NOT TO BE CITED WITHOUT PRIOR
REFERENCE TO THE AUTHOR(S)

Northwest Atlantic Fisheries Organization

Serial No. N6727 **NAFO SCR Doc. 17/058**

NAFO/ICES WG *PANDALUS* ASSESSMENT GROUP—September 2017

A note on the relationship between the survey abundance of 2-years old West Greenland Shrimp and the biomass two to four years later

by

Frank Rigét

and

AnnDorte Burmeister

Pinngortitaleriffik, Greenland Institute of Natural Resources
Box 570, 3900 Nuuk, Greenland

Summary

The correlation structure of the relationship between West Greenland offshore survey index of age-2 shrimps and the fishable biomass two to four years later was investigated using three different methods. Significant autocorrelation was found in the relationship between age-2 index and the fishable biomass index with a lag of two years, but not in case of three and four years ahead. The relationship between age-2 index and two years ahead was still significant after correction for autocorrelation.

Introduction

The West Greenland offshore survey index of age-2 shrimps has been applied as an index of recruitment two, three and four years later (Burmeister and Rigét 2017), which appears reasonable as the main contribution to the fishable biomass comes from sizes corresponding to ages 4 to 6. As both the recruitment index and the fishable biomass often are large or small for several years in a row, which mean that values in a given year are related to values in previos years and therefore maybe considered to be autocorrelated. If autocorrelation exists, it should be taken into account when evaluating the relationship between recruitment index and later fishable biomass.

Material and Methods

Input data are the West Greenland offshore survey fishable biomass index and the index of age-2 shrimps

Three different procedures were applied in order to correct for temporal autocorrelation. The Cochrane-Orcutt procedure (Cochrane and Orcutt, 1949), generalized least squares (gls in R library nlme, R Core Team (2016)) and the PROC AUTOREG procedure in SAS.

In the Cochrane-Orcutt procedure the regression estimates are found by iteration until they converge. The following steps are iterated 1) do an ordinary regression 2) analyse the time series structure of the residuals 3) estimate the autoregressive coefficients 4) adjust the original variables 5) do an ordinary regression of the adjusted variables. The asta package (Stoffer, 2016) and the orcutt package (Spada et al 2017) in R were used for this procedure.

In the generalized least squares (gls) in R and PROC AUTOREG in SAS an autoregressive error structure of the regression model is specified e.g. AR 1. For both methods, we used maximum likelihood estimation.

Results and Discussion

Results from ordinary regressions of fishable biomass two, three and four years ahead on the abundance index of age-2 shrimps were in all cases highly significant (Table 1).

The autocorrelation of the residuals from the three regressions were examined (Fig. 1). Autocorrelation of the residuals was only significant in case of the lag 1 (AR 1) coefficient of the residuals deriving from the regression of age-2 shrimp index and the fishable biomass two years ahead (FB2). In this case the autocorrelation was estimated to 0.647 (SE=0.158). In no other occasion, any pronounced autocorrelation was found (Fig. 1). However, for illustrative reasons AR 1 corrected regressions of fishable biomass three and four years ahead on age-2 shrimp index were also performed, although no autocorrelation of residuals were found in these ordinary regressions.

In case of age-2 shrimp and fishable biomasse 2 years ahead only minor differences were found in the results of the GLS and PROC AUTOREG procedures (Table 2, Appendix 1). The estimation by the C-O procedure resulted in a higher intercept and a slightly lower slope (Fig. 2). In all case's the corrected regressions showed a significant relationship between age-2 shrimp and fishable biomass 2 years ahead.

For the age-2 shrimp and fishable biomass 3 years ahead the slope of the corrected regression were much lower and no more significant compared to the ordinary regression (Fig. 2, Table 2). Note that the autocorrelation of residuals was not significant in this case.

For age-2 shrimp and fishable biomass 4 years ahead the autocorrected regressions resulted in lower slopes but were still significant in case of the GLS and PROC AUTOREG procedures compared to the ordinary regressions (Fig. 2, Table 2). Note again that the autocorrelation of residuals was not significant in this case and a correction should not be done.

Model comparisons between the GLS models and PROC AUTOREG models, and that of the ordinary regressions expressed by the difference in the Akaike information index (ΔAICc) corrected for small sample sizes. The Cochrane-Orcutt method could not be included in these comparisons because the method drops the first observation.

Fishable Biomasse	2 years ahead	3 years ahead	4 years ahead
	ΔAICc	ΔAICc	ΔAICc
GLS	-15.8	+0.3	+1.7
PROC AUTOREG	-18.9	-2.7	-1.4

In general, only minor differences were found between the GLS and the PROC AUTOREG procedures. Only in case of the fishable biomass 2 years ahead, the correction of autocorrelation of the residuals a considerably model improvement was obtained (the only case with significant autocorrelation of residuals).

In conclusion, we recommend to analyse the autocorrelation of residuals from regression analyses and if necessary to correct the regression using the GLS or PROC procedure.

References

BURMEISTER, AD., F.F RIGET. 2017. The West Greenland trawl survey for *Pandalus borealis*, 2017, with reference to earlier results. *NAFO SCR Doc*. 017/051, Ser. No. N6720. 39 pp.

COCHRANE, D., ORCUTT, G.H. 1949. Application of least squares regression to relationships containing auto-correlated error terms. Journal of the American Statistical Association 44: 32–61.

R CORE TEAM. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

SPADA, S., QUARTAGNO, M., TAMBURINI,, M., ROBINSON, D. 2017. Orcutt (Estimate Procedure in Case of First Order Autocorrelation) package in R.

STOFFER, D. 2016. Asta (Applied Statistical Time Series Analysis) package in R.

Table 1. Results of ordinary regressions (OR) of West Greenland survey abundance index of age-2 shrimps on fishable biomass 2, 3 and 4 years later, together with the lag-1 autocorrelation coefficients (AR 1) of the residuals. P-value is the result of Durbin-Watson test. SE given in brackets.

	$F\_B_{year\ i+2}$	$F\_B_{year\ i+3}$	$F\_B_{year\ i+4}$
Intercept	172.9 (35.3)	158.9 (29.8)	154.5 (26.5)
Slope	18.7 (4.8)	22.4 (4.1)	23.7 (3.6)
p-value	<0.001	<0.001	<0.001
R^2	0.42	0.60	0.69
AR 1	0.66	0.02	0.10
p-value	<0.001	0.70	0.41

Table 2. Results of regression between West Greenland survey abundance index of age-2 shrimps and fishable biomass 2, 3 and 4 years ahead by different approaches of correcting for autocorrelation of residuals. C-O = Cochrane-Orcutt procedure, GLS = general least square, SAS = autoreg procedure in SAS. SE given in brackets. resSE = residual SE.

	C-O	**GLS**	**PROC AUTOREG**
Age-2 shrimps and fishable biomass 2 years ahead			
Intercept	235.0 (35.3)	206.8 (54.3)	206.6 (56.4)
Slope	10.2 (4.8)	10.6 (3.7)	10.6 (3.9)
p-value	0.013	0.010	0.013
resSE, df	58.6, 20	89.2, 21[1]	59.7, 20
Age-2 shrimps and fishable biomass 3 years ahead			
Intercept	291.7 (29.8)	260.7 (58.0)	260.1 (59.5)
Slope	3.4 (4.1)	3.9 (4.8)	4.0 (5.0)
p-value	0.492	0.430	0.438
resSE, df	69.5, 19	98.7, 20	71.1 (19)
Age-2 shrimps and fishable biomass 4 years ahead			
Intercept	238.2 (26.5)	216.6 (43.1)	216.5 (46.7)
Slope	12.4 (3.6)	12.4 (4.4)	12.5 (5.2)
p-value	0.010	0.011	0.027
resSE, df	61.3, 18	77.2, 19	64.9 (18)

[1] Note that the GLS procedure reports higher df than PROC AUTOREG

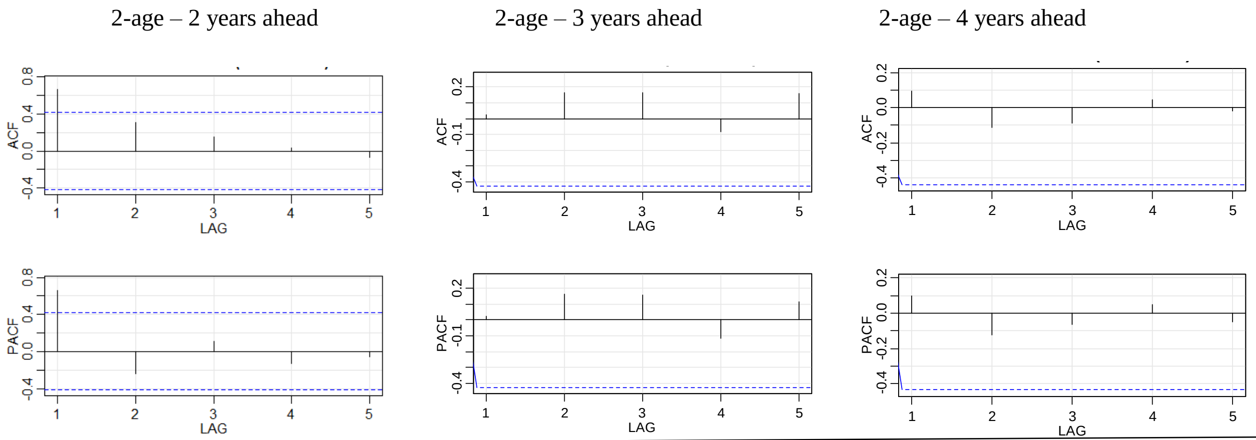

Fig. 1. Plot of autocorrelation (ACF) and partial autocorrelation (PACF) of the residuals from the three ordinary regressions.

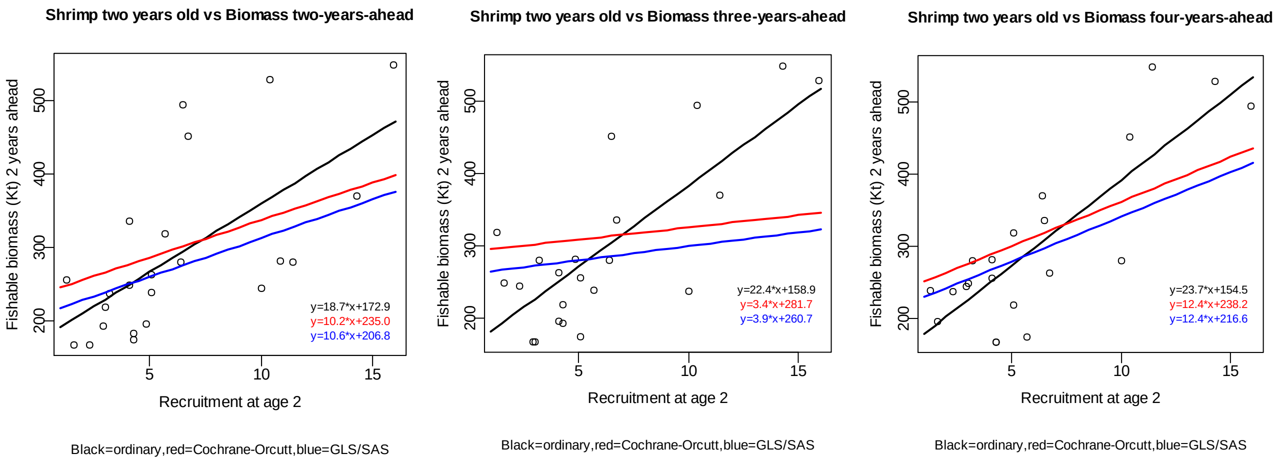

Fig. 2. *Pandalus borealis* in West Greenland: lagged fishable biomass vs. survey estimates of numbers at age 2 from 1993 to 2016. Ordinary regressions (black) and corrected for autocorrelated residuals.

Appendix 1. Fishable biomass two years ahead and recruitment index.
Detailed results of the Cochrane-Orcutt, the GLS and the PROC AUTOREG procedures to account for temporal autocorrelation.

Cochrane-Orcutt

```
lm(formula = fb2 ~ age2)

            Estimate Std. Error t value  Pr(>|t|)
(Intercept)  235.044     35.254   4.904 7.514e-05 ***
age2          10.216      4.769   3.916 0.0007939 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 58.6013 on 20 degrees of freedom
Multiple R-squared:  0.2705 ,  Adjusted R-squared:  0.2341
F-statistic: 7.4 on 1 and 20 DF,  p-value: < 1.309e-02

Durbin-Watson statistic
(original):    0.59120 , p-value: 1.616e-05
(transformed): 1.53676 , p-value: 1.332e-01
```

GLS

```
Generalized least squares fit by maximum likelihood
  Model: fb2 ~ age2
  Data: NULL
       AIC      BIC    logLik
  259.1003 263.6423 -125.5502

Correlation Structure: AR(1)
 Formula: ~1
 Parameter estimate(s):
      Phi
0.7811949

Coefficients:
                Value Std.Error  t-value p-value
(Intercept) 206.78549  54.29212 3.808757  0.0010
age2         10.57099   3.70978 2.849491  0.0096

 Correlation:
     (Intr)
age2 -0.45

Standardized residuals:
        Min          Q1         Med          Q3         Max
-0.88538981 -0.57926567 -0.03715444  0.48358419  2.45294673

Residual standard error: 89.15932
Degrees of freedom: 23 total; 21 residual
```

PROC AUTOREG

Estimates of Autoregressive Parameters

Lag	Coefficient	Standard Error	t Value
1	-0.634562	0.177309	-3.58

Maximum Likelihood Estimates

SSE	71054.4017	DFE	19
MSE	3740	Root MSE	61.15313
SBC	250.42691	AIC	247.153782
MAE	46.7985807	AICC	248.487116
MAPE	17.0467175	HQC	247.924833
Log Likelihood	-120.57689	Transformed Regression R-Square	0.2432
Durbin-Watson	1.4431	Total R-Square	0.7555
		Observations	22

Parameter Estimates

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	206.4569	58.6691	3.52	0.0023	
Age_2	1	10.3449	4.2652	2.43	0.0254	Age_2
AR1	1	-0.7848	0.1395	-5.62	<.0001	