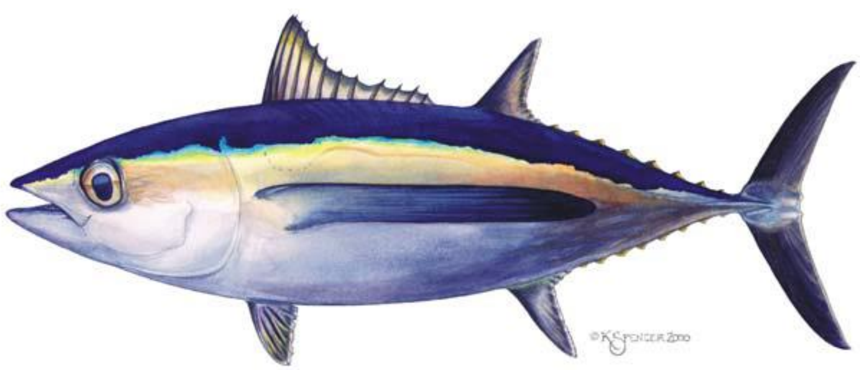
An update of the standardized abundance index of US and Canada Pacific albacore troll and pole-and-line fisheries on the West Coast of North America (1966-2011)[1]

Yi Xu[2], Steven L.H. Teo[2] and John Holmes[3]

[2]NOAA Fisheries
Southwest Fisheries Science Center
8901 La Jolla Shores Drive
La Jolla, CA 92037, USA

[3]Fisheries and Oceans Canada
Pacific Biological Station
3190 Hammond Bay Road
Nanaimo, B.C., V9T 6N7, Canada

Email: Yi.Xu@noaa.gov



[1]Working document submitted to the ISC Albacore Working Group Meeting, 14-28 April 2014, La Jolla, CA, 92037, USA. Document not to be cited without the authors' permission.

ABSTRACT

A merged US-Canada albacore troll/pole-and-line (surface) fisheries dataset was used to obtain a standardized abundance index from 1966 to 2011 for the upcoming 2014 stock assessment of North Pacific albacore tuna (*Thunnus alalunga*). During the November 2013 Working Group (WG) workshop, the WG decided to use the coastal region data and drop the 2012 data point in the index for the 2014 stock assessment. Therefore, in this paper we provide the final standardized abundance index from GLMs with three time periods (1966-1978, 1979-1998, and 1999-2011). In addition, the associated model diagnostics and summaries for the GLM models are reported. We recommend these standardized coastal ocean CPUE indices be used in a sensitivity run for the upcoming stock assessment.

INTRODUCTION

Albacore tuna have been targeted by the US and Canadian troll/pole-and-line (surface) fisheries in the North Pacific for over half a century. These fisheries are tightly integrated operationally since the 1980s owing to the influence of a bilateral treaty governing access to waters in each country. In the last albacore stock assessment, a generalized linear model (GLM) was used on a merged US-Canadian surface fisheries logbook dataset to obtain a standardized abundance index of North Pacific albacore tuna (ISC, 2011). The model considered areas and seasons as factors in the GLM with two general areas (north of 40°N and south of 40°N) and estimated a standardized CPUE index from 1966 to 2009. We previously updated the US-Canada standardized CPUE index time series to 2012 using the same analysis (Xu et al, 2013a) for the 2014 assessment. The updated CPUE index was based on several spatial regions based on an earlier analysis of the environment and CPUE distribution (Xu et al., 2013b) and was split into three time periods (1966-1978, 1979-1998, and 1999-2012) based on operational changes in fleet behavior, which are likely to affect catchability. This revised index was discussed during the November 2013 workshop of the ALBWG, where it was decided that the 2012 data point could not be used because access provisions in the bilateral treaty were suspended. The objective of this working paper is to re-do the standardization analysis for the coastal regions (within 200 nm of the coast), dropping the 2012 data based on fishery operational changes, as suggested by the working group.

DATA AND METHODS

Data

A database of catch and effort was assembled from the logbook records of US and Canadian albacore surface vessels. We used the same data structure as Teo et al. (2010) with updated US data from 1966-2011 and Canadian data from 2004-2011. Data without latitude and longitude or locations on land were removed from the dataset. Retained and discarded catch were summed to estimate total catch. We further aggregated the merged data into 1x1° spatial blocks by month. Strata with less than 3 boat-days were also removed from further analysis.

CPUE standardization

We calculated the standardized CPUE indices for the merged US-Canada dataset using GLMs (the same method as Teo et al., 2010) and updated the time series to 2011. We assigned the merged US-Canada data into three coastal regions (Fig 1) and used these area factors in the GLM to study the area effect on the abundance indices over time. The three regions are: inshore-north (region-1, 200 nm or less offshore, north of 48°N), inshore-central (region-2, 200 nm or less offshore, 40-48°N), inshore-south (region-3, 200 nm or less offshore, south of 40°N). Similarly, we assigned the data into 3 fishing seasons: season 2 (Apr-June), season 3 (Jul-Sep), and season 4 (Oct-Dec). We further split the time series into three periods:

1966-1978, 1979-1998, and 1999-2011, as suggested by the working group. By splitting the time series and running GLMs separately, there is more flexibility in accounting for catchability changes between the different time periods. For each of the 3 periods, the log-transformed CPUE was related to year, area and season by

$$\ln(CPUE_{ijk} + 1) = X + Y_i + S_j + A_k + \varepsilon_{ijk},$$

where $CPUE_{ijk}$ is the CPUE (fish per boat fishing day) in year i, season j, and area k, and X is the intercept term, and ε_{ijk} is the random error term. The standardized CPUE indices, I_t, were obtained by calculating the population marginal means of the above model for each given year and back-transforming the result using

$$I_t = \exp(\hat{\alpha}_t + \hat{\sigma}_t^2/2),$$

where $\hat{\alpha}_t$ is the estimated year factor and $\hat{\sigma}_t$ is the standard error of the estimated year factor, which reduces the log-transformation bias. Confidence intervals of the abundance indices were estimated from 1000 bootstrap runs.

RESULTS AND DISCUSSION

The overall trends in the abundance indices did not change substantially when we dropped the 2012 data (Fig 2). Similar to previous findings, the confidence intervals for 1966-1978 are larger compared with 1979-1998 and 1998-2011 because when bootstrapping was performed separately, the 1966-1978 period had fewer data with larger variability. The standardized CPUE appeared to have a reduced variability relative to nominal CPUE (Fig 2, black line).

The diagnostics of three GLM models associated with three time periods are in Fig 3. The residual plots seem to show a skewed distribution, with higher magnitudes of negative residuals (over estimating expected values) and this seems to be reflected in the QQ plots. The summaries of three GLM model results are listed in the Appendix. We continue to recommend that this standardized coastal ocean CPUE index, separated into three time periods, be used in a sensitivity run for the upcoming stock assessment.

REFERENCES

ISC, 2011. Annex 9. Report of 11th meeting the Albacore Working Group Workshop. 20-25 July 2011, San Francisco, California, USA. Stock assessment of albacore tuna in the North Pacific Ocean in 2011.

Teo, S.L.H., Holmes, J., and Kohin, S., 2010. Joint standardized abundance index of US and Canada albacore troll fisheries in the North Pacific. ISC Working Paper (ISC/10-3/ALBWG/01).

Xu, Y., Teo, S.L.H., and Holmes, J., 2013a. An update of the standardized abundance index of US and Canada albacore troll fisheries in the North Pacific (1966-2012). ISC Working Paper (ISC-13-ALBWG-03-06).

Xu, Y., Nieto, K., Teo, S.L.H., McClatchie, S., Holmes, J., 2013. Influence of subtropical fronts on the spatial distribution of albacore tuna (*Thunnus alalunga*) in the Northeast Pacific over the past 30 years (1982-2011). (*Progress in Oceanography*, in review).

FIGURES

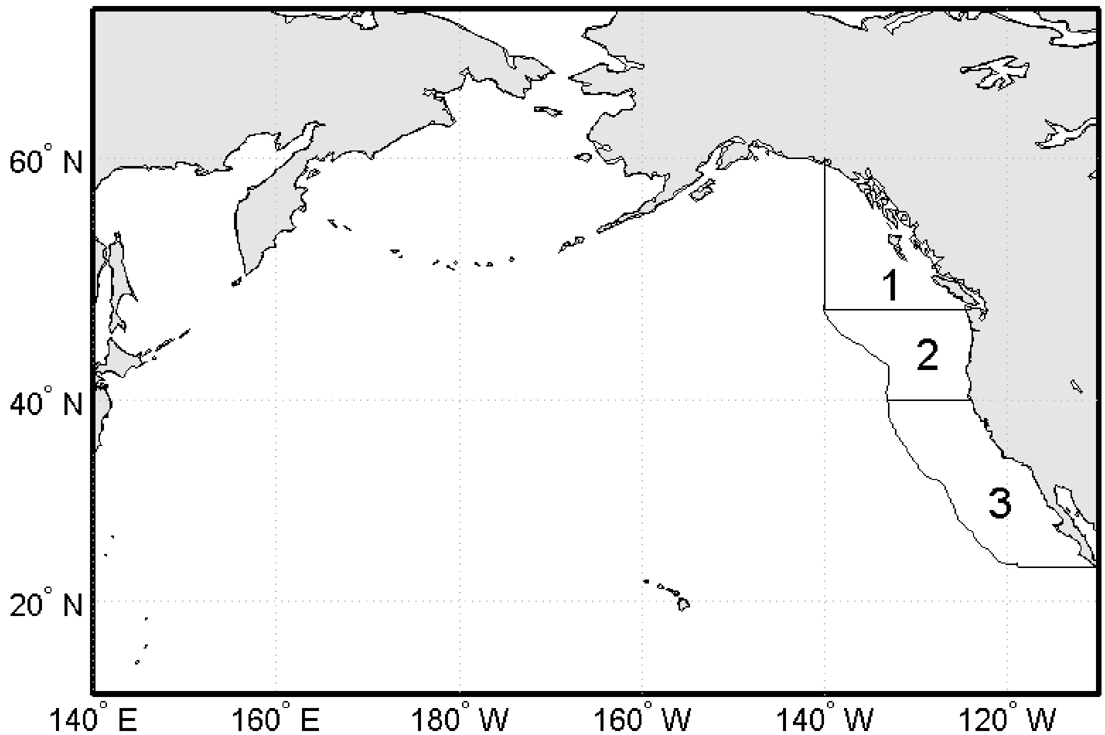


Figure 1. Map showing subdivision regions for constructing standardized CPUE abundance indices for the EPO surface fisheries: inshore-north (Region-1, 200 nm or less offshore, north of 48°N), inshore-central (Region-2, 200 nm or less offshore, 40-48°N), and inshore-south (Region-3, 200 nm or less offshore, south of 40°N).

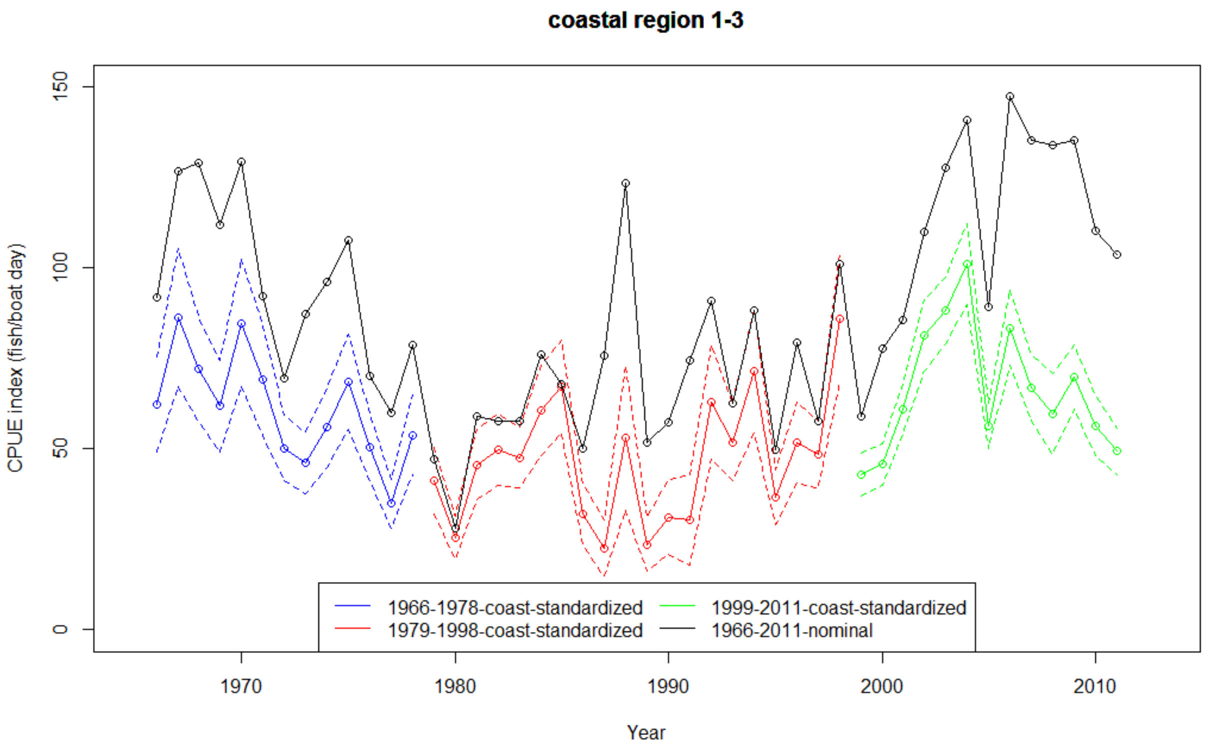


Figure 2. Comparison of standardized CPUE indices for coastal ocean, Regions 1-3 combined. GLM model run separately for to 1966-1978 (blue), 1979-1998 (red), 1999-2011 (green). Black line is nominal CPUE.

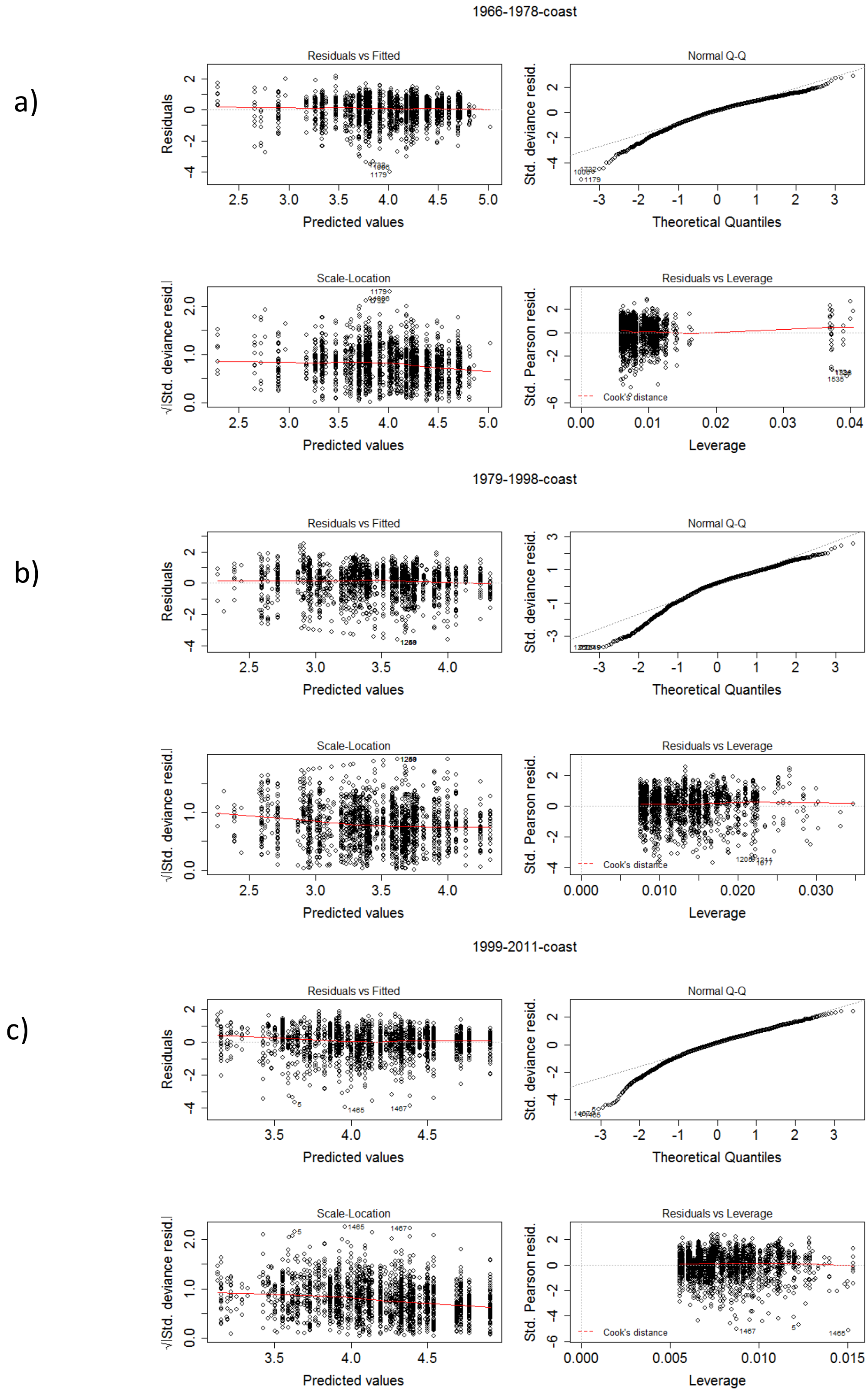


Figure 3. GLM diagnostic plots for coastal ocean indices (Regions 1-3 combined) of three periods: a) 1966-1978; b) 1979-1998; and c) 1999-2011.

Appendix 1. GLM results summary for Fig 2, CPUE standardization with area and season effect during 1966-1978, 1979-1998, and 1999-2011.

a) 1966-1978

```
Call:
glm(formula = alb.model, family = gaussian, data = CPUE.IN)

Deviance Residuals: 
    Min       1Q   Median       3Q      Max  
-4.0162  -0.4167   0.1180   0.5227   2.1604  

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept)  4.713054   0.097511  48.333  < 2e-16 ***
year.f1967   0.328126   0.107646   3.048  0.00233 ** 
year.f1968   0.151072   0.105020   1.439  0.15046    
year.f1969  -0.005041   0.097691  -0.052  0.95885    
year.f1970   0.310731   0.098725   3.147  0.00167 ** 
year.f1971   0.108285   0.104321   1.038  0.29940    
year.f1972  -0.212496   0.095500  -2.225  0.02619 *  
year.f1973  -0.299361   0.099173  -3.019  0.00257 ** 
year.f1974  -0.103683   0.096019  -1.080  0.28036    
year.f1975   0.098201   0.095391   1.029  0.30340    
year.f1976  -0.209347   0.092110  -2.273  0.02315 *  
year.f1977  -0.578652   0.092252  -6.273 4.39e-10 ***
year.f1978  -0.142566   0.093525  -1.524  0.12759    
season.f2   -1.050730   0.141977  -7.401 2.03e-13 ***
season.f3   -0.442221   0.052009  -8.503  < 2e-16 ***
area.f2     -0.321456   0.066304  -4.848 1.35e-06 ***
area.f3     -0.795026   0.067084 -11.851  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 0.5726061)

    Null deviance: 1486.9  on 1903  degrees of freedom
Residual deviance: 1080.5  on 1887  degrees of freedom
AIC: 4360.7

Number of Fisher Scoring iterations: 2
```

b) 1979-1998

```
Call:
glm(formula = alb.model, family = gaussian, data = CPUE.IN)

Deviance Residuals: 
    Min       1Q   Median       3Q      Max  
-3.6197  -0.5119   0.1781   0.6647   2.5185  

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  3.58877    0.12312  29.149  < 2e-16 ***
year.f1980  -0.48185    0.13805  -3.490 0.000494 ***
year.f1981   0.10069    0.13121   0.767 0.442953    
year.f1982   0.19041    0.12817   1.486 0.137548    
year.f1983   0.14352    0.12586   1.140 0.254308    
year.f1984   0.38643    0.12497   3.092 0.002017 ** 
year.f1985   0.48754    0.13123   3.715 0.000209 ***
year.f1986  -0.24915    0.14920  -1.670 0.095101 .  
year.f1987  -0.60499    0.14464  -4.183 3.02e-05 ***
year.f1988   0.25346    0.16000   1.584 0.113343    
year.f1989  -0.55817    0.15334  -3.640 0.000280 ***
year.f1990  -0.28037    0.16930  -1.656 0.097892 .  
```

```
year.f1991  -0.30223    0.17544  -1.723 0.085121 .
year.f1992   0.42167    0.15167   2.780 0.005488 **
year.f1993   0.23058    0.15323   1.505 0.132548
year.f1994   0.54895    0.15754   3.485 0.000505 ***
year.f1995  -0.12004    0.15799  -0.760 0.447474
year.f1996   0.22549    0.14534   1.551 0.120985
year.f1997   0.16372    0.13343   1.227 0.219987
year.f1998   0.73459    0.14389   5.105 3.66e-07 ***
season.f2   -0.27668    0.14414  -1.919 0.055082 .
season.f3   -0.33086    0.07612  -4.346 1.46e-05 ***
area.f2     -0.07367    0.07903  -0.932 0.351420
area.f3     -0.39078    0.09066  -4.310 1.72e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 0.9752574)

    Null deviance: 2047.2  on 1815  degrees of freedom
Residual deviance: 1747.7  on 1792  degrees of freedom
AIC: 5133.9

Number of Fisher Scoring iterations: 2
```

c) 1999-2011

```
Call:
glm(formula = alb.model, family = gaussian, data = CPUE.IN)

Deviance Residuals: 
    Min       1Q   Median       3Q      Max  
-3.9603  -0.3731   0.0994   0.5059   1.8706  

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  3.59177    0.06950  51.678  < 2e-16 ***
year.f2000   0.06413    0.08516   0.753 0.451505    
year.f2001   0.35272    0.08297   4.251 2.22e-05 ***
year.f2002   0.63844    0.08656   7.376 2.34e-13 ***
year.f2003   0.72143    0.09009   8.008 1.92e-15 ***
year.f2004   0.85643    0.08821   9.709  < 2e-16 ***
year.f2005   0.27393    0.08108   3.379 0.000742 ***
year.f2006   0.66578    0.08825   7.544 6.74e-14 ***
year.f2007   0.44522    0.08277   5.379 8.32e-08 ***
year.f2008   0.32975    0.09372   3.518 0.000443 ***
year.f2009   0.48752    0.08474   5.753 1.01e-08 ***
year.f2010   0.27434    0.08218   3.338 0.000858 ***
year.f2011   0.14058    0.08087   1.738 0.082299 .  
season.f2   -0.42518    0.06443  -6.599 5.23e-11 ***
season.f3   -0.40581    0.04382  -9.261  < 2e-16 ***
area.f2      0.46399    0.04484  10.347  < 2e-16 ***
area.f3     -0.03631    0.06430  -0.565 0.572348    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 0.603626)

    Null deviance: 1597.9  on 2116  degrees of freedom
Residual deviance: 1267.6  on 2100  degrees of freedom
AIC: 4958.1

Number of Fisher Scoring iterations: 2
```