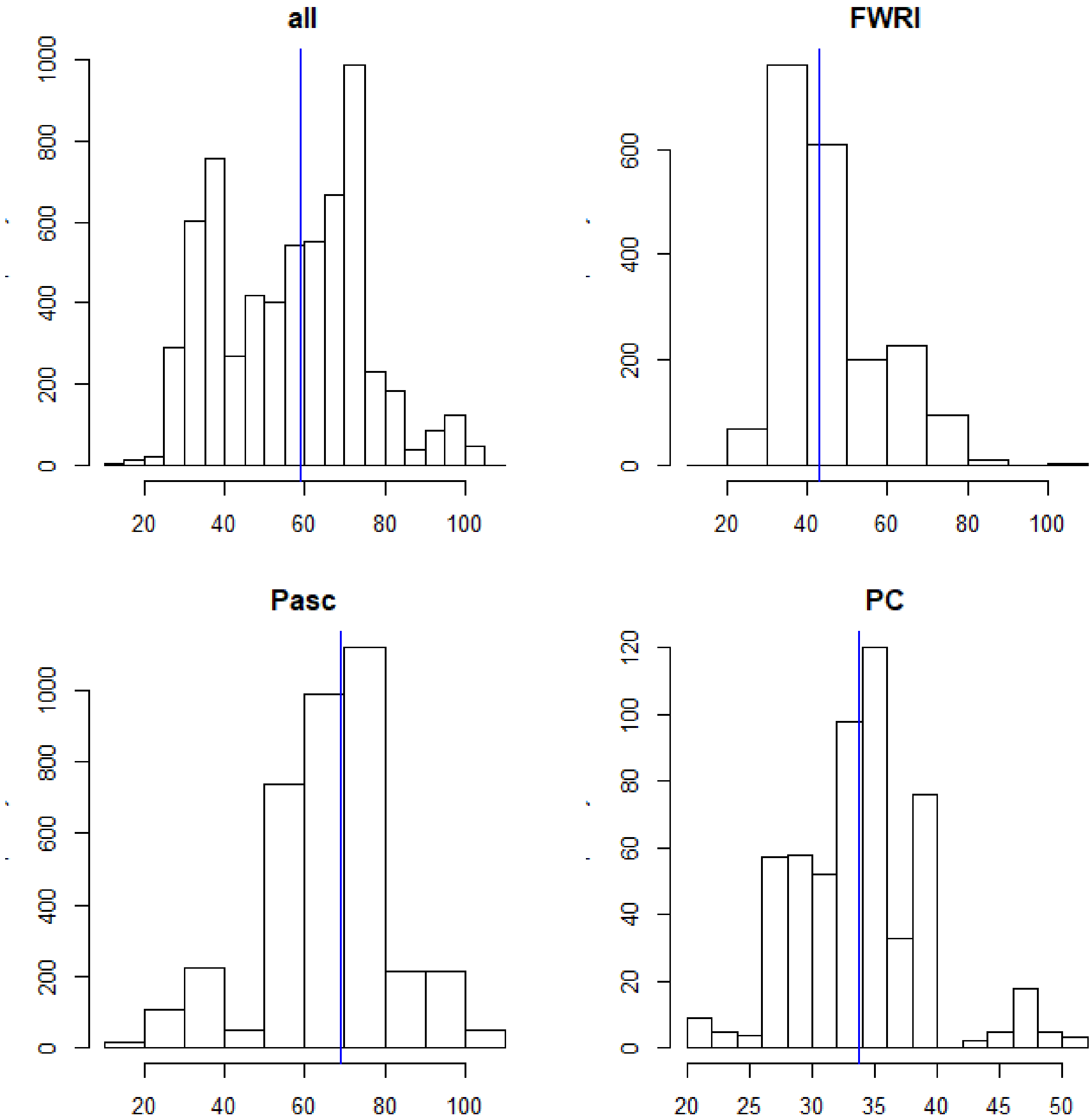

Figure 3. Depth (m) and median (blue line) overall and for each study.

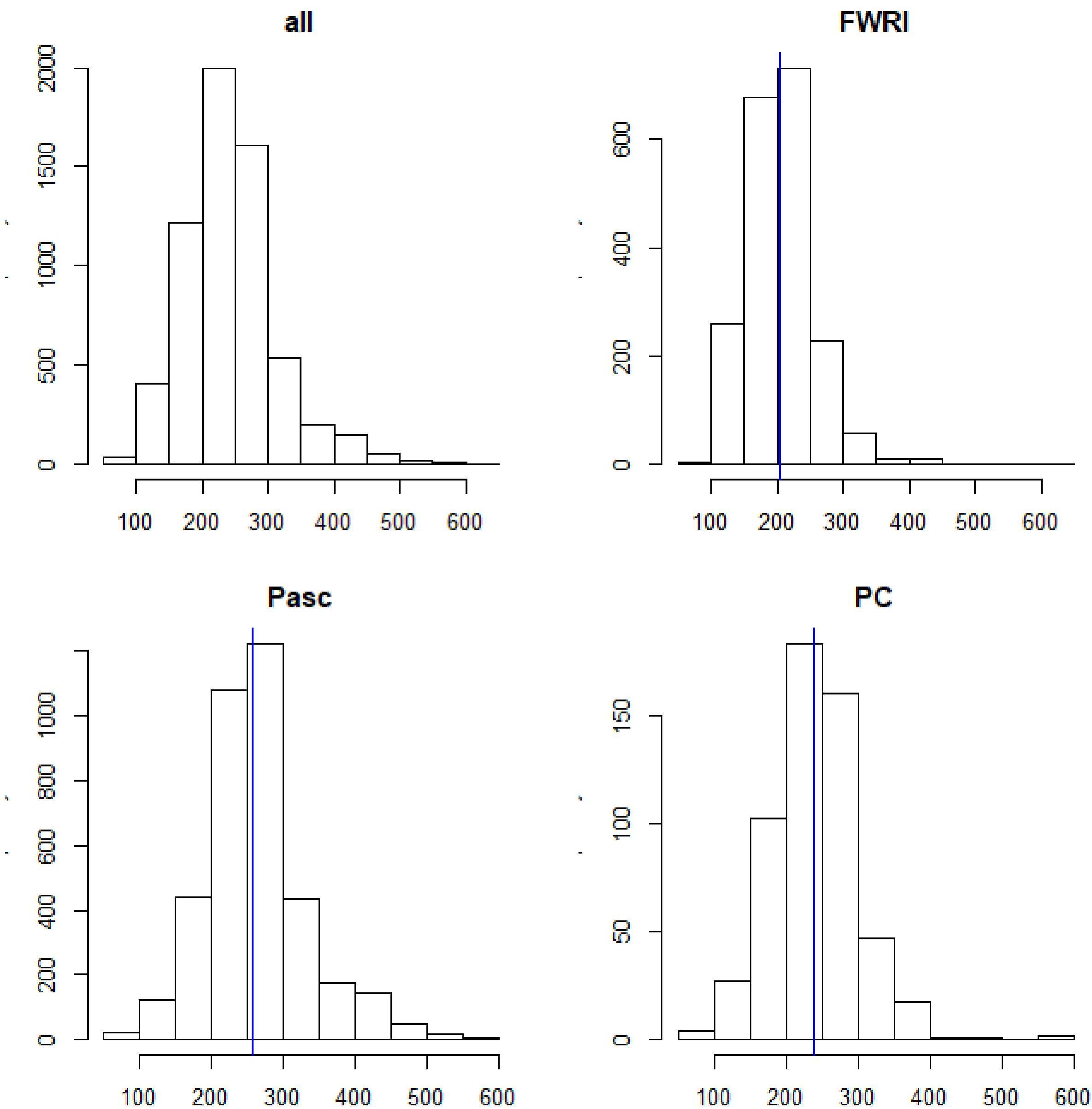

Figure 4. FL (mm) overall and for each dataset.

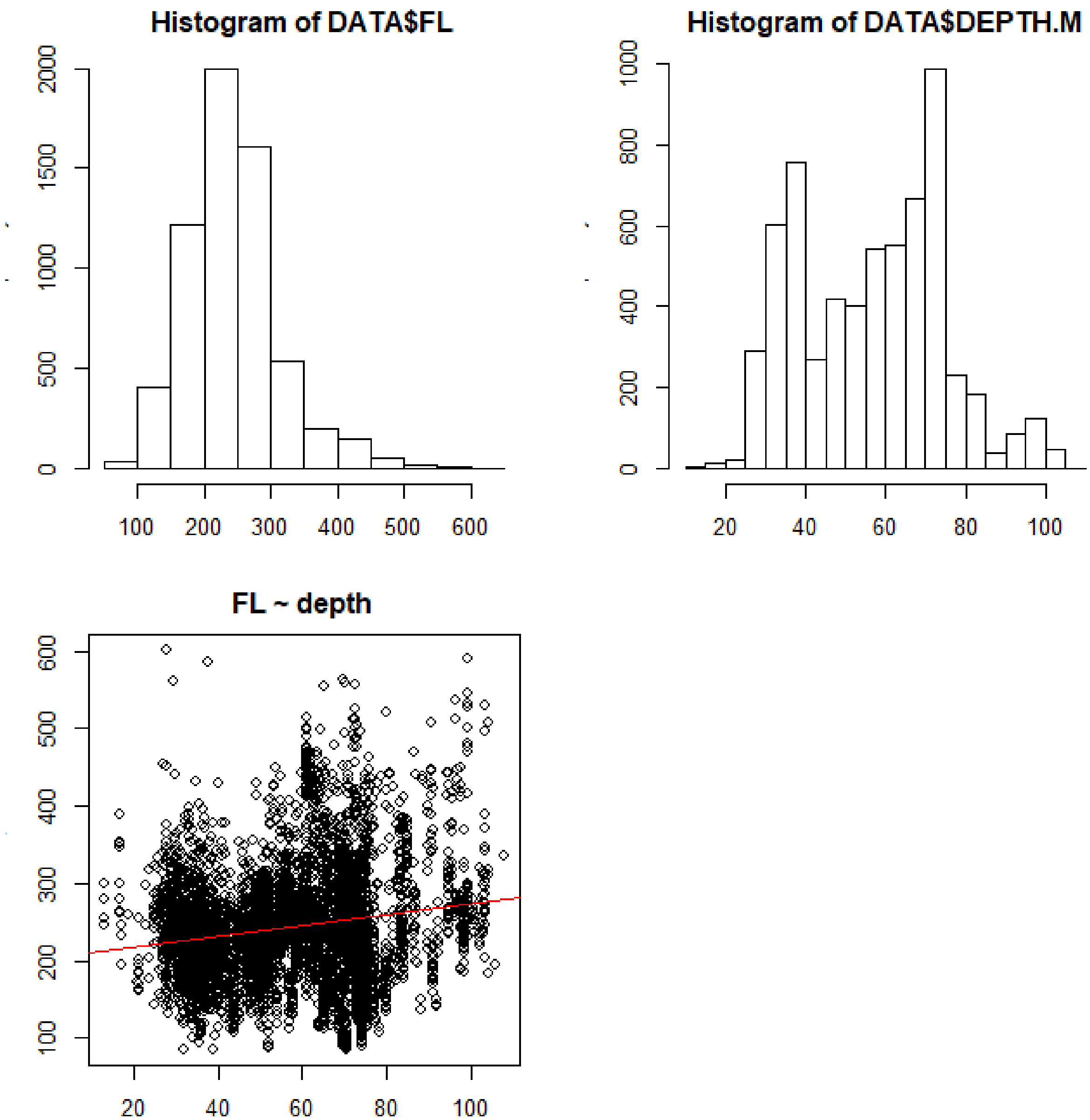

Figure 5. FL (mm) vs depth (m)

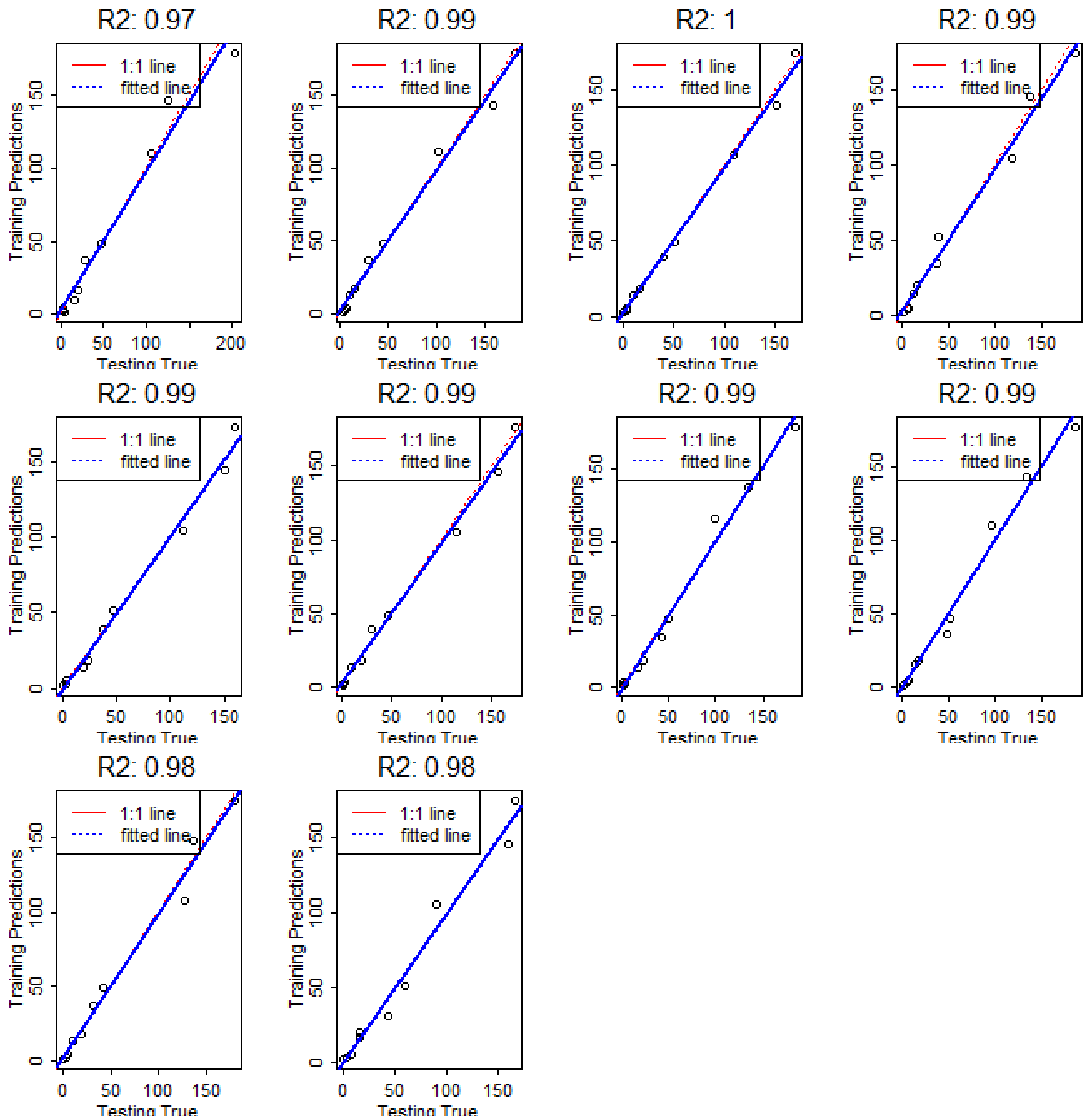

Figure 6. K-fold cross-validation of model performance of testing true data values versus training predictions.

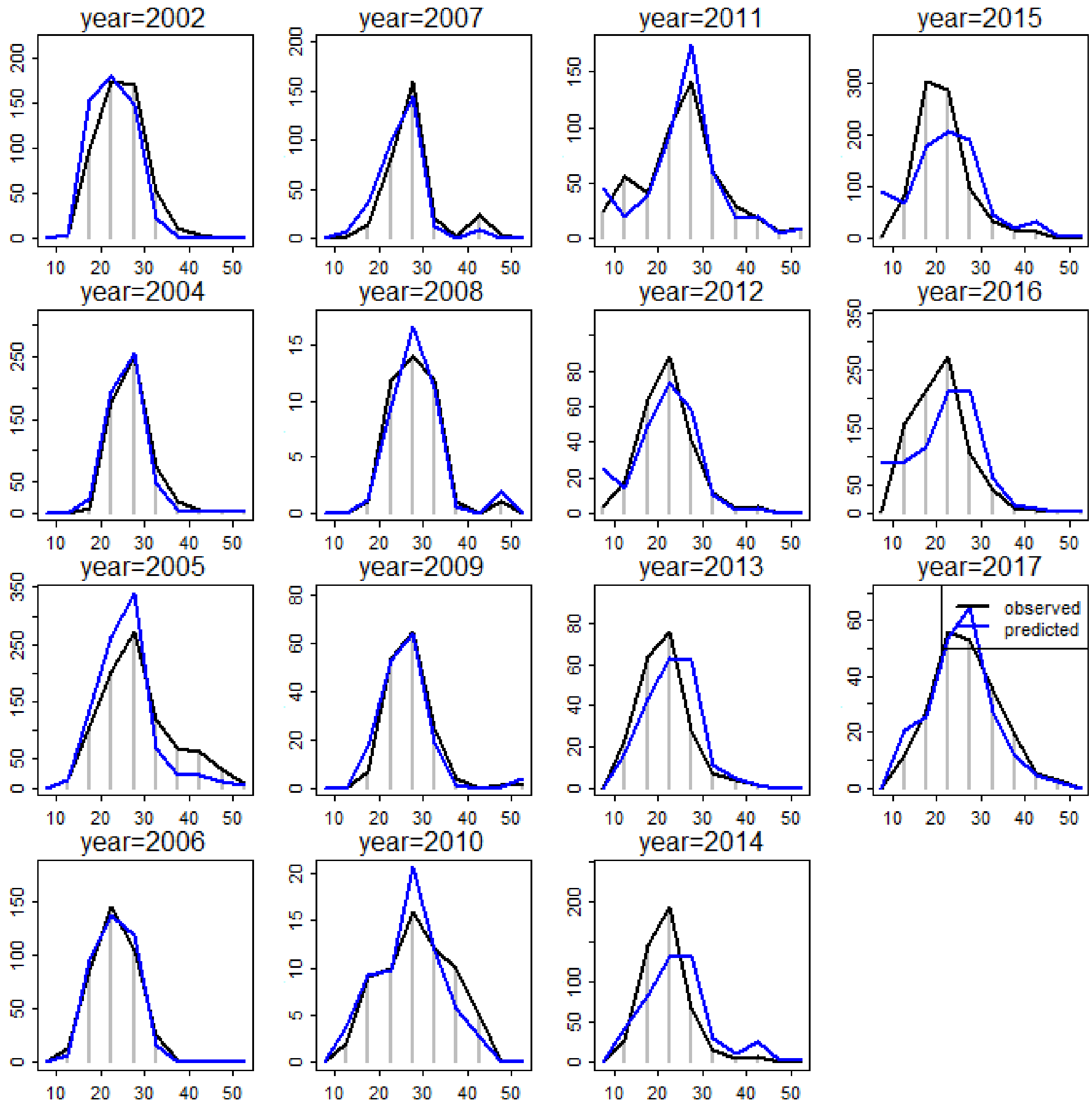

Figure 7. Observed and model-predicted length frequencies.

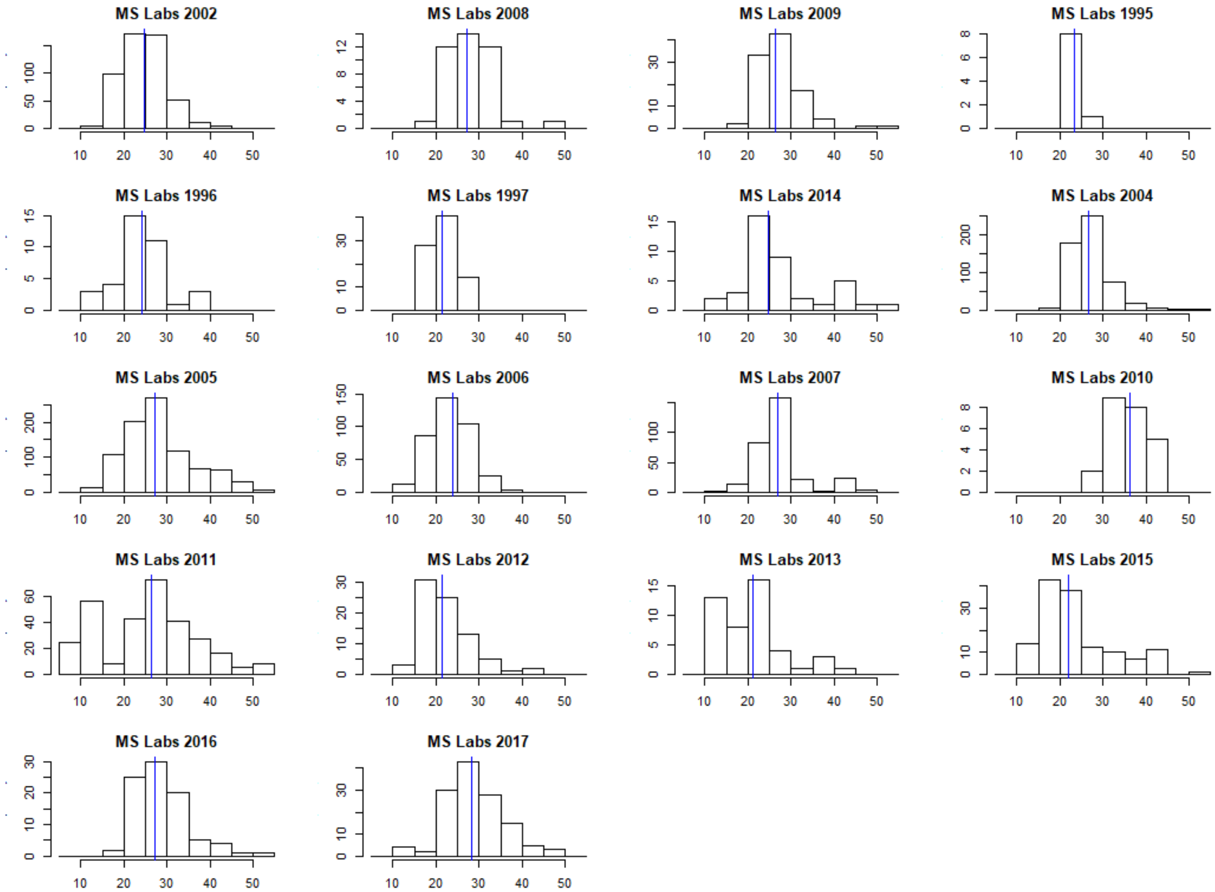

Figure 8. Mississippi Labs length frequency by year.

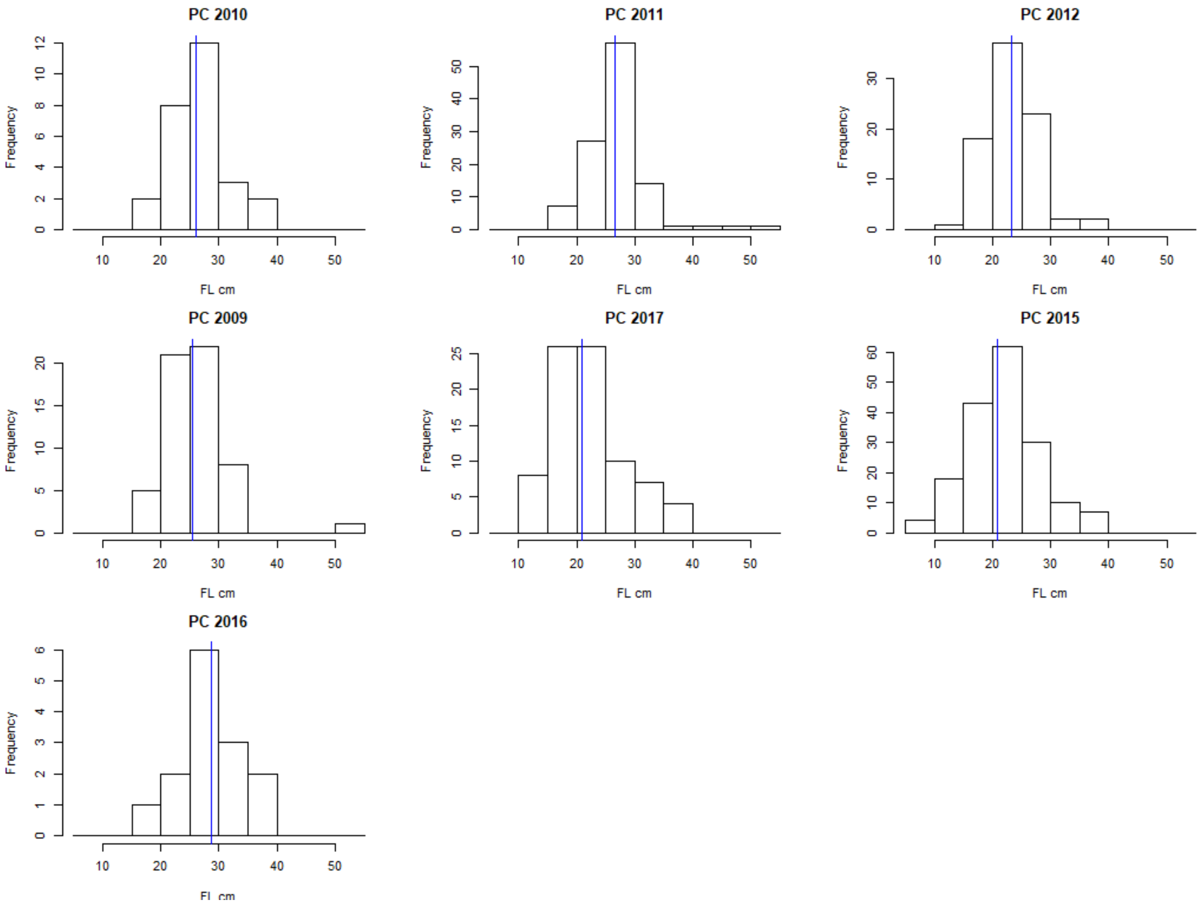

Figure 9. Panama City length frequency by year.

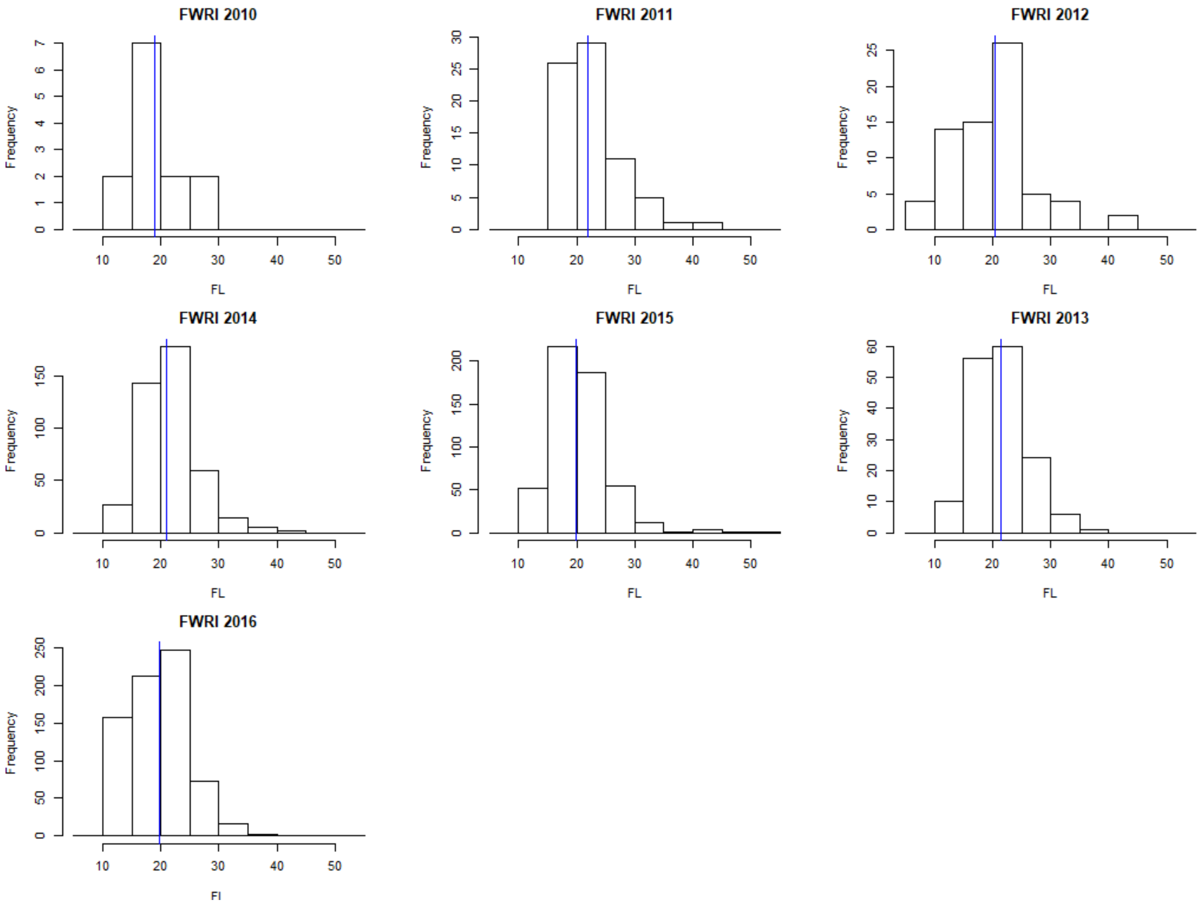

Figure 10. FWRI length frequency by year.