Fisheries and Oceans Canada | Pêches et Océans Canada
NOAA FISHERIES
NATIONAL MARINE FISHERIES SERVICE

CERT

Comité d'évaluation des ressources transfrontalières

Document de référence 2010/12

Ne pas citer sans
autorisation des auteurs

TRAC

Transboundary Resources Assessment Committee

Reference Document 2010/12

Not to be cited without
permission of the authors

Exploratory Analysis of Fishery Data for Georges Bank Yellowtail Flounder

Saang-Yoon Hyun[1], Steve X. Cadrin[2], and Sally Roman[1]

[1]University of Massachusetts at Dartmouth
School for Marine Science and Technology
Department of Fisheries Oceanography
200 Mill Road, Suite 325
Fairhaven MA 02719

[2]National Marine Fisheries Service
Northeast Fisheries Science Center
166 Water Street
Wood's Hole, MA, 02543

Ce document est disponible sur l'Internet à :
This document is available on the Internet at :
http://www.mar.dfo-mpo.gc.ca/science/TRAC/trac.html

Canada

TABLE OF CONTENTS

ABSTRACT

Data on yellowtail flounder caught by the otter trawl fishery on Georges Bank during 2000-2010 indicated that there was no trend in catch per unit effort (CPUE) by depth (i.e., neither deeper nor shallower) over year, although there was annual and seasonal variability in the CPUE. Depth, bottom water temperature, their interaction term, and the quadratic term of depth and water temperature were explanatory variables in the best linear model where the CPUE was the response variable. The significant interaction term means that the effect of depth or bottom water temperature on the CPUE is not independent. The significant interaction term of depth and bottom water temperature means that the effect of either on the CPUE is not independent of the other. It implies that other factors such as prey at depth could be interacting with bottom water temperature.

RÉSUMÉ

Aucune tendance dans les prises par unité d’effort (PUE) selon la profondeur n’a été décelée au fil des ans dans les données sur les prises de limande à queue jaune réalisées de 2000 à 2010 dans le cadre de la pêche au chalut à panneaux sur le banc Georges, bien que les PUE variaient selon le moment de l’année et la saison. La profondeur, la température de l’eau au fond, le terme interaction et le terme quadratique de la profondeur et de la température de l’eau étaient des variables explicatives dans le meilleur modèle linéaire où les PUE étaient la variable de réponse. Le fait que le terme interaction soit significatif signifie que l’effet de la profondeur ou de la température de l’eau au fond sur les PUE n’est pas indépen-dant. Le fait que le terme de l’interaction entre la profondeur et la température de l’eau au fond soit significatif signifie que l’effet de l’un sur les PUE n’est pas indépendant de l’autre. Cela veut dire que d’autres facteurs comme la quantité de proies selon la profondeur peuvent inter-agir avec la température de l’eau au fond.

INTRODUCTION

Stock assessment of Georges Bank yellowtail flounder has considerable uncertainties associated with survey catchability (Legault et al. 2009). The current stock assessment method splits survey series in 1994/1995, resulting in the appearance of a substantial increase in survey catchability after 1994 (up to six times greater at some ages). Although apparent changes in catchability may be aliasing other model misspecifications, potential causes of increased survey catchability should be explored. Information from survey data suggests a shift in geographic distribution of yellowtail flounder to more northern and deeper habitats (Nye et al. 2009), and reports from fishermen also indicate a shift in yellowtail to deeper waters (S. Roman, personal communication). Previous analyses of survey data suggest that yellowtail flounder prefer a relatively narrow depth range and tolerate a wide range of temperature on Georges Bank (Murawski and Finn, 1988) and from Cape Hatteras to the Scotian Shelf (Helser and Brodziak, 1996). This analysis was conducted to explore fishery-dependent data for patterns of catch by depth and temperature.

METHODS

Data

We examined data from the study fleet program of the School for Marine Science & Technology (SMAST). The program was established in November 2000 as a cooperative effort between SMAST and the Massachusetts Fisheries Recovery Commission and continued with support from the Massachusetts Marine Fisheries Institute. The fleet deployed otter trawl gear, targeting fish on Georges Bank, in the Gulf of Maine, and in Southern New England waters in a multispecies groundfish fishery. The project also collected environmental data as well as catch data. We analyzed data on yellowtail flounder catch (weight in pounds), fishing time (year and month), fishing effort (tow duration in hours), location (latitude and longitude), water depth (fathoms), and bottom water temperature (^{o}C). We defined catch as the sum of pounds kept and discarded.

Statistical Model

We used catch per unit effort (CPUE) as the response variable, and considered the following data explanatory variables: year, month, depth, and water temperature. Testing the effect of locations of latitude and longitude on the CPUE was meaningless because the area fished (40.0535 N ~ 42.2530 N, and 70.29117 W ~ 66.42033 W) was limited only to part of Georges Bank.

We transformed CPUE data by the natural logarithm to meet the assumption that the response variable is from a normal (Gaussian) distribution. Later we checked the validity of the assumption. Using analysis of deviance (ANODEV), we built a linear model to do hypothesis tests, and to detect a relationship between the response and explanatory variables. We started from a null model Ψ to a full model full (saturated) model Ω, where deviance of the former model (D_Ψ) is larger than that of the latter model (D_Ω), and the degrees of freedom of the former (df_Ψ) is also larger than that of the latter (df_Ω). We used a F-statistic to test an explanatory variable added to the null model (McCullagh and Nelder 1989; Faraway 2006).

$$\frac{\Delta D / \Delta df}{\phi} \sim F_{\Delta df,\ df_\Omega} \qquad (1)$$

where $\Delta D = D_\Psi - D_\Omega$, $\Delta df = df_\Psi - df_\Omega$, and ϕ is a dispersion parameter. Degrees of freedom is the number of data points minus the number of coefficients in a model of interest. In a Gaussian model, the dispersion parameter is estimated as follows:

$$\hat{\phi} = \frac{RSS_\Omega}{df_\Omega} \tag{2}$$

where RSS = residual sum of squares also called error sum of squares (SSE). In a Gaussian model, deviance is RSS (McCullagh and Nelder 1989; Neter, et al 1989): i.e.,

$$D = RSS = \sum_i \left(y_i - \hat{y}_i \right)^2 \tag{3}$$

where y_i is the response variable (= log(cpue$_i$)), and $\hat{y}_i$ is the fitted value from model of interest. Among candidate explanatory variables, year and month are categorical variables whereas bottom water temperature and depth are continuous.

<u>Year, Month, Depth, and Temperature</u>

To test the hypothesis that yellowtail flounder's vertical distribution is different by year and month, we used year, month and depth as candidate explanatory variables. We removed data from 2000 and 2010, because the fishery took place only during one month (December) in 2000, and 2010 data are not complete. Otherwise, it would be too obvious that the CPUE differs by year and month, because of the unbalanced design. When year and month were considered, we should not use water temperature as another covariate, because of a multicollinearity problem.

We tested a hypothesis that yellowtail flounder's vertical distribution differs by water temperature, considering depth and water temperature as candidate explanatory variables. In the case where only depth and temperature are considered, we included all 2000 and 2010 data.

RESULTS AND DISCUSSION

Year, Month and Depth

Year was the most significant factor and month was the second most significant. Although depth was significant when it was the only factor in the model, its effect was not significant (p-value = 0.065) when year and month were included (Table 1). Thus depth was dropped as a candidate explanatory variable. Furthermore, the interaction of year and month was significant. Thus the form of final model was as follows:

$$\log(\mathrm{cpue}_i) = b_0 + b_1 \cdot year_i + b_2 \cdot month_i + b_3 \cdot year_i * month_i \tag{4}$$

where year and month are categorical variables. Coefficients of the model (eq. (4)) are shown in Appendix Table 1. The main effects of year and month indicate year-to-year and month-to-month variability in the CPUE. The interaction term of year and month means that monthly CPUE is different over year, or yearly CPUE is different by month (Fig. 1). However, the significant interaction term of year and month appears to be due to the unbalanced design of data (Fig. 1). For example, CPUE data are available only during August through December in 2002 while those are available only during January through July in 2004 (Fig. 1).

Depth and Temperature

Environmental data such as bottom water temperature change by year and month, and thus we dropped those categorical variables (year and month) when exploring the temperature as a candidate explanatory variable.

CPUE data were mainly distributed over depth of 19-60 fathoms, and over temperature of 5-12^{o}C. Water temperature, depth, and their interaction term were all significant. Furthermore, the quadratic terms of depth and temperature were significant, and improved the model fitness (e.g. lower AIC, and better residual distribution; Table 2 and Fig. 2).

$$\log(\mathrm{cpue}_i) = 4.929 - 0.291 \cdot \mathrm{temp}_i - 0.015 \cdot \mathrm{depth}_i + 0.013 \cdot \mathrm{temp}_i * \mathrm{depth}_i - 0.001(\mathrm{depth}_i^{\,2}) - 0.014 \cdot (\mathrm{temp}_i^{\,2}) \qquad (5)$$

We considered the model in eq. (5) the final model, based on AIC, Q-Q plot of residuals, and a biological meaning. The final model had the lower AIC of 16 585.7 (Model B6 in Table 2), and its residuals indicated a better normality (Fig. 2).

The quadratic terms of depth and temperature respectively were significantly negative (P-value = 0.000; Table 3). The negative quadratic term led to a convex shape of the CPUE against depth and water temperature (Fig. 3), and the shape indicates there is an optimal depth and temperature range. Predicted CPUE from final model B6 was highest over depth range of 40-60 fathoms and temperature range of 5-8^{o}C (Fig. 3). However, we caution about interpretation of the effects of depth and water temperature. Although they are all significant, they cannot be singled out because of their significant interaction term, which means that the effect of depth on CPUE is not independent of water temperature, and vice versa. For this reason, we provided 3-dimensional figures, illustrating the CPUE against both depth and temperature (Fig. 4).

CONCLUSION

There was year-to-year and month-to-month variability in CPUE data, but there was no trend in depth of CPUE data (deeper or shallower) over year. Depth alone did not significantly reduce the deviance of a model after year and month reduced it. A change of CPUE data in depth seems to be more due to water temperature. The best model includes depth, water temperature, their interaction term, and the quadratic term of depth and water temperature. The significant interaction term of depth and water temperature implies that other factors such as prey availability in addition to water temperature could co-act.

ACKNOWLEDGEMENTS

We were funded by Massachusetts Marine Fisheries Institute, and data were provided by the SMAST study fleet program. We thank Chris Legault for his guidance.

REFERENCES

Faraway, J.J. 2006. Extending the linear model with R : Generalized Linear, Mixed Effects and Nonparametric Regression Models. Chapman and Hall / CRC, Boca Raton.

Helser, T.E., and J.K.T. Brodziak. 1996. Influence of Temperature and Depth on Distribution and Catches of Yellowtail Flounder, Atlantic Cod and Haddock in NEFSC Bottom Trawl Surveys. NEFSC Ref. Doc. 96- 05e.

Legault, C.M., L.Alade, and K.J. Clark. 2009. Stock Assessment of Georges Bank Yellowtail Flounder for 2009. TRAC Ref. Doc. 2009/03

McCullagh, P., and J.A. Nelder. 1989. Generalized Linear Models. Chapman and Hall, London ; New York.

Murawski, S.A., and J.T. Finn. 1988. Biological Bases for Mixed-Species Fisheries: Species Co-Distribution in Relation to Environmental and Biotic Variables. Can. J. Fish. Aquat. Sci. 45: 1720-1735.

Neter, J., W. Wasserman, and M.H. Kutner. 1989. Applied Linear Regression Models. Irwin, Burr Ridge, Ill.

Nye, J.A., J.S. Link, J.A. Hare, and W.J. Overholtz. 2009. Changing Spatial Distribution of Fish Stocks in Relation to Climate and Population Size on the Northeast United States Continental Shelf. MEPS 393: 111-129.

Table 1. Analysis of Deviance (ANODEV) table where effects of year (yr), month (mo) and depth (dep) were tested on the logarithm of the catch per unit effort (CPUE). Catch was measured in weight (pounds), and effort was in tow duration (hours). Data from 2001-2009 were used. Year and month were categorical variables whereas depth (in fathoms) was continuous. Selected model was A5. 1 in the linear form denotes intercept.

Model	Linear Form	*D*	*df*	ΔD	Δdf	p-value	AIC
A1	Null	14518.0	4491	NA	NA	NA	18021.3
A2	1 + yr	12674.5	4484	1843.6	7	0.000	17425.3
A3	1 + yr + mo	11680.9	4473	993.6	11	0.000	17080.6
A4	1 + yr + mo + dep	11570.1	4443	110.8	30	0.065	16958.9
A5	**1 + yr + mo + yr*mo**	**9944.4**	**4429**	**1736.6**	**44**	**0.000**	**16445.6**

Table 2. Analysis of Deviance (ANODEV) table where effects of water temperature (tem) and depth (dep) were tested on the logarithm of the catch per unit effort (CPUE). Catch was measured in weight (pounds), and effort was in tow duration (hours). Data from 2000-2010 were used. Selected model was B6. 1 in the linear form denotes intercept.

Model	Linear Form	*D*	*df*	ΔD	Δdf	p-value	AIC
B1	Null	14929.2	4574	NA	NA	NA	18398.2
B2	1 + tem	14037.7	4380	891.5	194	0.000	17543.3
B3	1 + dep	13766.2	4336	271.5	44	0.000	17331.2
B4	1 + tem + dep + tem*dep	12295.1	4335	1471.1	1	0.000	16842.9
B5	1 + tem + dep + tem*dep + dep^2	11625.0	4334	670.1	1	0.000	16601.7
B6	**1 + tem + dep + tem*dep + dep^2 + tem^2**	**11576.9**	**4333**	**48.0**	**1**	**0.000**	**16585.7**

Table 3. Coefficients (coef) and standard errors (SE) of explanatory variables in model B6. The response variable was the logarithm of catch per unit effort. Catch was measured in weight (pounds), and effort was in tow duration (hours). The units of depth and temperature were fathom and degrees Celcius, respectively. *t* is t-statistic (=(coef-0)/SE) testing null hypothesis H_0, coef = 0, and P-value is the value of testing the null hypothesis (=Pr($|t_{df}|$ > *t*, where df = model's residual degrees of freedom).

Explanatory Variables	Coef	SE	*t*	P-value
Intercept	4.929	0.426	11.6	0.000
tem	-0.291	0.072	-4.0	0.000
dep	-0.015	0.009	-1.6	0.100
dep*temp	0.013	0.001	18.7	0.000
dep^2	-0.001	0.000	-16.2	0.000
tem^2	-0.014	0.003	-4.2	0.000

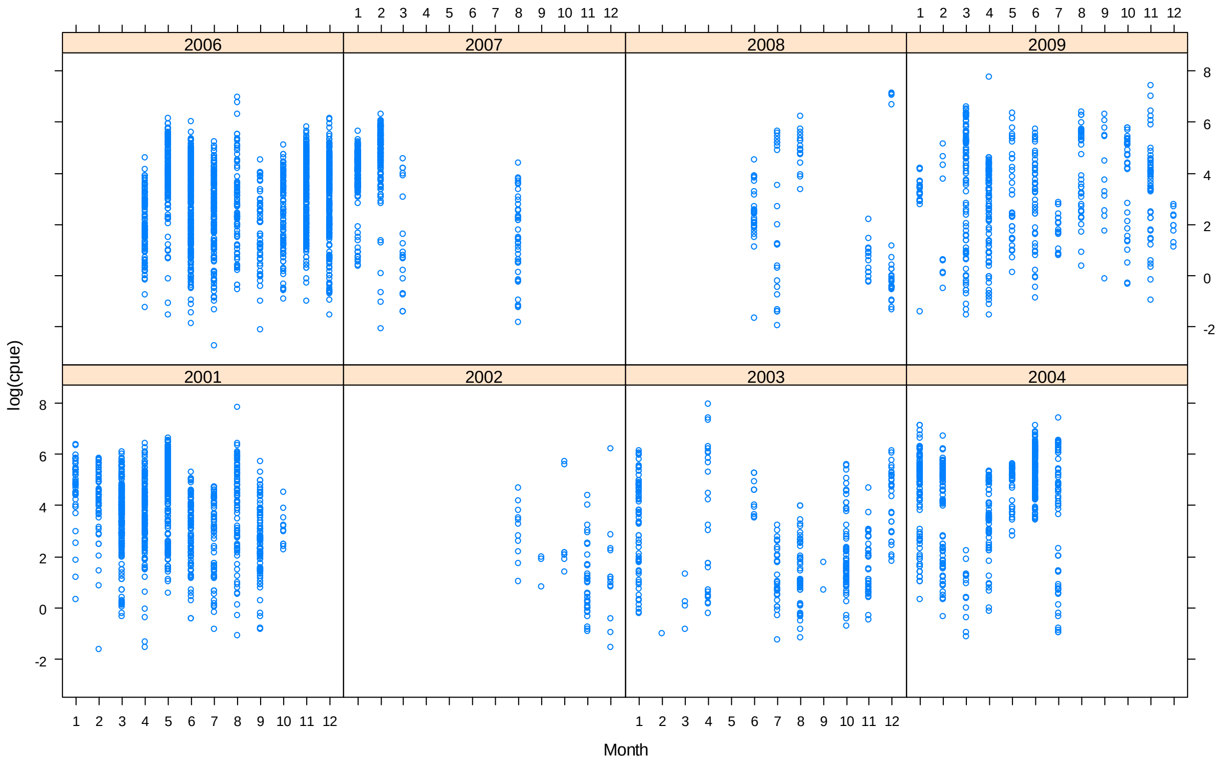


Fig. 1. Logarithm of catch per unit effort (CPUE) against month by year. Catch is in weight (lbs) and effort is in tow duration (hours).

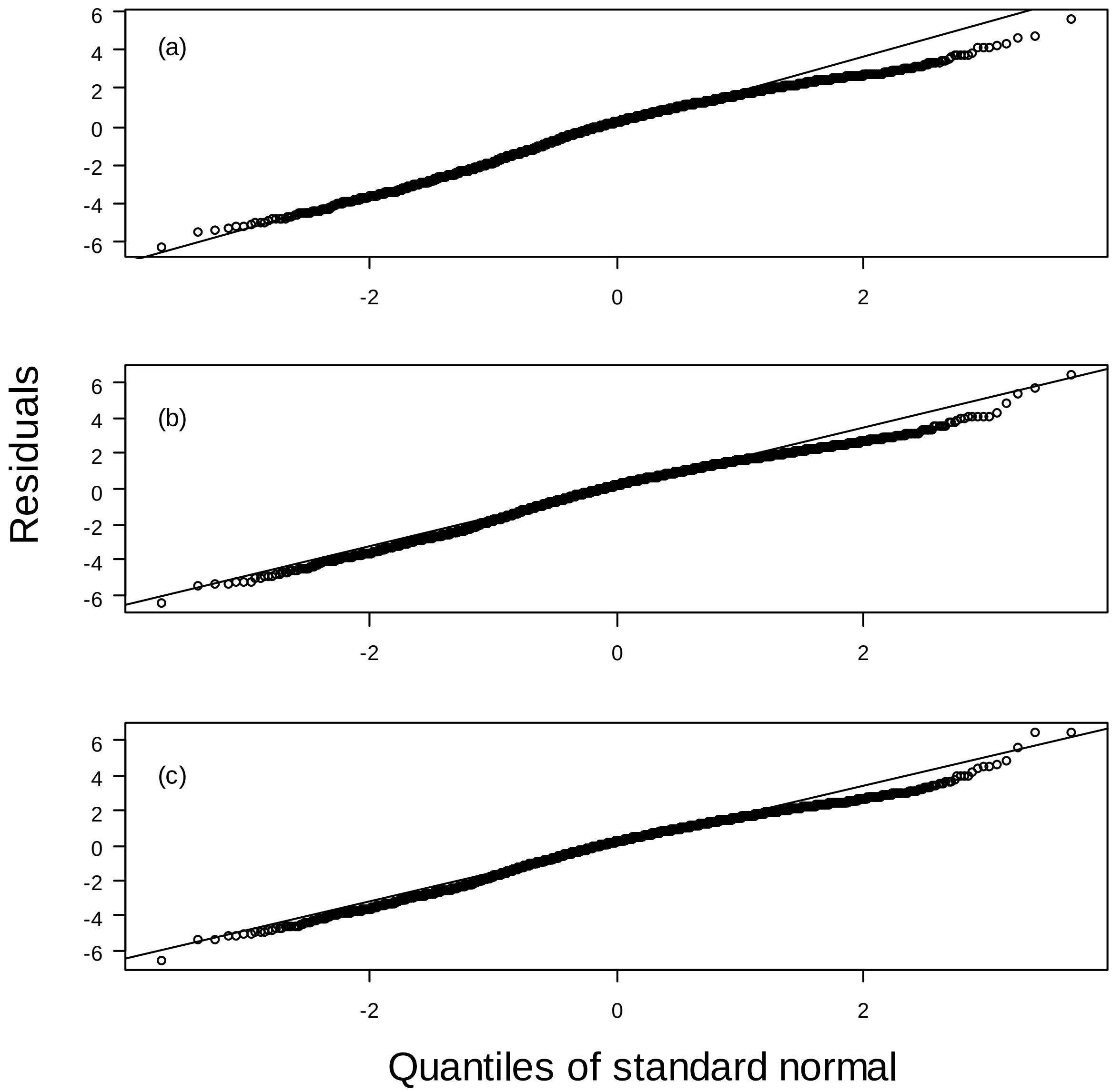
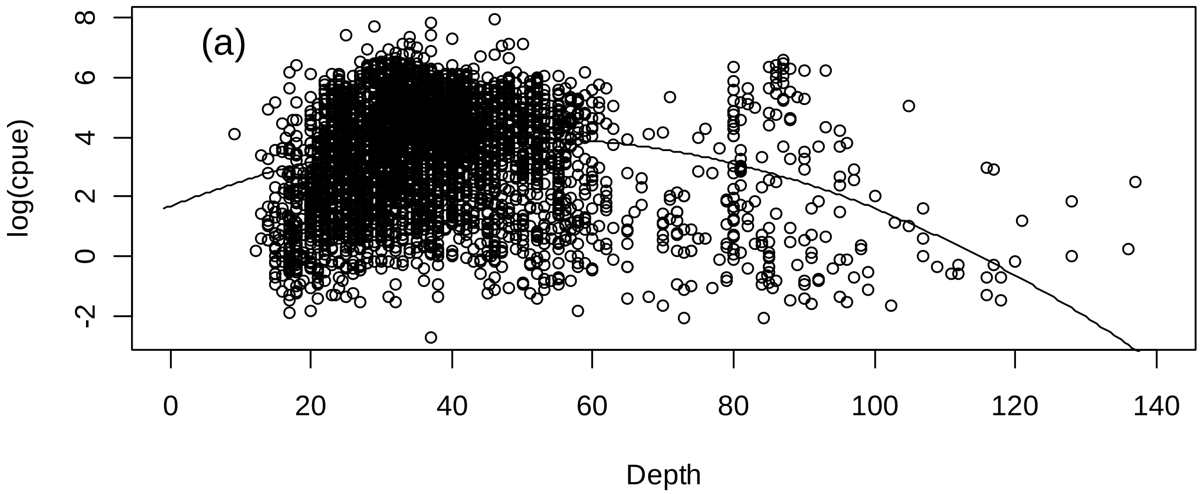


Fig. 2. The Q-Q plot of residuals of competitive models: B4 (*a*), B5 (*b*), and B6 (*c*).



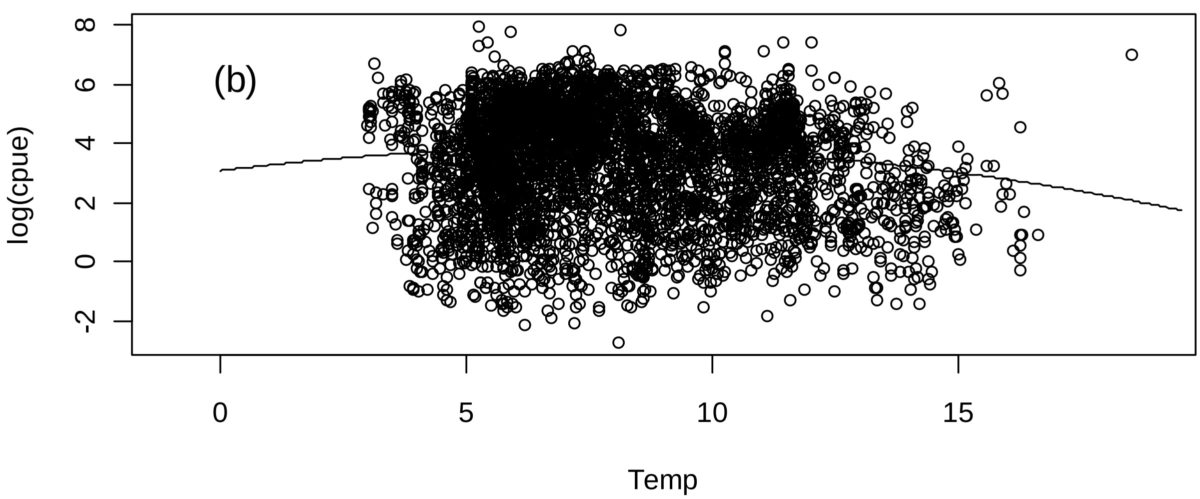


Fig. 3. Effect of the respective quadratic term of depth (fathom) and bottom water temperature (^{o}C) on catch per unit effort (CPUE) (Table 3). Circles are actual data, and the line was the predicted values generated by model B6 with simulated data where other variable but variable of interest (depth or temperature) was constant as the mean of the actual data. The convex shape is due to a negative coefficient of the respective quadratic term of depth and temperature in model B6.

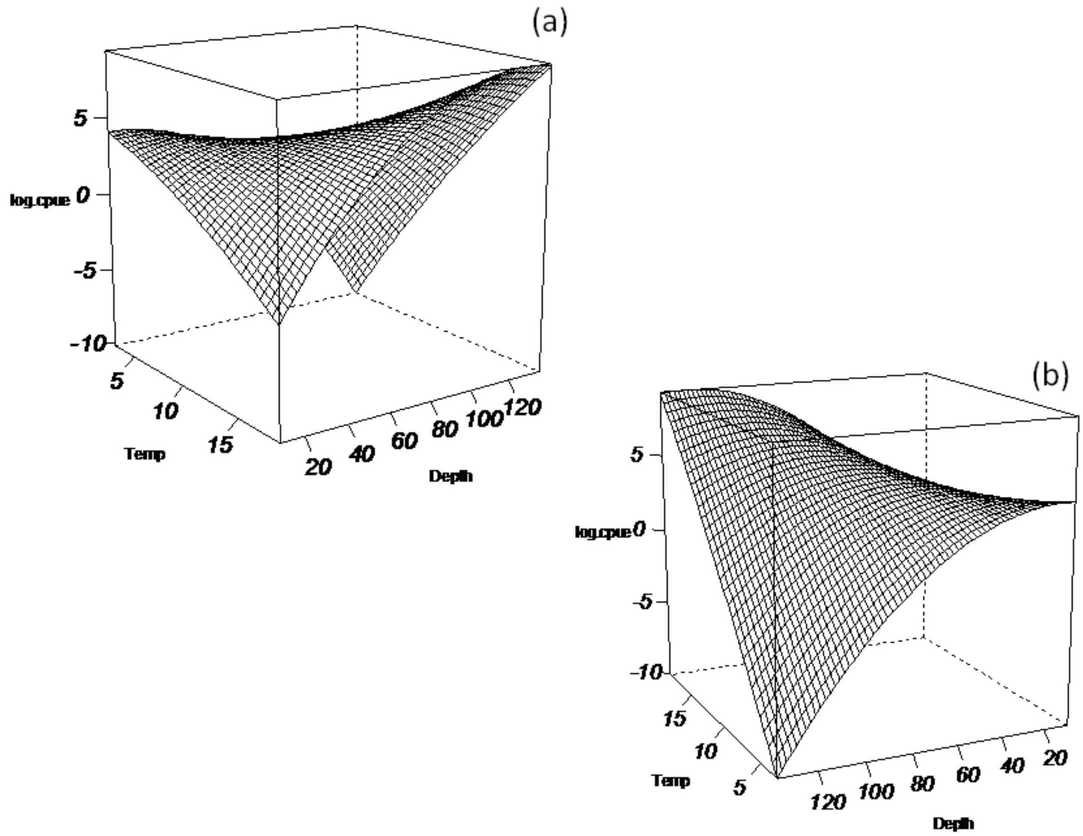


Fig. 4. Illustration of the significant interaction term of bottom water temperature (oC) and depth (fathom). Two panels a and b show the same relationship with catch per unit effort with being viewed in a different angle. The surface over temperature and depth was predicted by final model B6 with simulated data on temperature and depth.

Appendix Table 1. Coefficients (coef) and standard errors (SE) of year and month in model A5 (Table 1) whose linear form = 1 + Year + Month + Year*Month, where 1 denotes intercept term. Year and month are categorical variables: 8 years of 2001-2004 and 2006-2009, and 12 months of Jan through December. *t* is t-statistic (=(coef-0)/SE) testing null hypothesis H_0, coef = 0, and P-value is the value of testing the null hypothesis (=Pr($|t_{df}|$ > *t*, where df = model's residual degrees of freedom).

	Coef	SE	*t*	P-value
Intercept	4.670	0.229	20.44	0.000
yr2002	-2.292	0.745	-3.08	0.002
yr2003	-1.266	0.287	-4.41	0.000
yr2004	-0.124	0.274	-0.45	0.651
yr2006	-0.362	0.625	-0.58	0.562
yr2007	-0.646	0.249	-2.59	0.010
yr2008	-2.875	0.678	-4.24	0.000
yr2009	-1.572	0.364	-4.32	0.000
mo2	-0.427	0.304	-1.41	0.160
mo3	-1.042	0.252	-4.14	0.000
mo4	-0.908	0.259	-3.51	0.000
mo5	0.083	0.247	0.33	0.738
mo6	-1.655	0.274	-6.04	0.000
mo7	-2.037	0.300	-6.78	0.000
mo8	-0.710	0.280	-2.54	0.011
mo9	-2.049	0.279	-7.34	0.000
mo10	-1.582	0.506	-3.12	0.002
mo11	0.412	0.356	1.16	0.248
mo12	-1.073	0.574	-1.87	0.062
yr2003:mo2	-3.951	1.539	-2.57	0.010
yr2004:mo2	-0.252	0.379	-0.67	0.506
yr2007:mo2	1.092	0.344	3.17	0.002
yr2009:mo2	-0.582	0.650	-0.90	0.370
yr2003:mo3	-2.152	0.809	-2.66	0.008
yr2004:mo3	-2.862	0.509	-5.62	0.000
yr2007:mo3	-1.556	0.445	-3.50	0.000
yr2009:mo3	1.479	0.413	3.59	0.000
yr2003:mo4	0.984	0.433	2.27	0.023
yr2004:mo4	-0.164	0.359	-0.46	0.647
yr2006:mo4	-1.427	0.651	-2.19	0.028
yr2009:mo4	0.577	0.409	1.41	0.159
yr2004:mo5	0.186	0.369	0.50	0.614
yr2006:mo5	-0.379	0.643	-0.59	0.556

	Coef	SE	*t*	P-value
yr2009:mo5	0.004	0.453	0.01	0.994
yr2003:mo6	2.582	0.574	4.49	0.000
yr2004:mo6	2.542	0.333	7.62	0.000
yr2006:mo6	0.181	0.649	0.28	0.780
yr2008:mo6	2.247	0.754	2.98	0.003
yr2009:mo6	1.528	0.452	3.38	0.001
yr2003:mo7	-0.104	0.448	-0.23	0.817
yr2004:mo7	1.222	0.387	3.16	0.002
yr2006:mo7	0.226	0.665	0.34	0.734
yr2008:mo7	2.160	0.769	2.81	0.005
yr2009:mo7	0.802	0.557	1.44	0.151
yr2002:mo8	1.356	0.886	1.53	0.126
yr2003:mo8	-1.310	0.394	-3.32	0.001
yr2006:mo8	-0.759	0.666	-1.14	0.254
yr2007:mo8	-1.815	0.376	-4.82	0.000
yr2008:mo8	3.840	0.791	4.85	0.000
yr2009:mo8	1.698	0.465	3.65	0.000
yr2002:mo9	1.241	1.153	1.08	0.282
yr2003:mo9	-0.112	1.110	-0.10	0.919
yr2006:mo9	-0.473	0.678	-0.70	0.486
yr2009:mo9	2.801	0.575	4.87	0.000
yr2002:mo10	2.081	1.019	2.04	0.041
yr2003:mo10	0.332	0.569	0.58	0.560
yr2006:mo10	-0.165	0.783	-0.21	0.834
yr2009:mo10	2.295	0.636	3.61	0.000
yr2002:mo11	-1.667	0.731	-2.28	0.023
yr2003:mo11	-2.161	0.469	-4.61	0.000
yr2006:mo11	-1.395	0.560	-2.49	0.013
yr2008:mo11	-1.561	0.731	-2.14	0.033
yr2003:mo12	1.900	0.654	2.90	0.004