Standardized Catch Rates

Humber A. Andrade

Applied Statistical Modeling
Fisheries and Aquaculture Department – DEPAq
Federal Rural University of Pernambuco – UFRPE
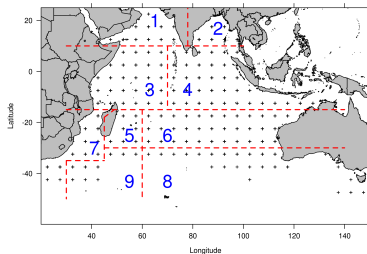
Summary

Variables

- Response variable $(Y) \Rightarrow$ Number of fish
- Explanatory variables $(\boldsymbol{x})$
  - Year $\Rightarrow$ factor
  - Month $\Rightarrow$ factor
  - Area $\Rightarrow$ factor

Model Especification

Expectation of response variable

$$E(Y_i) = \mu_i = n_i \theta_i$$

Mean rate

$$\theta_i = e^{\boldsymbol{x_i}^T \boldsymbol{\beta}}$$

Development

$$E(Y_i) = \mu_i = n_i e^{\boldsymbol{x_i}^T \boldsymbol{\beta}}$$

$$\log \mu_i = \log n_i + \boldsymbol{x_i}^T \boldsymbol{\beta}$$

Distribution of response variable

$$Y_i \sim Poisson(\mu_i)$$

Components

$n \Rightarrow$ number of hooks

link function $\Rightarrow \log(\cdot)$

offset $\Rightarrow \log(n)$

Order of the Explanatory Variables

Steps:

1. Fit simple models with each of explanatory variables once at a time;

2. Select the "most important" variable based on some numerical criterion (*e.g.* Deviance, AIC, ...);

3. Fit models with variable selected in the above step plus each of explanatory variables once at a time;

4. Select the "next most important" variable based on some numerical criterion;

5. Fit models with variables selected in the above steps plus each of explanatory variables once at a time;

6. Select the "next most important" variable based on some numerical criterion;

7. Repeat the above steps until all variables are in the model.

Selection of Variables and Models

Interactions:

1. Fit a full model with all variables, main effects and first order interactions;

2. Calculate some numerical criterion (*e.g.* AIC) for the base case (full model);

3. Fit models more simple than the base case by discarding each interaction once at a time;

4. Calculate the numerical criterion for all the models and select the new base case relying in that criterion;

5. Repeat the procedure until the base case does not change anymore.

The procedure used to select (or discard) <u>main effects</u> is similar to those described above. However a main effect was dropped off only if it is not included in interactions that were not discard.

Summary Statistics and Diagnostics

1. Residuals Diagnostics
2. Table of the Analysis of Deviances
3. Chi-squared goodness of fit related to standardized residuals

$$
\begin{aligned}
X^2 &= \sum r_i^2 = \sum \left(\frac{o_i - e_i}{\sqrt{e_i}} \right)^2 \\
&= \sum \frac{(o_i - e_i)^2}{e_i} \sim \chi^2_{(N-p)}
\end{aligned}
$$

4.

$$
\text{Pseudo } R^2 = \frac{l(\hat{\boldsymbol{\beta}}_{min}) - l(\hat{\boldsymbol{\beta}})}{l(\hat{\boldsymbol{\beta}}_{min})}
$$

Japan Dataset – Year x Month
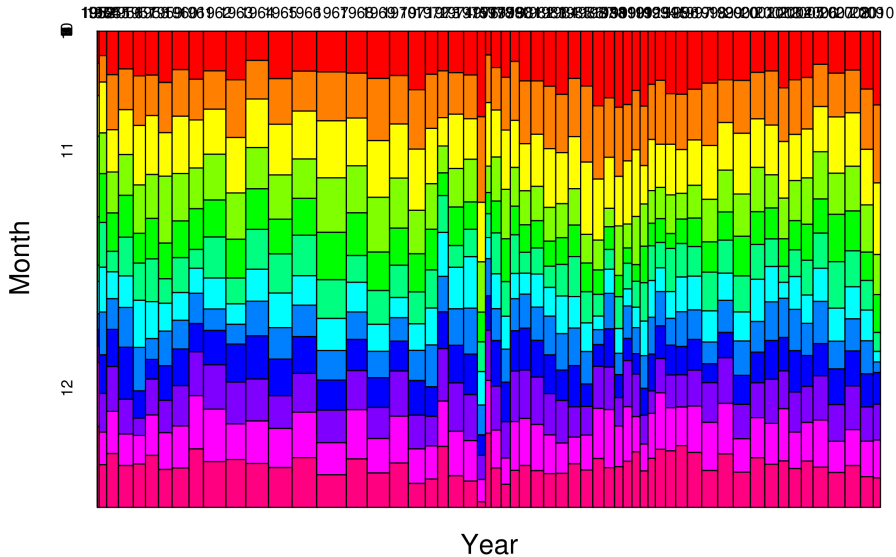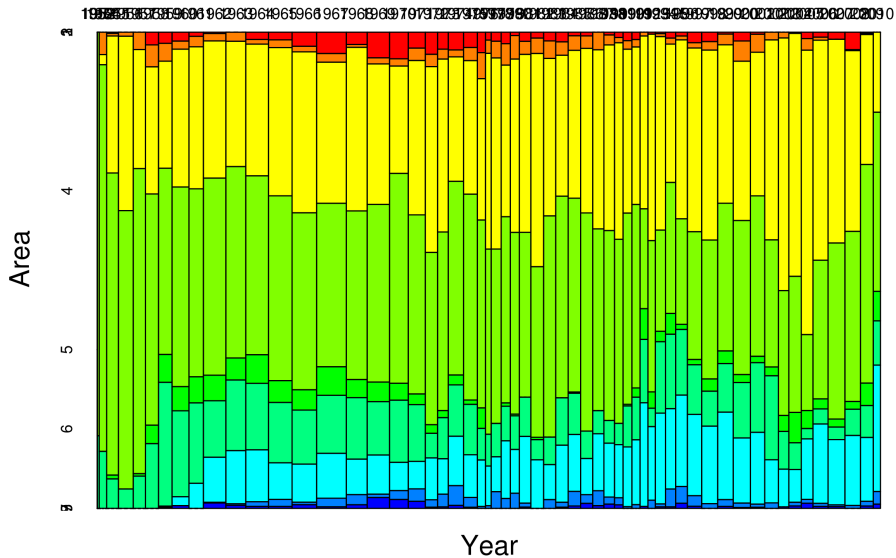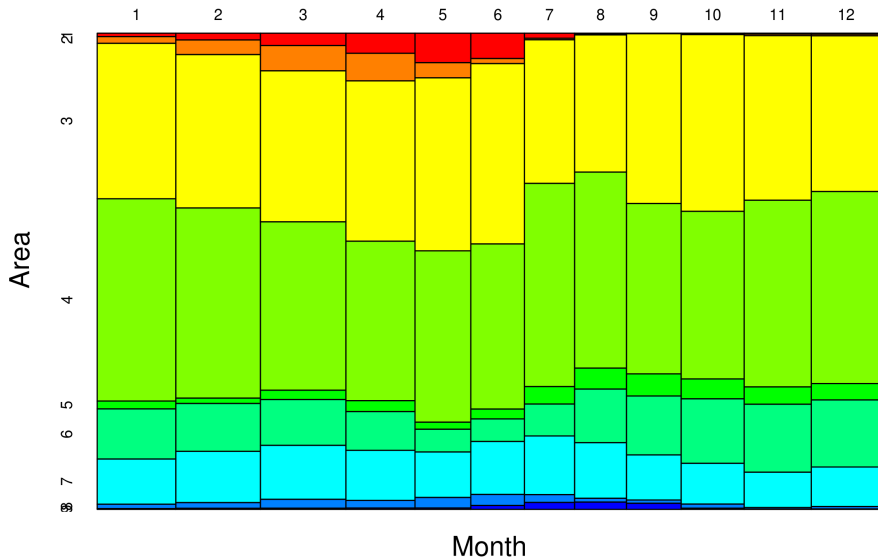
Balance Year x Area

Balance Month x Area

Taiwan Dataset – Year x Month
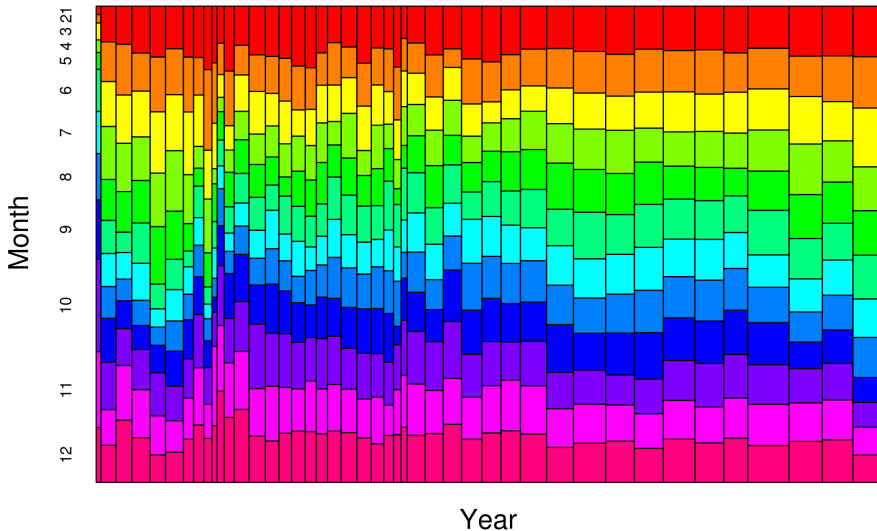
Balance Year x Area
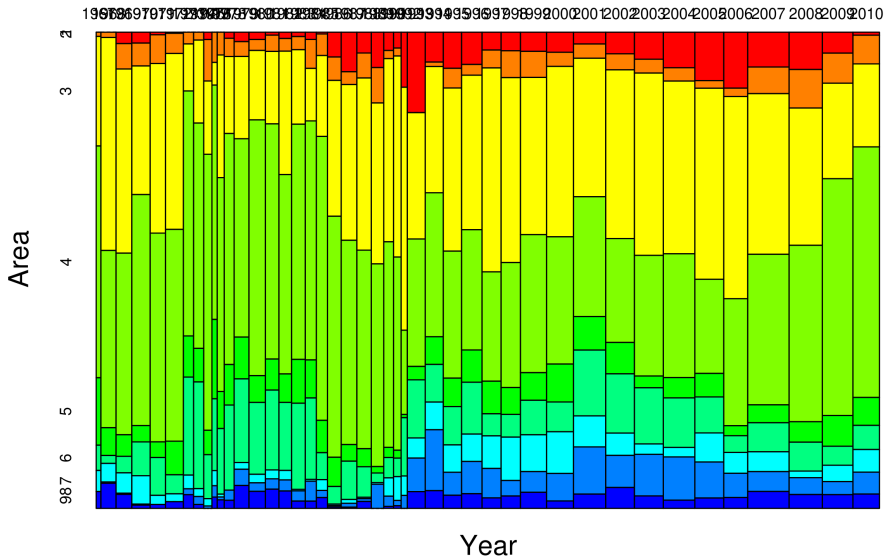
Balance Month x Area
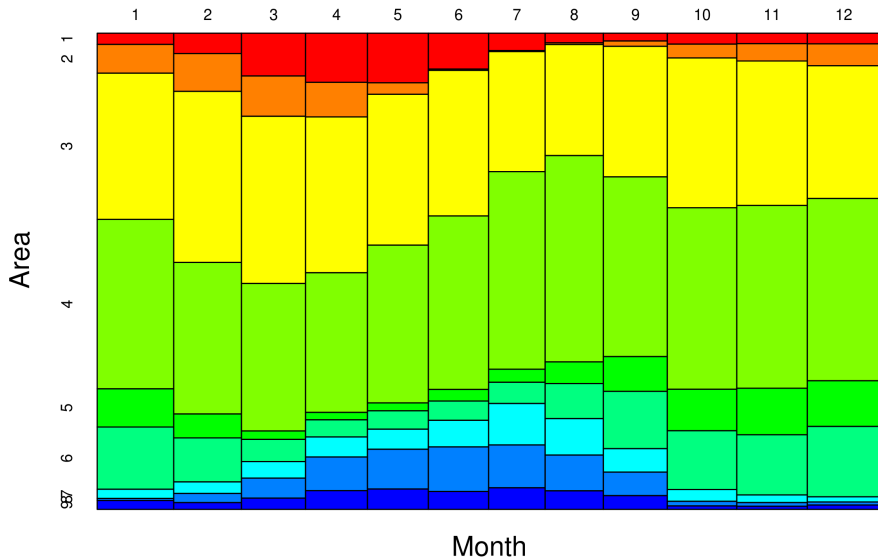
Blue Marlin – Japan Dataset

Selected Models

- Complex (calculated in a computer with more memory)

$$Y \sim year + area + month + year : area + year : month + area : month$$
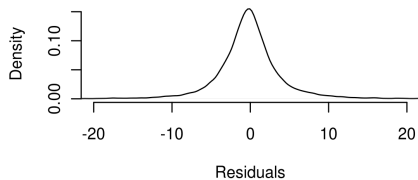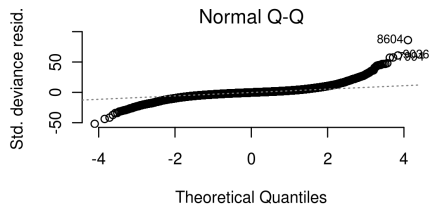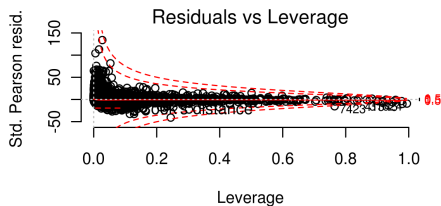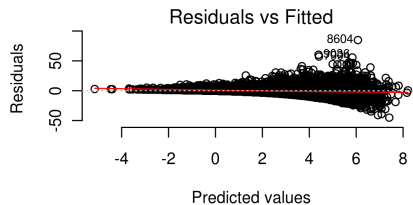
- Simple (calculated in this computer)

$$Y \sim year + area + month + year : area + area : month$$

Residuals Diagnostics – Complex Model
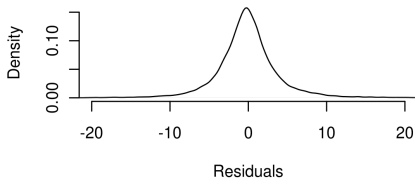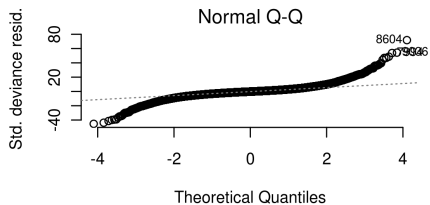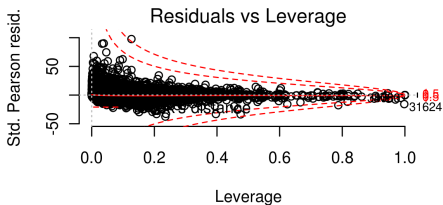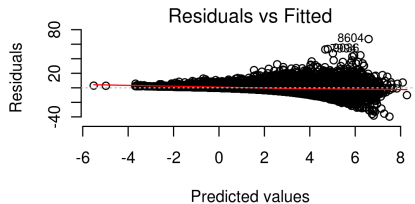
Residuals Diagnostics – Simple Model

Summary Statistics – Complex Model

Analysis of Deviance

	Df	Deviance	Resid. Df	Resid. Dev	P(>|Chi|)	Dev.Exp2
NULL			24684	2222120.41		
Year	58	1174053.32	24626	1048067.09	0.00	52.83
area	8	271009.12	24618	777057.97	0.00	12.20
MonthStart	11	74036.84	24607	703021.12	0.00	3.33
Year:area	391	96150.61	24216	606870.52	0.00	4.33
Year:MonthStart	628	76571.76	23588	530298.76	0.00	3.45
area:MonthStart	88	45186.79	23500	485111.96	0.00	2.03

$X^2 \sim \chi^2_{(N-p)}$

statistic	p.value
632936.24	0.00

Pseudo R^2

0.75

Summary Statistics – Simple Model

Analysis of Deviance

	Df	Deviance	Resid. Df	Resid. Dev	P(>|Chi|)	Dev.Exp2
NULL			24684	2222120.41		
Year	58	1174053.32	24626	1048067.09	0.00	52.83
area	8	271009.12	24618	777057.97	0.00	12.20
MonthStart	11	74036.84	24607	703021.12	0.00	3.33
Year:area	391	96150.61	24216	606870.52	0.00	4.33
area:MonthStart	88	54681.97	24128	552188.54	0.00	2.46

$X^2 \sim \chi^2_{(N-p)}$

statistic	p.value
739652.98	0.00

Pseudo R^2

0.72
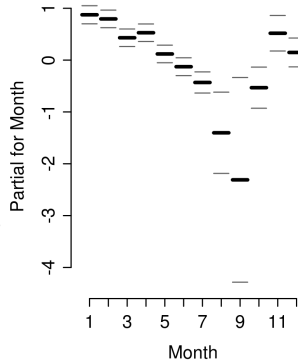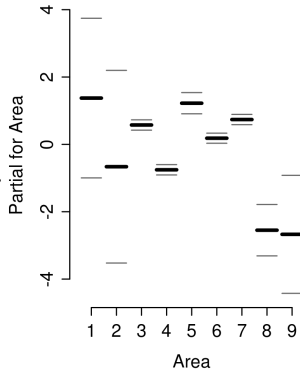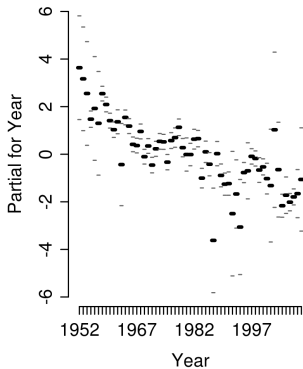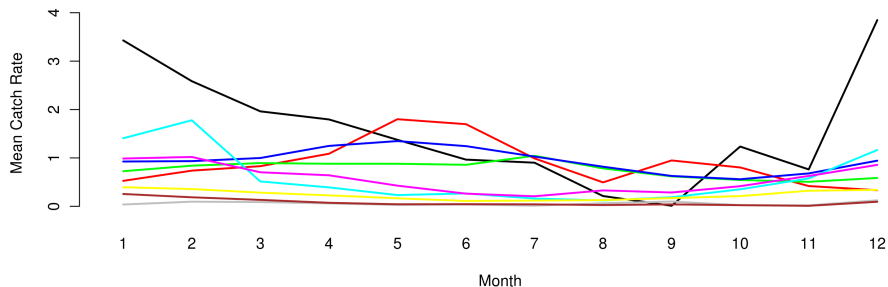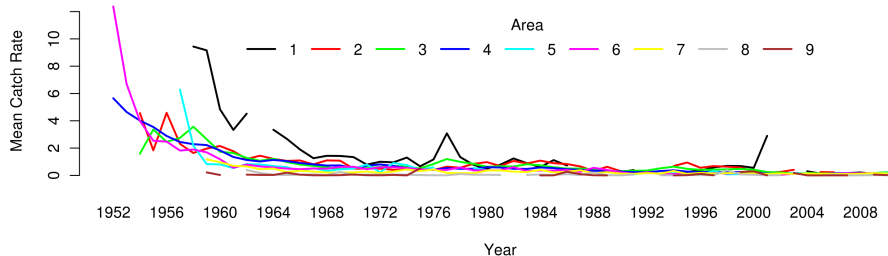
Partial Effects

Partial Effects

Standardazed Catch Rates
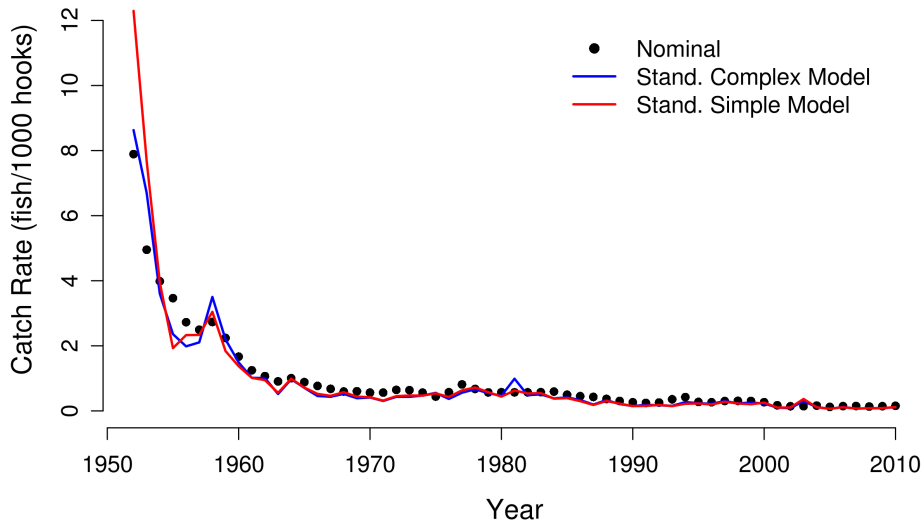
Catch x Standardazed Catch Rates
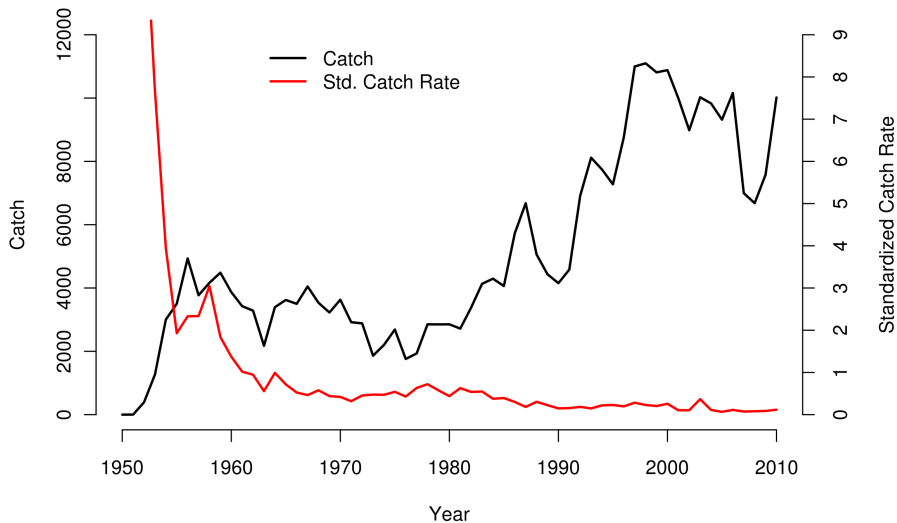
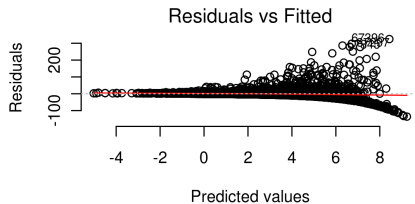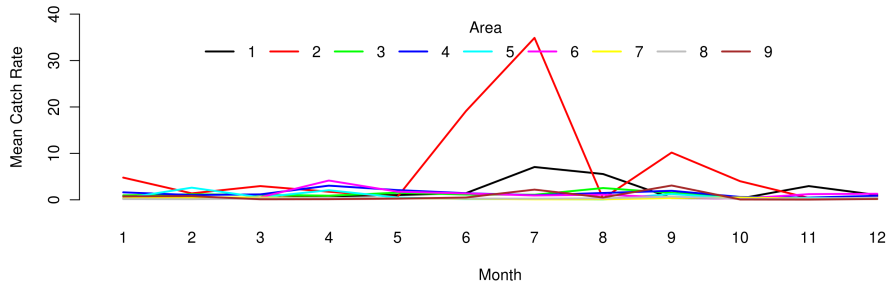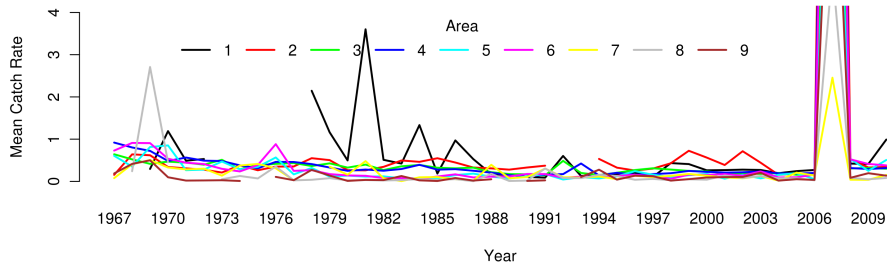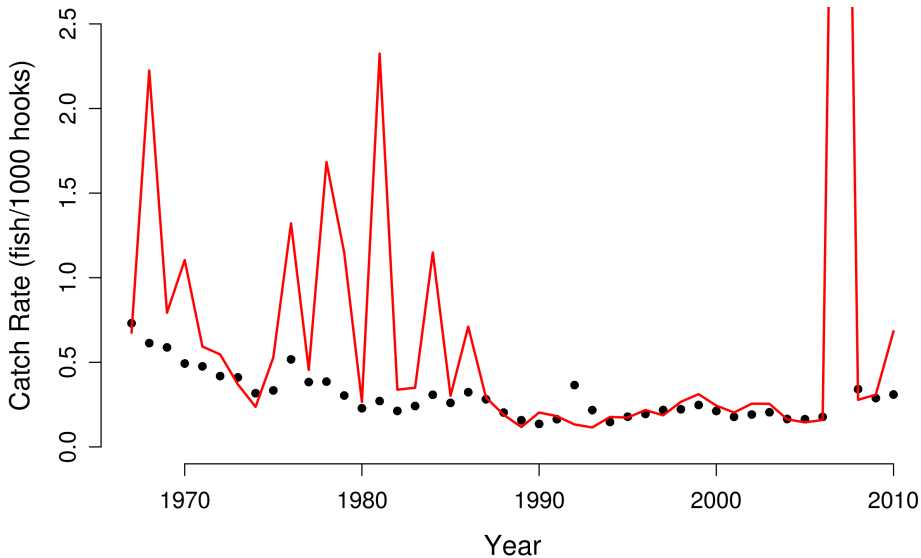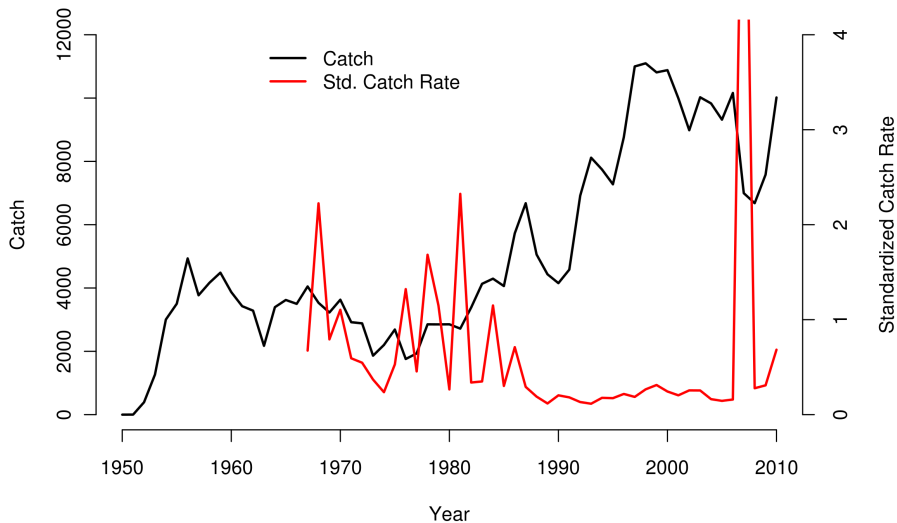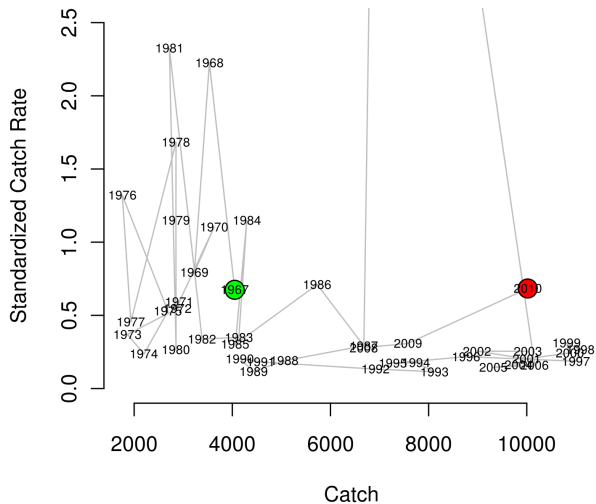# Catch x Standardazed Catch Rates
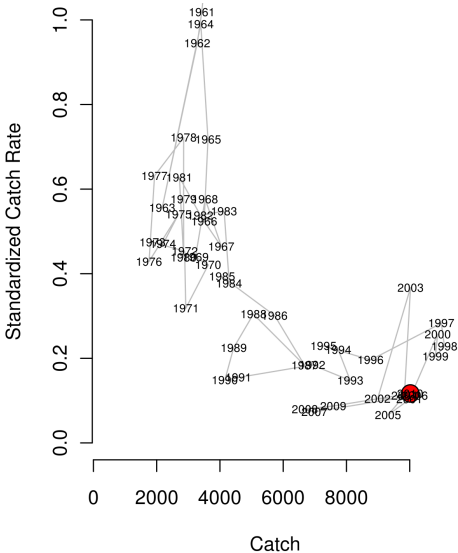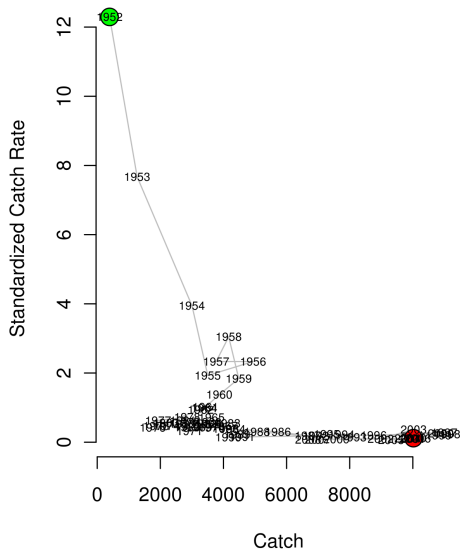
Blue Marlin – Taiwan Dataset

Selected Model

Simple (for illustration purposes)

$$Y \sim year + area + month + year : area + area : month$$

Residuals Diagnostics
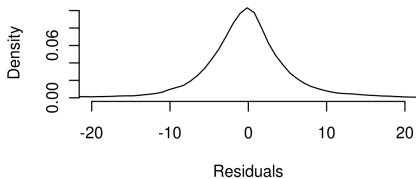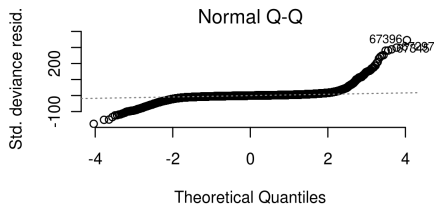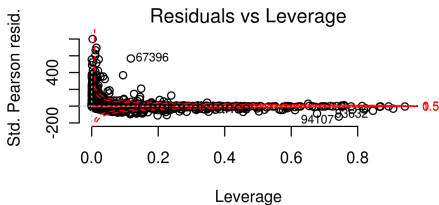
Summary Statistics

Analysis of Deviance

	Df	Deviance	Resid. Df	Resid. Dev	P(>|Chi|)	Dev.Exp2
NULL			18633	7195741.65		
Year	43	2552503.34	18590	4643238.32	0.00	35.47
area	8	375630.55	18582	4267607.77	0.00	5.22
MonthStart	11	131870.30	18571	4135737.47	0.00	1.83
Year:area	329	325724.79	18242	3810012.68	0.00	4.53
area:MonthStart	88	300632.28	18154	3509380.40	0.00	4.18

$X^2 \sim \chi^2_{(N-p)}$

statistic	p.value
11181216.02	0.00

Pseudo R^2

0.51

Partial Effects
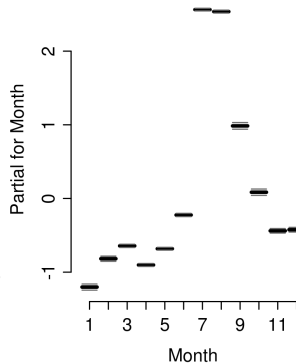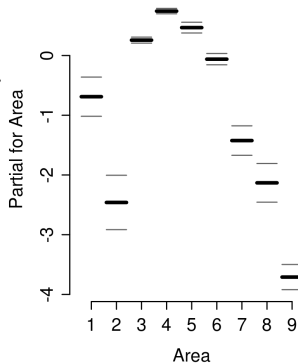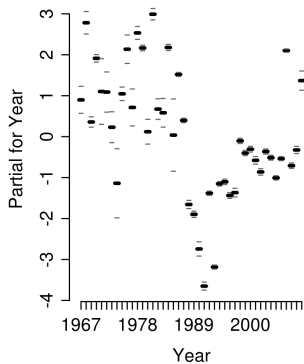
Partial Effects

Standardazed Catch Rates

Catch x Standardazed Catch Rates

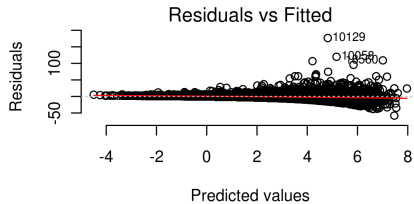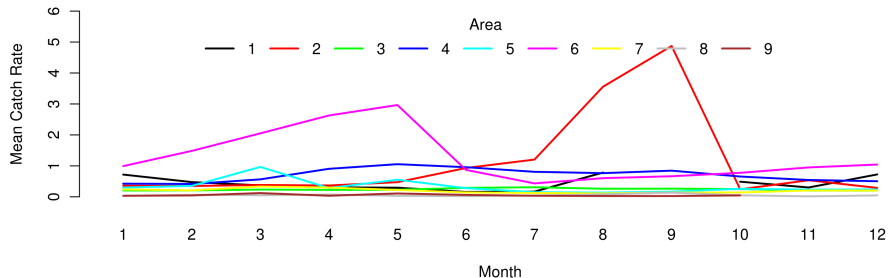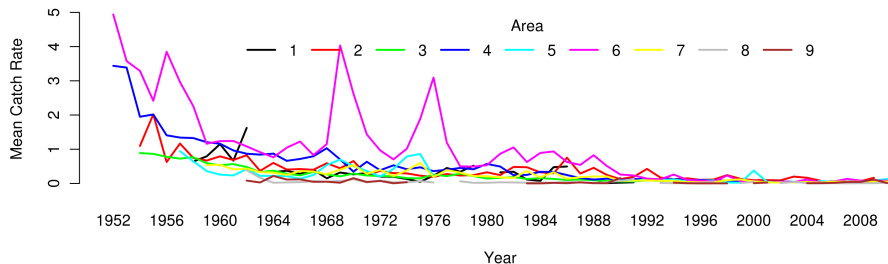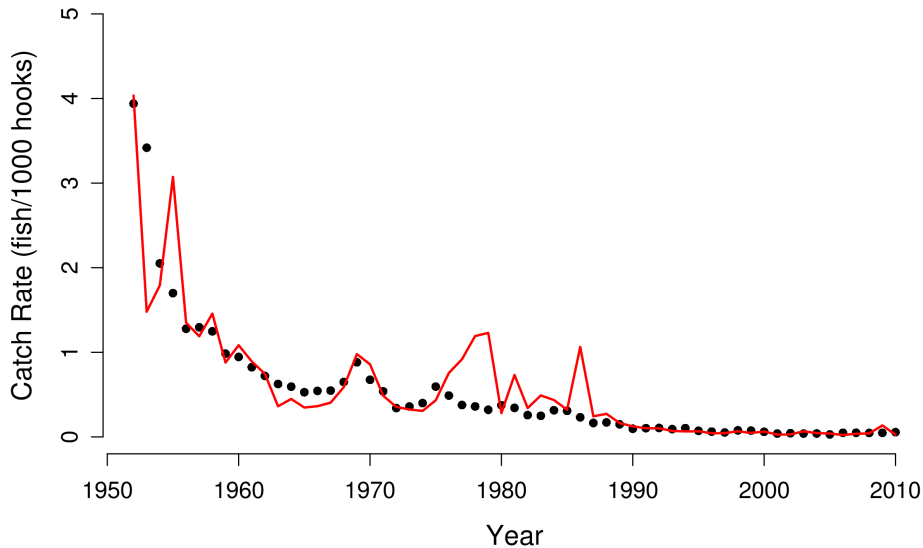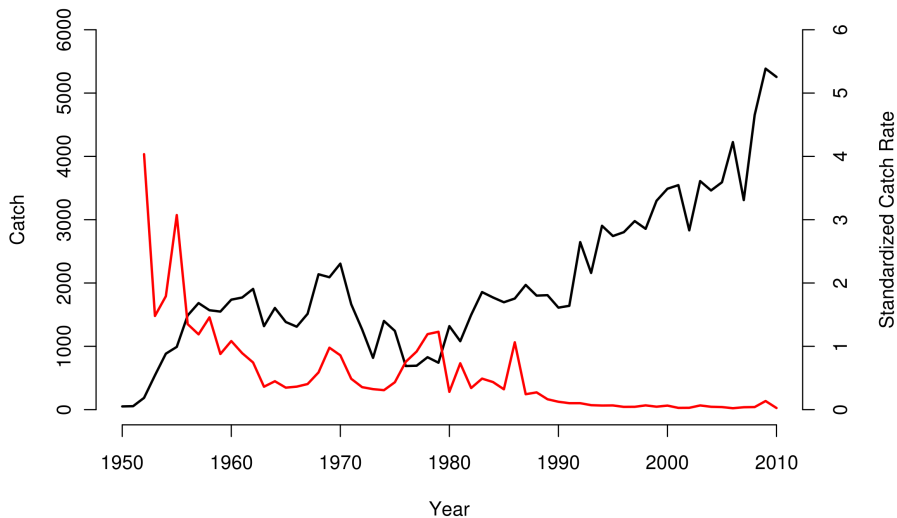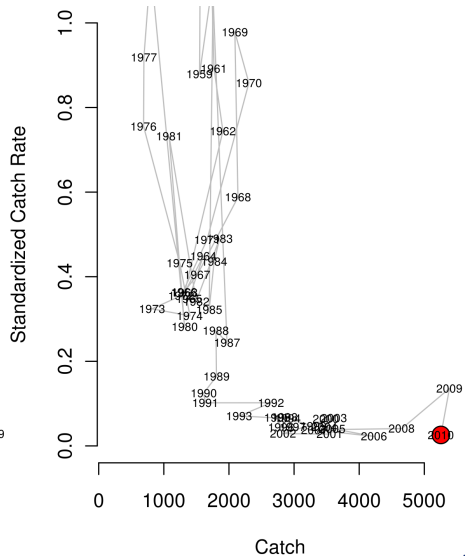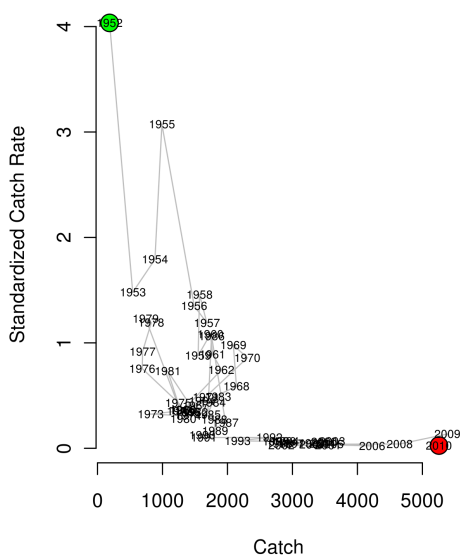# Catch x Standardazed Catch Rates

Black Marlin – Japan Dataset

Selected Model

Simple (for illustration purposes)

$$Y \sim year + area + month + year : area + area : month$$

Residuals Diagnostics
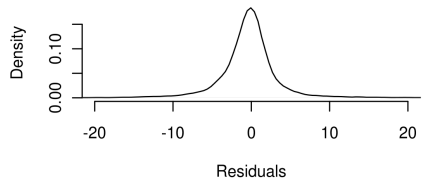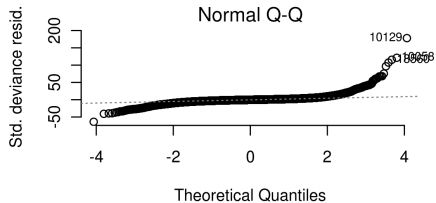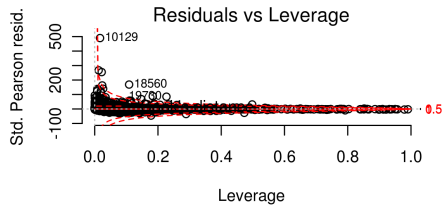
Summary Statistics – Simple Model

Analysis of Deviance

	Df	Deviance	Resid. Df	Resid. Dev	P(>|Chi|)	Dev.Exp2
NULL			20943	1820122.75		
Year	58	835269.42	20885	984853.34	0.00	45.89
area	8	191395.79	20877	793457.55	0.00	10.52
MonthStart	11	22837.28	20866	770620.27	0.00	1.25
Year:area	381	145216.46	20485	625403.80	0.00	7.98
area:MonthStart	86	65579.26	20399	559824.54	0.00	3.60

$X^2 \sim \chi^2_{(N-p)}$

statistic	p.value
1159675.41	0.00

Pseudo R^2

0.73

Partial Effects
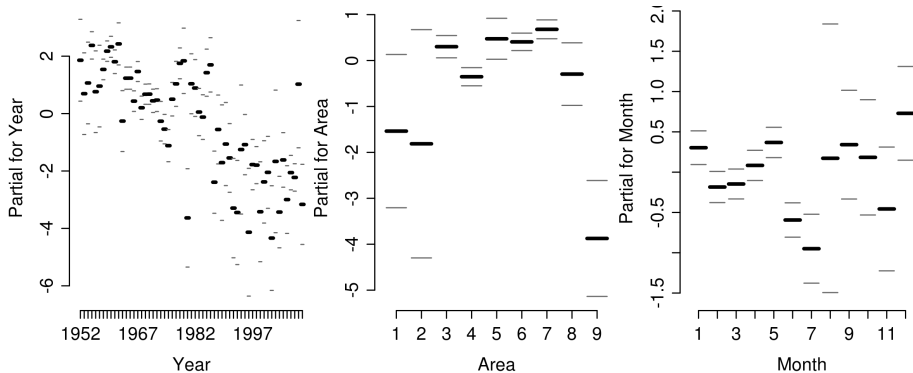
Partial Effects

Standardazed Catch Rates

Catch x Standardazed Catch Rates

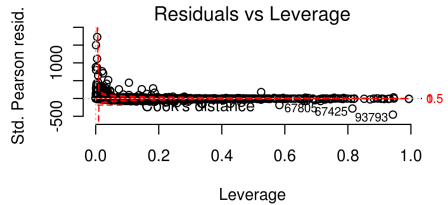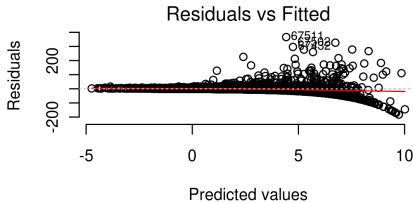# Catch x Standardazed Catch Rates
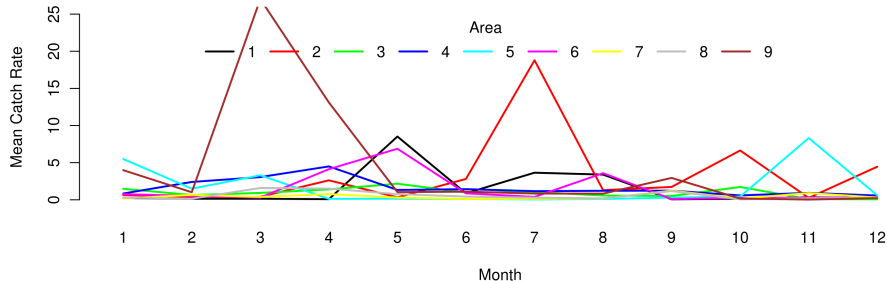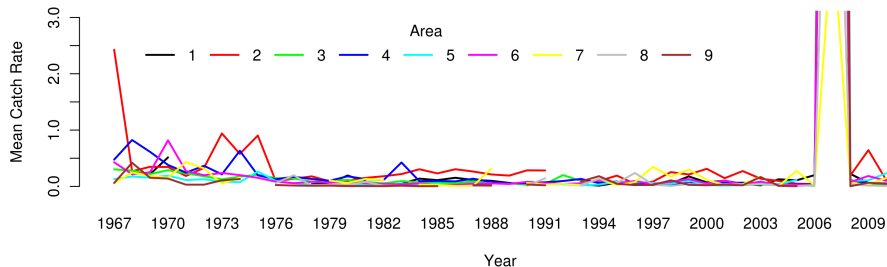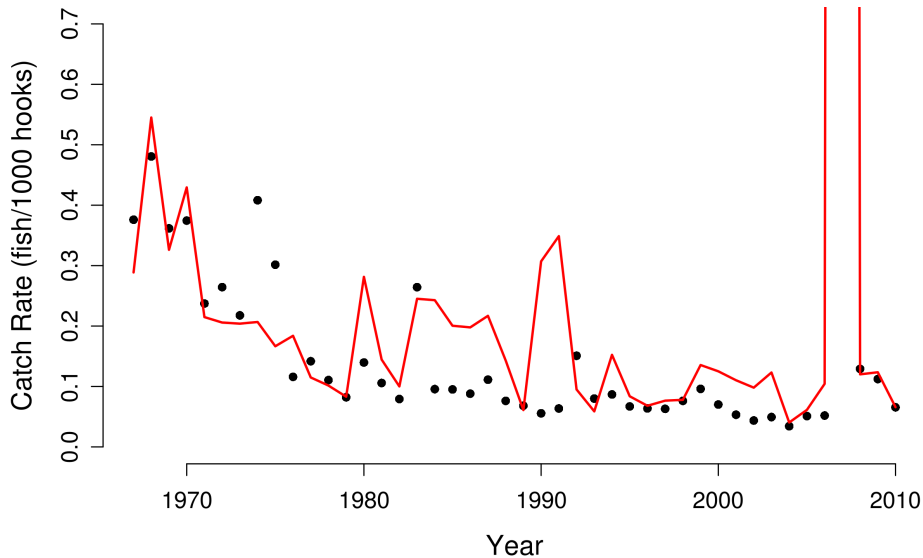
Black Marlin – Taiwan Dataset

Selected Model

Simple (for illustration purposes)

$$Y \sim year + area + month + year : area + area : month$$

Residuals Diagnostics
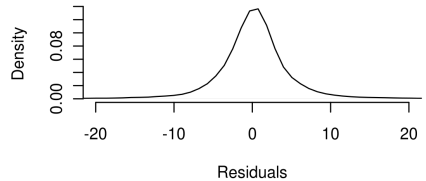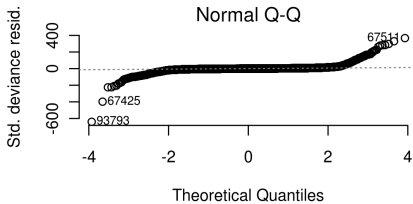
Summary Statistics

Analysis of Deviance

	Df	Deviance	Resid. Df	Resid. Dev	P(>|Chi|)	Dev.Exp2
NULL			11671	7854664.10		
Year	43	3891575.11	11628	3963088.99	0.00	49.54
area	8	460106.53	11620	3502982.46	0.00	5.86
MonthStart	11	108451.85	11609	3394530.61	0.00	1.38
Year:area	310	142989.24	11299	3251541.36	0.00	1.82
area:MonthStart	88	450830.07	11211	2800711.30	0.00	5.74

$X^2 \sim \chi^2_{(N-p)}$

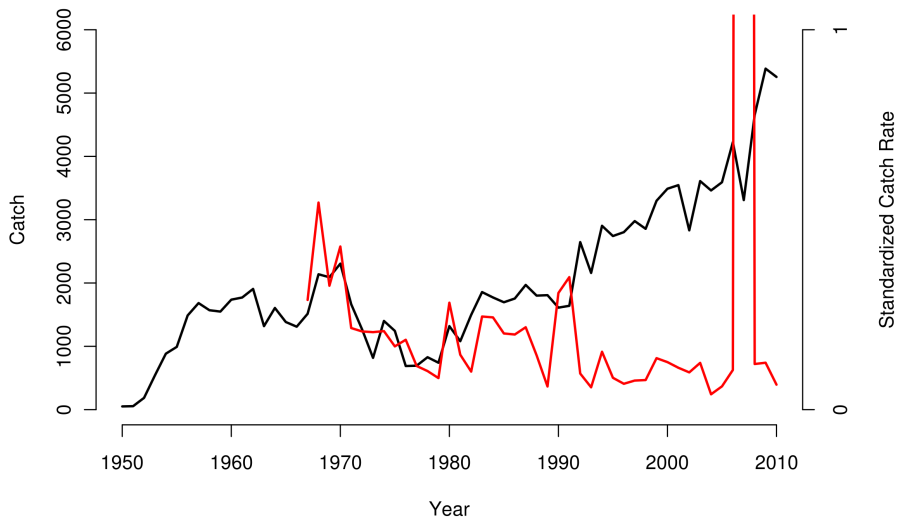| statistic	p.value
14914544.66	0.00

Pseudo R^2
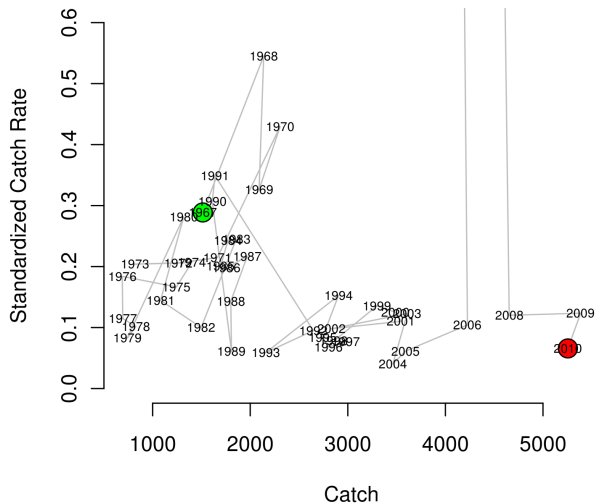
0.38

Partial Effects
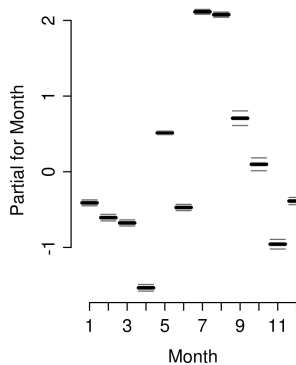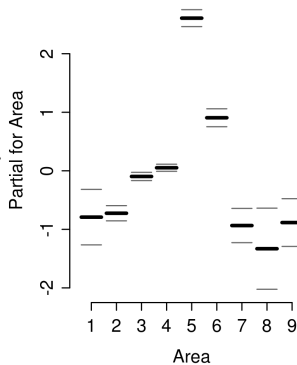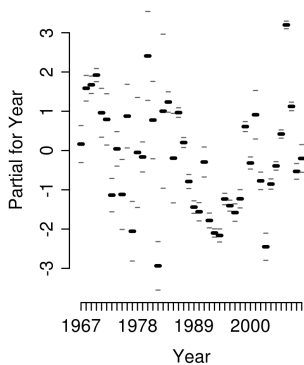
Partial Effects

Standardazed Catch Rates

Catch x Standardazed Catch Rates

Catch x Standardazed Catch Rates

Striped Marlin – Japan Dataset

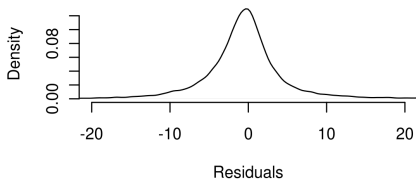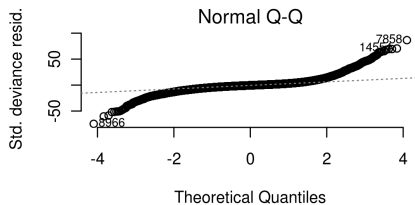# Selected Model
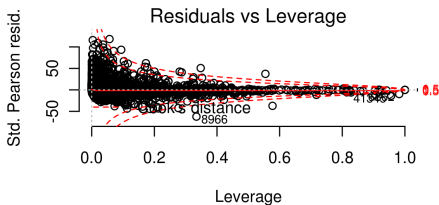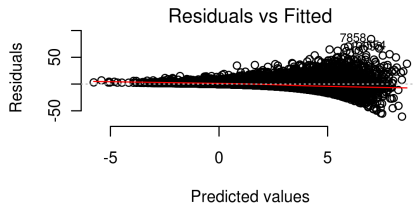
Simple (for illustration purposes)

$$Y \sim year + area + month + year : area + area : month$$
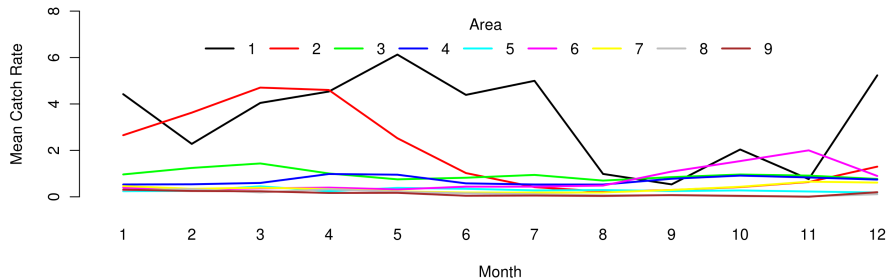
Residuals Diagnostics

Summary Statistics – Simple Model

Analysis of Deviance

	Df	Deviance	Resid. Df	Resid. Dev	P(>|Chi|)	Dev.Exp2
NULL			23999	3794957.47		
Year	58	1424925.77	23941	2370031.71	0.00	37.55
area	8	868970.42	23933	1501061.29	0.00	22.90
MonthStart	11	72718.20	23922	1428343.09	0.00	1.92
Year:area	403	267517.00	23519	1160826.09	0.00	7.05
area:MonthStart	88	174439.81	23431	986386.28	0.00	4.60

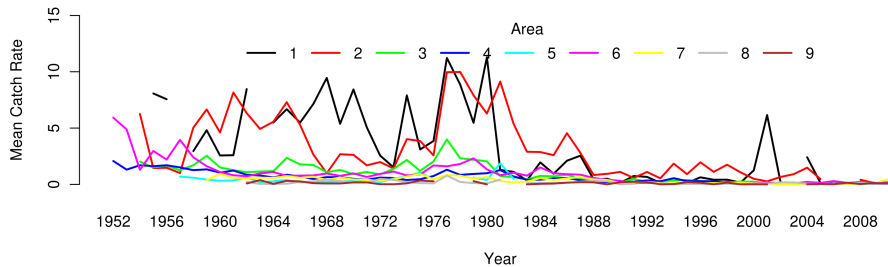### $X^2 \sim \chi^2_{(N-p)}$

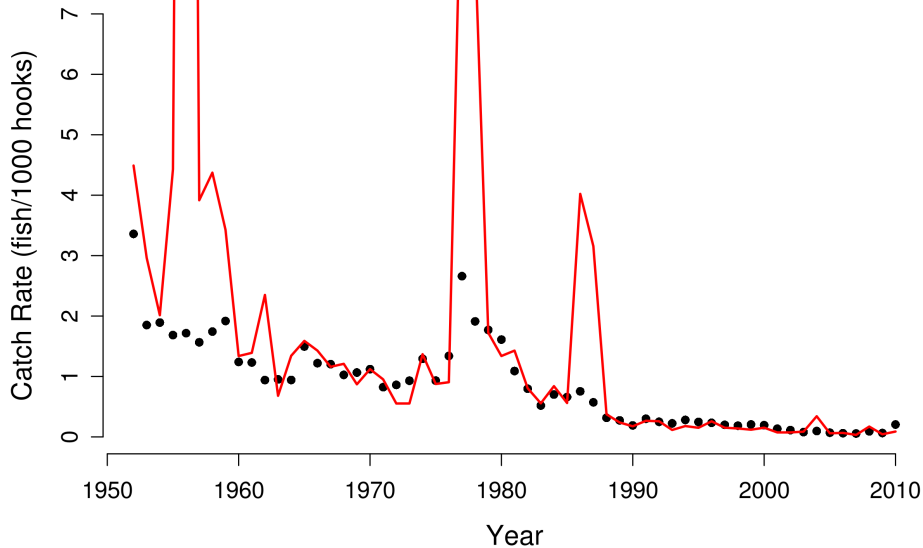| statistic	p.value
1403093.38	0.00

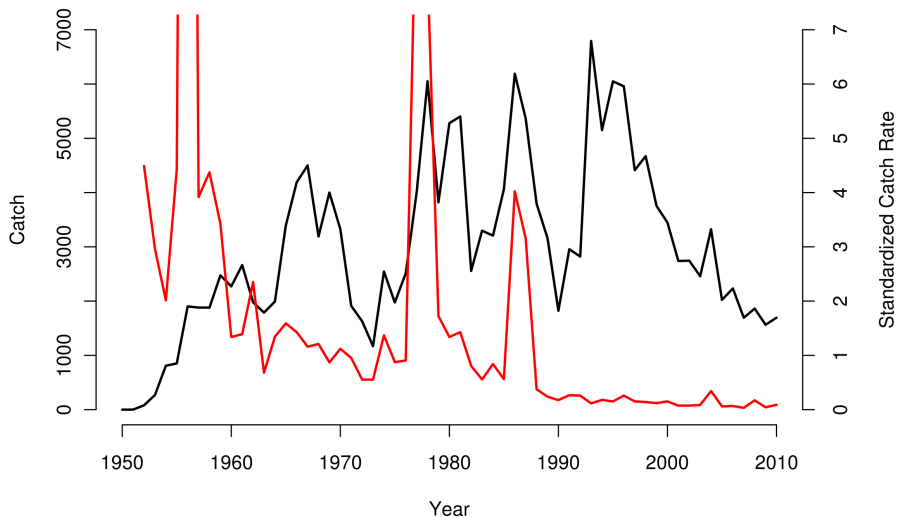### Pseudo R^2
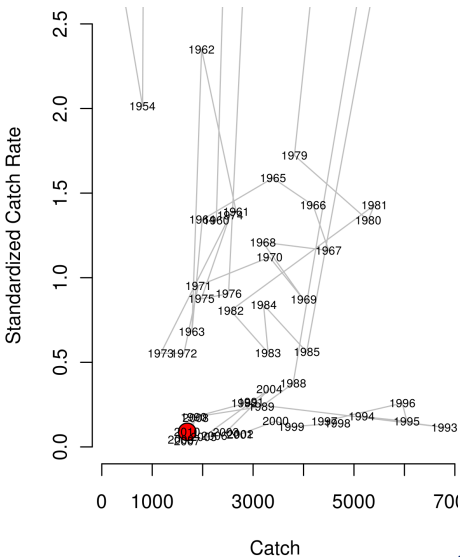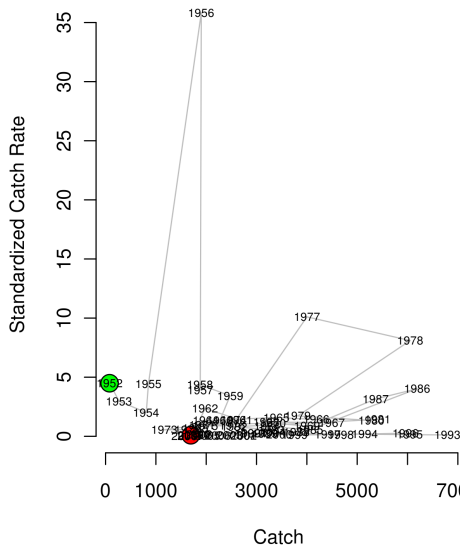
0.72
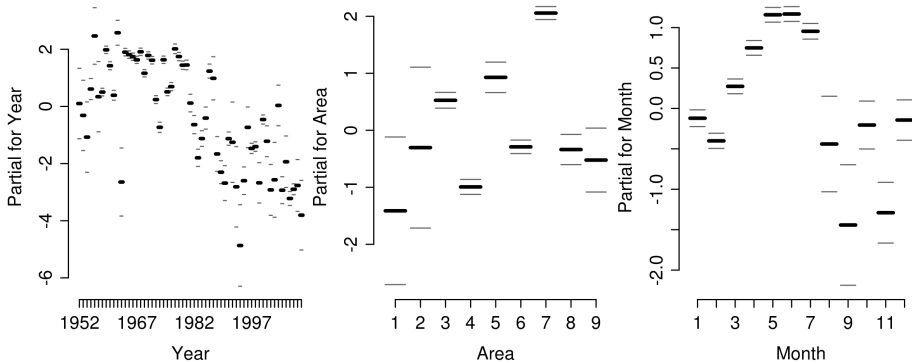
Partial Effects

Partial Effects

Standardazed Catch Rates

Catch x Standardazed Catch Rates
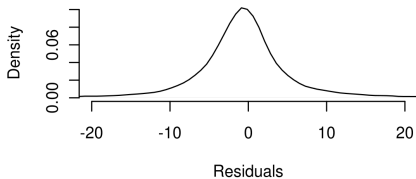
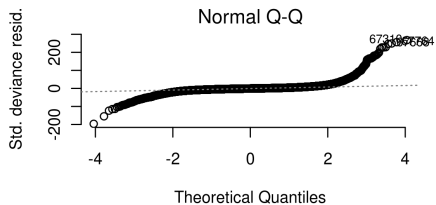# Catch x Standardazed Catch Rates

Striped Marlin – Taiwan Dataset

Selected Model
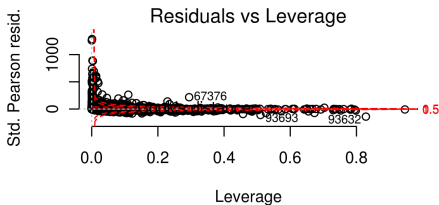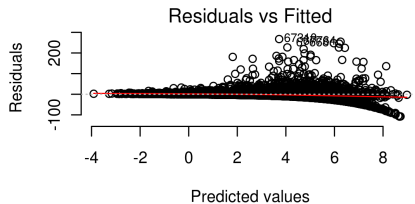
Simple (for illustration purposes)

$$Y \sim year + area + month + year : area + area : month$$

Residuals Diagnostics

Summary Statistics

Analysis of Deviance

	Df	Deviance	Resid. Df	Resid. Dev	P(>|Chi|)	Dev.Exp2
NULL			18989	5921280.84		
Year	43	1699796.39	18946	4221484.45	0.00	28.71
area	8	468414.50	18938	3753069.95	0.00	7.91
MonthStart	11	70022.51	18927	3683047.44	0.00	1.18
Year:area	332	374045.51	18595	3309001.93	0.00	6.32
area:MonthStart	88	243639.81	18507	3065362.12	0.00	4.11

$X^2 \sim \chi^2_{(N-p)}$

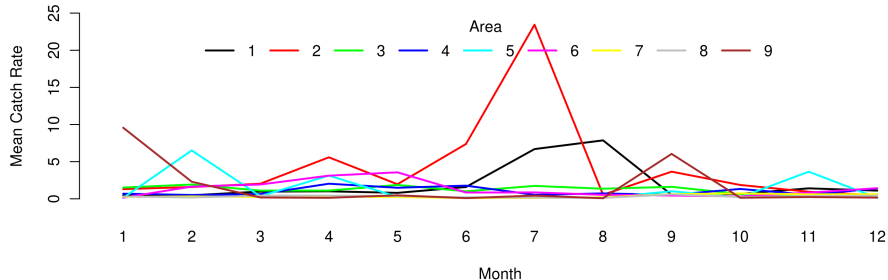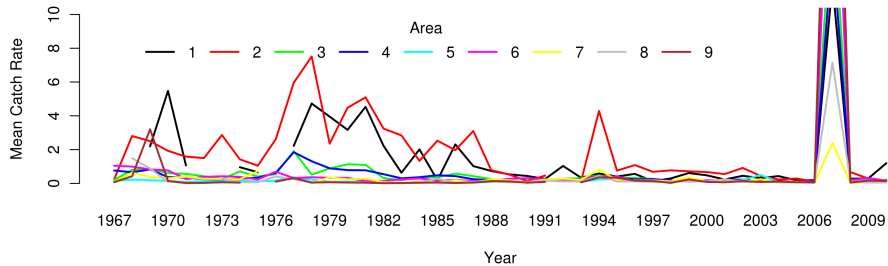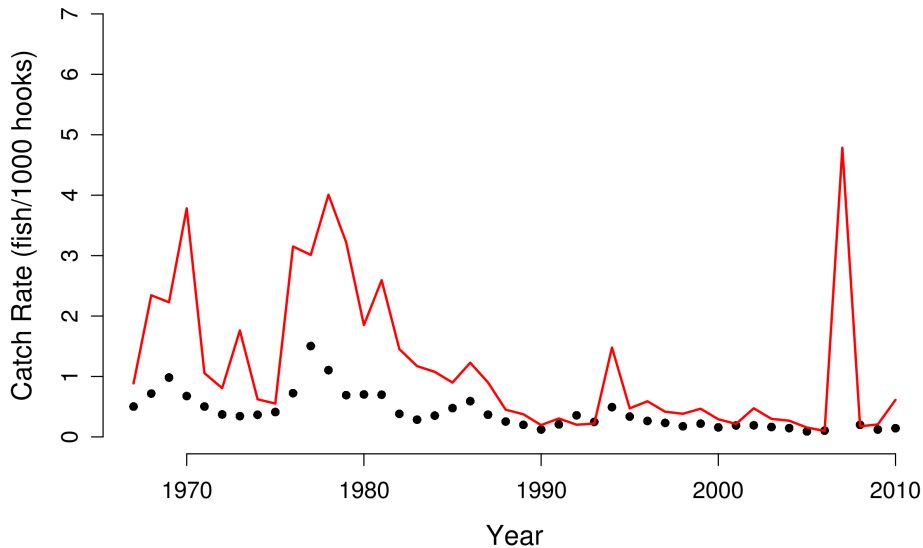| statistic	p.value
16240742.79	0.00

Pseudo R^2
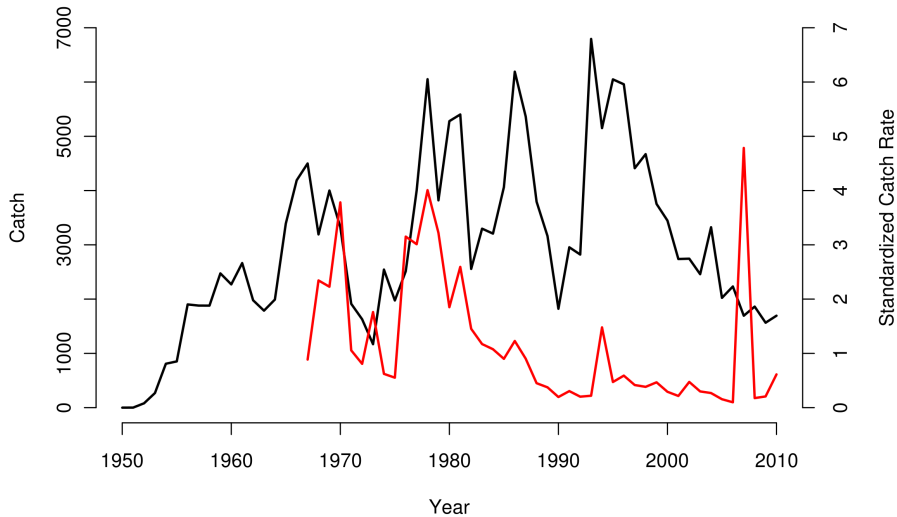
0.48

Partial Effects

Partial Effects

Standardazed Catch Rates

Catch x Standardazed Catch Rates

Catch x Standardazed Catch Rates