Impacts of spatial scales of fisheries and environmental data on catch per unit effort standardisation

Siquan Tian\textsuperscript{A,B,C}, Yong Chen\textsuperscript{A,C}, Xinjun Chen\textsuperscript{A,B,D}, Liuxiong Xu\textsuperscript{A,B} and Xiaojie Dai\textsuperscript{A,B}

\textsuperscript{A}Key Laboratory of Sustainable Exploitation of Oceanic Fisheries Resources, Ministry of Education, China, and Key Laboratory of Shanghai Education Commission for Oceanic Fisheries Resources Exploitation, Shanghai Ocean University, Shanghai 201306, China.
\textsuperscript{B}College of Marine Sciences, Shanghai Ocean University, 999 Huchenghuan Avenue, Lingang New City, Shanghai 201306, China.
\textsuperscript{C}School of Marine Sciences, University of Maine, Orono, ME 04469, USA.
\textsuperscript{D}Corresponding author. Email: xjchen@shou.edu.cn

Abstract. Spatial scale is an important factor that needs to be considered in data collection and analysis in ecological studies. Studies focusing on the quantitative evaluation of impacts of spatial scales are, however, limited in fisheries. Using the Chinese squid-jigging fishery in the north-western Pacific Ocean as an example, we evaluated impacts of spatial scale used in grouping fisheries and environmental data on the standardisation of fisheries catch per unit effort (CPUE). We developed 18 scenarios of different spatial scales with a combination of three latitudinal levels (0.5°, 1° and 2°) and six longitudinal levels (0.5°, 1°, 2°, 3°, 4° and 5°) to aggregate the data. We then applied generalised additive models to analyse the 18 scenarios of data for the CPUE standardisation, and quantified differences among the scenarios. This study shows that longitudinal and latitudinal spatial scale and size of the spatial area for data aggregation can greatly influence the standardisation of CPUE. We recommend that similar studies be undertaken whenever possible to evaluate the roles of spatial scales and to identify the optimal spatial scale for data aggregations in the standardisation of CPUE and fisheries stock assessment.

Additional keywords: CPUE standardisation, environmental variables, generalised additive models, north-western Pacific Ocean, Ommastrephes bartramii.

Introduction

Information essential to fisheries stock assessment and management is usually collected in fishery-independent scientific surveys and fishery-dependent monitoring programs (Fox and Starr 1996; Booth 2000; OSB 2000). In fishery-independent surveys designed based on statistical principles, data are usually collected with a standard gear under a standard protocol over a long time period on a defined spatial and/or temporal scale (Stelzenmüller \textit{et al.} 2005). Such a survey can provide unbiased estimates of fish population parameters (Hilborn and Walters 1992). However, the quality of these estimates often depends on the spatial and/or temporal scales of the surveys. The finer the spatial and/or temporal scales are, the more precise the estimates tend to be. The choice of spatial and/or temporal scales is, however, often dictated by the availability of funds for the survey.

Fishery-dependent monitoring programs often include data collected from fishers’ logbooks and port and sea sampling of commercial fishing vessels (OSB 2000; Scheirer \textit{et al.} 2004). Because fishing activity is often a non-random process, estimates may be biased in representing the targeted fish population (Bordalo-Machado 2006). The spatial and temporal scales of data collected in fisheries monitoring programs are often much finer than those in fishery-independent scientific surveys, in particular for fisheries with vessel monitoring systems. However, for confidentiality, fisheries data are often required to be aggregated in use so that the exact fishing location for a fisher is not disclosed.

Catch per unit effort (CPUE), whether derived from a fishery-independent or fishery-dependent sampling program, is a key statistic for fisheries stock. It is commonly used as an index of fish stock abundance, implying that a proportional change in CPUE is expected to represent the same proportional change in stock size (FAO 1999). The data collected from fisheries are usually grouped at a defined spatial and/or temporal scale, and CPUE is commonly calculated by dividing the total catch by the corresponding fishing effort over a specific spatial scale and time (Hilborn and Walters 1992). The distribution of fishing effort with respect to targeted fish is usually assumed random within the defined spatial scale. Thus, the spatial scale at which fisheries data are collected directly determines whether the assumption regarding the distribution of fish and fishing effort is violated. Clearly the resolution of spatial scale influences the estimation of CPUE.

Although the ecological implications of spatial scales are well known (Nally and Quinn 1998), few peer-reviewed publications...
are available on the quantitative evaluation of impacts of spatial scales on data analysis in fisheries. Spatial scales for data collection in many fishery-independent surveys are usually determined by financial resources available to the studies, and the impacts of spatial scales on data analyses are often not evaluated. Fishery-dependent data analyses are usually based on whatever spatial scales are available through data collection. Data may also be aggregated so as not to disclose the exact commercial fishing locations. The effect of spatial scale on data analyses is rarely considered. The spatial scales commonly used in data grouping include 10-min squares (in some coastal fisheries, e.g. Chinese coastal trawl fisheries; Chen et al. 2004), 0.5° longitude × 0.5° latitude (e.g. Chinese squid-jigging fisheries statistical rectangles; Wang and Chen 2005), 1° longitude × 0.5° latitude (e.g. International Council for the Exploration of the Sea statistical rectangles; http://www.ices.dk/indexfla.asp, accessed 18 March 2008), and 1° longitude × 1° latitude (e.g. the statistical rectangles of the South Pacific Commission for pole-and-line and purse seine fisheries; http://www.spc.int/corp/, accessed 18 March 2008). The spatial scale of 5° longitude × 5° latitude used in most longline tuna fisheries (e.g. the statistical rectangles of the Inter-American Tropical Tuna Commission for longline fishery; http://www.iattc.org/HomeENG.htm, accessed 18 March 2008) is perhaps the coarsest scale used in grouping commercial fisheries data.

Spatial scale is considered critical in the application of geostatistics to address fisheries problems. Geostatistical analysis is commonly used in fisheries to optimise sampling strategies for studying fish distributions (Petitgas 1996) and to estimate fish biomass (Warren 1997; Maynou 1998; Fernandes and Rivoirard 1999; Grabowski et al. 2005). Stelzenmüller et al. (2005) employed geostatistics to investigate the spatial structure of several fish species on different spatial scales in the northern North Sea and showed that the spatial scale of the survey is important for estimating the mean biomass. However, limited consideration is given to comparing the impacts of different spatial scales on data analysis in fisheries. The only study we could find was Campbell et al. (1996), which compared two different spatial scales (1° × 1° and 5° × 5°) of catch and effort data when analysing indices of abundance for southern bluefin tuna. In the present study, we grouped data using different spatial scales, and evaluated and quantified the impacts of different spatial scales on the standardisation of CPUE.

Standardisation of CPUE is common in fisheries stock assessment. Many factors may influence CPUE, such as spatial dispersion of resources (Lange 1991), fishing strategy (He et al. 1997), and abiotic and biotic environmental variables (Maundry and Punt 2004). Thus, it is necessary to standardise fisheries CPUE to remove the influence of factors other than stock abundance before the CPUE can be used as an index of stock abundance (Hinton and Maundry 2004).

The Chinese squid-jigging fishery in the north-western Pacific Ocean, which targets the neon flying squid, Ommastrephes bartramii, was used as an example in the present study to illustrate the impacts of spatial scale on the standardisation of CPUE. We chose this fishery because the neon flying squid has a short life span and its population dynamics are influenced by environmental factors on a fine spatial scale (Yatsu et al. 1997, 1998). This squid plays an important role in the pelagic ecosystem and supports a valuable fishery (Wang and Chen 2005). Chinese fishing vessels have targeted this squid stock since 1993, and yielded annual landings of 80 to 120 thousand tonnes. We developed a statistical model to standardise the CPUE data grouped based on 18 different combinations of spatial scales, and then compared their standardisation results to evaluate how the choice of different spatial scales might influence the standardisation of CPUE.

Materials and methods

Fisheries data

Four stocks of neon flying squid are defined in the North Pacific Ocean: west stock of the winter–spring cohort, central-east stock of the winter–spring cohort, central stock of the fall cohort and east stock of the fall cohort (Yatsu et al. 1998). The west stock of the winter–spring cohort of neon flying squid is the main target of the Chinese squid-jigging vessels (Wang and Chen 2005) and is mainly distributed in the area west of 170°E in the north-western Pacific Ocean (Bower and Ichii 2005). The main fishing period of this squid fishery is from June to November. Commercial fisheries data of the west stock of winter–spring cohort in the waters west of 170°E from June to November during 1995 to 2004 were acquired from the Chinese Squid-jigging Technology Group. The data include fishing dates, fishing locations (longitude and latitude), monthly number of fishing vessels and monthly catch per fishing vessel.

Environmental data

Previous studies (e.g. Yatsu and Watanabe 1996; Chen and Chiu 1999; Fan 2004; Chen and Tian 2005) showed that the CPUE of neon flying squid was related to marine environmental variables such as salinity, sea level height and sea surface temperature (SST). The neon flying squid exhibits vertical diel movement. They inhabit water depths of 0–40 m at night and of 150–350 m during the day (Murata and Nakamura 1998; Tanaka 2001).

The corresponding environmental data were acquired from the website http://iridl.ldeo.columbia.edu/SOURCES/.CARTON-GIESE/ (accessed 10 September 2007). From the database, we selected the temperature at depths of 35 m and 317 m, which reflect favourable depths for the squid during night and day, respectively (Wang and Chen 2005). SST and water temperatures at depths of 35 m and 317 m, sea surface salinity (SSS) and sea level height (SLH) were selected as environmental variables in this study. Clearly the list may not include all the environmental variables that can influence the squid CPUE. However, inclusion of these variables that were deemed important to squid abundance and distribution in previous studies is sufficient, as the present study focuses on the impacts of spatial scale on the standardisation of CPUE.

Statistical methods for CPUE standardisation

Statistical modelling is commonly used in quantifying the relationships between the performance of fisheries (e.g. CPUE) and variables that may influence it (e.g. environmental factors and spatio-temporal factors). The developed statistical models are then used to remove the impacts of the variables to derive standardised CPUE (Hilborn and Walters 1992; Punt et al. 2000). Of the statistical methods used, generalised linear models (GLM)
and generalised additive models (GAM) are probably most commonly used in the standardisation of CPUE of commercial fisheries (e.g. Bigelow et al. 1999; Punt et al. 2000; Campbell 2004). The GAM is an extension of GLM in which a link function is chosen depending on the assumed error structure of the model. This function then transforms the response variable into a differentiable and continuous function (Hastie and Tibshirani 1990).

The GAM is often used to study possible interactions between fisheries variables and geographic and environmental variables for understanding and predicting fish population dynamics, and tends to perform well in dealing with spatially explicit data (Bigelow et al. 1999; Maury et al. 2001; Denis et al. 2002). Relationships between fisheries and environmental variables are typically non-linear (Brander 1994; Bigelow et al. 1999). The GAM can deal with non-linear relationships well because it is not tied to a particular linear functional relationship for its link function and because it is less restrictive in assumptions about the underlying statistical distribution of the data (Maravelias and Reid 1997). Polynomials, step-functions, splines and smoothers are some of the commonly used functions for assessing the underlying relationship of independent predictors. Compared with GLM, GAM can describe more robustly the non-linear and/or skewed ecological responses that are typical of ecological processes (Yee and Mitchell 1991), and is a powerful tool for the standardisation of CPUE (Maunder and Punt 2004).

For the above reasons, GAM was used to standardise CPUE of the neon flying squid fishery in the present study. The GAM used in the present study is specified as (Bellido et al. 2001):

$$Y = \alpha + \sum_{j=1}^{p} f_j(x_j) + \varepsilon$$  \hspace{1cm} (1)

where $\alpha$ represents the intercept term in the fitted model, $f_j$ is a smooth function (a spline or loess smoother), $x_j$ are the independent variables and $\varepsilon$ is an error term with $\varepsilon \sim N(0, \sigma^2)$ and $E(\varepsilon) = 0$.

The natural log-transformed CPUE was assumed to have normally distributed errors in the GAM modelling. This assumption is common in studies involving the standardisation of CPUE (Maunder and Punt 2004). Thus, the GAM for the standardisation of CPUE in this study can be written as:

$$\ln(\text{CPUE} + c) = \alpha + \text{year} + \text{month} + s(\text{longitude}) + s(\text{latitude}) + s(\text{SST}) + s(T35m) + s(T317m) + s(\text{SLH}) + s(\text{SSS}) + \text{longitude} \times \text{latitude} + \varepsilon$$  \hspace{1cm} (2)

where $\alpha$ is a constant, $s$ is a spline smoother function, and $T35m$ and $T317m$ are water temperatures at depths of 35 m and 317 m, respectively. The constant $c$ was used in log-CPUE because there were '0' values in the CPUE data. To make log-transformation of CPUE valid, a constant is usually added to the catch rate (Punt et al. 2000). Various values have been assumed for the constant. However, Campbell et al. (1996) suggested that setting the constant equal to 10% of the mean overall catch rate used in the analysis could minimise biases resulting from adjusting the catch rate in this manner. We set constant $c$ at 10% of the mean CPUE. In the GAM, year and month were treated as factors, $\text{longitude} \times \text{latitude}$ was the interaction of longitude and latitude, and the remaining covariates were dealt with by spline smoothers.

A pseudo-coefficient of residual determination ($PCf$) was calculated for measuring the goodness of fit for the GAM model (Swartzman et al. 1992):

$$PCf = 1 - \frac{RD}{ND}$$  \hspace{1cm} (3)

where $RD$ is the residual deviance, i.e. the deviance of the parsimonious model, and $ND$ is the null deviance, which is the deviance of the model only having the intercept. S-plus (ver. 6.2) was used to implement GAM analysis.

**Setting spatial scales**

To evaluate the impacts of spatial scale on grouping fisheries data, we set three spatial scale levels for the latitude: 0.5°, 1° and 2°; and six levels for the longitude: 0.5°, 1°, 2°, 3°, 4° and 5°. The combination of these latitudinal and longitudinal scales results in 18 scenarios in the GAM (Table 1). The level of scales along the latitudinal direction was fewer than that along the longitudinal direction for two reasons: (1) the distribution of fishing grounds of Chinese squid-jigging fishery was 30° longitudinally (140°–170°E) but only 8° latitudinally (37°–45°N; Wang and Chen 2005); and (2) the latitudinal gradients of marine environmental variables were larger than the longitudinal gradients on the squid fishing grounds in the north-western Pacific Ocean. This difference is caused by the northward Kuroshio Current and southward Oyashio Current (Chen and Tian 2005).

The mean nominal CPUE (t per vessel) for each scenario was calculated as:

$$\text{CPUE} = \frac{\sum \text{catch}}{\sum \text{fishing days}}$$  \hspace{1cm} (4)

where $\sum \text{catch}$ is sum of monthly catch for all the fishing vessels within the grid of a specific scenario, and $\sum \text{fishing days}$ is the sum of monthly fishing days for all fishing vessels in the grid defined for each scenario. We chose month as the time step in grouping CPUE for each scenario.

The environmental variables were reported on the spatial scale of 0.5° × 0.5° (i.e. the first scenario in Table 1). For all other scenarios with coarser spatial scales (Table 1), the mean values of all 0.5° × 0.5° grids within the defined areas were calculated and used in the corresponding CPUE standardisation. For example, scenario VIII of scale 1° × 1° includes four original grids of 0.5° × 0.5°. The value of each environmental variable was calculated as the mean of the values of these four original grids. The environmental data and CPUE were grouped by the corresponding spatial scale defined in each scenario (Table 1).

**Evaluating spatial scales of data**

We plotted the spatial distribution of CPUE and environmental variables (using the Kriging interpolation method) to illustrate the differences when different spatial scales were used in grouping fisheries and environmental data. The large amount of data for different months and years and 18 scenarios of different spatial scales make it impractical to plot all the data in this paper.
Table 1. The goodness of fit (pcf) and attained level of significance (P) for environmental variables in the generalised additive model (GAM) analysis for each scenario

We only included the plots for one set of CPUE and environmental variables in the GAM analysis.

The CPUE is a relative abundance index. Thus, its temporal variation is most important, and the impacts of data spatial scales on the temporal variability of CPUE should be evaluated. We calculated the coefficient of variation (CV) of the annual and monthly standardised CPUEs from the GAM to evaluate how the choice of spatial scale of the analysis might influence the precision of the estimated CPUE, and then compared differences in CVs between the scenarios. The CV value was calculated as:

\[ CV = \sqrt{\frac{\sum_{i=1}^{n}(CPUE_{ij} - \bar{CPUE})^2}{n - 1}} / \bar{CPUE} \]  

where \( CPUE_{ij} \) is standardised CPUE and \( \bar{CPUE} \) is the average of standardised CPUE of the \( j \)th year or month for each scenario, and \( n \) is the number of data points in the calculation (\( n \) varies with scenarios).

Because Scenario I has the finest spatial scale (0.5° × 0.5°), Table 1) and all other scenarios were derived from this scenario, we used the standardised CPUE derived for Scenario I as the base case and compared it with those derived for other scenarios. Differences in the derived CPUE between Scenario I and other scenarios reflect the impacts of spatial scales used in grouping fisheries and environmental data. A mean relative difference index (MRDI) was calculated for each scenario to quantify such an impact:

\[ MRDI_i = \left( \frac{\sum_j |CPUE_{ij} - CPUE_{ij}|}{CPUE_{ij}} \right) \ast 100 / n \]  

where \( MRDI_i \) is the mean relative absolute difference in CPUE for scenario \( i \), \( CPUE_{ij} \) is the standardised CPUE of the \( j \)th year for each scenario. All other scenarios were derived from Scenario I (i.e. 0.5° × 0.5°) and all other scenarios were derived from this scenario, were not significant (\( P > 0.05 \)) and were excluded from the GAM analysis. Longitude and Latitude were included in all the scenarios. Although longitude was not significant, it was also included for the consideration of impacts of spatial scales on the standardisation of CPUE.

We only included the plots for one set of CPUE and environmental data (i.e. September 1998) for two spatial scales: Scenarios I (i.e. 0.5° × 0.5°) and XV (i.e. 2° × 2°).

For a feasible comparison of the standardised CPUE of different scenarios, each of the 18 standardised indices derived from GAM was standardised to 1. This was done using the following equation:

\[ CPUE_{ij} = \frac{CPUE_{ij}}{\text{max} \{ CPUE_{ij} \}} \]  

where \( CPUE_{ij} \) is the standardised CPUE of the \( i \)th scenario in the \( j \)th month or year, \( \text{max} \{ CPUE_{ij} \} \) is the standard CPUE derived from GAM for each scenario and was the maximum GAMCPUE of the \( j \)th month or year for each scenario.

Both annual standardised CPUE and monthly standardised CPUE were estimated for each scenario in the GAM analysis. The values of standardised CPUEs among the scenarios were compared using the relative index (RI) calculated as:

\[ RI_{ij} = \frac{CPUE_{ij} - CPUE_{ij}^{\text{max}}}{CPUE_{ij}^{\text{max}}} \ast 100 \]  

where \( RI_{ij} \) is relative index of the \( i \)th scenario in the \( j \)th month or year. \( CPUE_{ij} \) is the standardised CPUE for each scenario and \( CPUE_{ij}^{\text{max}} \) is the average standardised CPUE of the \( j \)th month or year for each scenario. The RI value was used to compare differences in standardised CPUE estimated for different scenarios.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Intervals of longitude (°)</th>
<th>Intervals of latitude (°)</th>
<th>LON</th>
<th>LAT</th>
<th>SST</th>
<th>Tem-35 m</th>
<th>Tem-317 m</th>
<th>SSS</th>
<th>SLH</th>
<th>pcf</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>0.5</td>
<td>0.5</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.23</td>
</tr>
<tr>
<td>II</td>
<td>1</td>
<td>0.5</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.26</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>2</td>
<td>0.5</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>3</td>
<td>0.5</td>
<td>0.00</td>
<td>0.03</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.03</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>4</td>
<td>0.5</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>VI</td>
<td>5</td>
<td>0.5</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.29</td>
<td></td>
</tr>
<tr>
<td>VII</td>
<td>0.5</td>
<td>1</td>
<td>0.00</td>
<td>0.04</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.04</td>
<td>0.26</td>
<td></td>
</tr>
<tr>
<td>VIII</td>
<td>1</td>
<td>1</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>IX</td>
<td>2</td>
<td>1</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>3</td>
<td>1</td>
<td>0.03</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>XI</td>
<td>4</td>
<td>1</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>XII</td>
<td>5</td>
<td>1</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>XIII</td>
<td>0.5</td>
<td>2</td>
<td>0.00</td>
<td>0.03</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.03</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>XIV</td>
<td>1</td>
<td>2</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>XV</td>
<td>2</td>
<td>2</td>
<td>0.03</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.03</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>XVI</td>
<td>3</td>
<td>2</td>
<td>0.01</td>
<td>0.04</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>XVII</td>
<td>4</td>
<td>2</td>
<td>0.00</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>XVIII</td>
<td>5</td>
<td>2</td>
<td>0.03</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.33</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Spatial scales in CPUE standardisation

Marine and Freshwater Research

1277

Comparing observed CPUE and environmental variables

The spatial distribution of observed CPUE showed obvious differences between the two scenarios of different spatial scales (i.e. Scenarios I and XV were selected for the plots) in September 1998 (Fig. 1). The proportion of areas where CPUE had values larger than 3.0 for Scenario I was higher than that for Scenario XV (Fig. 1). Differences were also observed in the spatial distribution of environmental variables between Scenarios I and XV (Fig. 2). The environmental variables’ isolines for Scenario I were more tortuous than those for Scenario XV. Thus, the spatial scales used to group data could greatly influence the spatial distributions of observed CPUE and environmental variables.

Comparing standardised CPUE

Annual and monthly standardised CPUEs of each scenario derived from GAMs were compared, the variables are included in each GAM and the goodness of fit for these models are summarised in Table 1. The variables found to be insignificant in the GAM are excluded from the models used for the standardisation of CPUE.

The RIs calculated according to Eqn 6 are shown in Table 2 for the annual standardised CPUEs and in Table 3 for the monthly standardised CPUEs. The ranges of annual and monthly standardised CPUEs over all the scenarios are shown in Fig. 3a and Fig. 3b, respectively. For the annual standardised CPUEs, the temporal trends of the standardised CPUEs were similar, but the values varied among different scenarios. The largest difference of RI values occurred in 2004 for which the maximum value was 9.26% and minimum value was −15.86%, and the smallest difference of RI values occurred in 1998 for which the maximum and minimum values were 0.24% and −1.68%, respectively. For the monthly standardised CPUEs, the monthly trends increased from June to September, and were similar among the scenarios. The largest difference of RI values occurred in August with the maximum and minimum values of 8.76% and −12.76%, respectively, and the smallest difference of RI values was found in July with the maximum and minimum values of 3.88% and −8.55%, respectively.

Comparing CVs of different scenarios

We calculated the CVs using Eqn 7 for both the annual and monthly standardised CPUEs for each scenario. The results are shown in Fig. 4a and Fig. 4b. For each scenario, CVs of annual standardised CPUEs ranged from ~0.26 to 0.35 as opposed to CVs for monthly standardised CPUEs, which ranged from 0.15 to 0.25. Identically sized scenarios with different longitude and latitude scales also had different CV. When the latitude scale was fixed, the CV for the annual standardised CPUEs tended to increase with the longitude scales from 0.5° to 2°, but fluctuated as the longitude scale increased past 2° (Fig. 5a). The lowest CV
was found when the longitude scale was set at 0.5°, but the spatial scale of the highest CV varied with different longitude scales. The CVs for the monthly CPUEs tended to increase with longitude scale when the latitude scale was fixed (Fig. 5b). To compare impacts of different longitude scales, the scenarios were divided into six groups by their longitude scales (i.e. 0.5°, 1°, 2°, 3°, 4° and 5°). The CVs for the annual standardised CPUEs tended to increase with the latitude scales and different trends occurred when the longitude scale was 2° or 3° (Fig. 6a). The trends of the CVs for the monthly standardised CPUEs relating to latitude scales were not clear (Fig. 6b). From both Fig. 6a and Fig. 6b, it appeared that the CVs tended to increase with longitude scales.

Comparing CPUEs of Scenario I v. other scenarios

The MRDI values measured the departure of a scenario from Scenario I (i.e. the finest spatial scale) and were plotted for annual (Fig. 7a) and monthly standardised CPUEs (Fig. 7b). MRDI values for annual standardised CPUEs increased quickly with spatial scale, but the increase lessened when the spatial scale surpassed 1.5-degree squares (Fig. 7a). MRDI values for monthly standardised CPUEs showed much variation across the area of spatial grids (Fig. 7b). Scenario XV had the highest MRDI of 8.07% for the annual CPUEs, and most scenarios had MRDI values from 4% to 8%. The largest and smallest MRDI values of monthly CPUEs were 9.02% and 0.83%, respectively. Scenarios with the same-sized square but different longitude and latitude

Fig. 2. The spatial distributions of environmental variables for Scenario I (left panel) and Scenario XV (right panel) defined in Table 1. The thick solid lines are the environmental variables’ isolines derived from the kriging interpolation.
Table 2. The relative index (RI) calculated from Eqn 5 for the annual standardised catch per unit effort (CPUE) for each scenario described in Table 1

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>-1.89</td>
<td>-1.57</td>
<td>-3.52</td>
<td>0.22</td>
<td>4.62</td>
<td>4.11</td>
<td>4.91</td>
<td>3.86</td>
<td>9.14</td>
<td>9.26</td>
</tr>
<tr>
<td>II</td>
<td>-2.68</td>
<td>-0.63</td>
<td>-3.43</td>
<td>0.19</td>
<td>4.60</td>
<td>4.49</td>
<td>5.71</td>
<td>4.31</td>
<td>9.39</td>
<td>8.49</td>
</tr>
<tr>
<td>III</td>
<td>-3.00</td>
<td>-2.46</td>
<td>-3.43</td>
<td>-0.01</td>
<td>2.24</td>
<td>4.55</td>
<td>4.59</td>
<td>3.42</td>
<td>4.09</td>
<td>7.57</td>
</tr>
<tr>
<td>IV</td>
<td>-2.10</td>
<td>-1.93</td>
<td>-4.28</td>
<td>0.23</td>
<td>4.05</td>
<td>2.66</td>
<td>1.94</td>
<td>-0.40</td>
<td>2.93</td>
<td>3.22</td>
</tr>
<tr>
<td>V</td>
<td>2.70</td>
<td>7.71</td>
<td>-0.92</td>
<td>0.22</td>
<td>-0.40</td>
<td>1.74</td>
<td>4.71</td>
<td>6.93</td>
<td>5.16</td>
<td>5.63</td>
</tr>
<tr>
<td>VI</td>
<td>3.03</td>
<td>7.47</td>
<td>-0.46</td>
<td>0.22</td>
<td>0.14</td>
<td>1.99</td>
<td>4.69</td>
<td>7.44</td>
<td>5.14</td>
<td>5.35</td>
</tr>
<tr>
<td>VII</td>
<td>-3.54</td>
<td>-2.88</td>
<td>-5.51</td>
<td>0.23</td>
<td>3.51</td>
<td>2.07</td>
<td>-0.65</td>
<td>-3.18</td>
<td>-0.64</td>
<td>-0.98</td>
</tr>
<tr>
<td>VIII</td>
<td>-1.79</td>
<td>-2.98</td>
<td>-0.27</td>
<td>0.23</td>
<td>0.96</td>
<td>2.44</td>
<td>-0.79</td>
<td>-1.00</td>
<td>-3.16</td>
<td>-1.20</td>
</tr>
<tr>
<td>IX</td>
<td>5.10</td>
<td>8.02</td>
<td>1.78</td>
<td>0.22</td>
<td>-1.36</td>
<td>2.04</td>
<td>2.75</td>
<td>6.82</td>
<td>6.80</td>
<td>4.91</td>
</tr>
<tr>
<td>X</td>
<td>-1.14</td>
<td>-2.95</td>
<td>-2.49</td>
<td>0.24</td>
<td>0.87</td>
<td>2.41</td>
<td>-1.69</td>
<td>-2.73</td>
<td>-6.32</td>
<td>-4.27</td>
</tr>
<tr>
<td>XI</td>
<td>1.69</td>
<td>1.87</td>
<td>4.09</td>
<td>0.05</td>
<td>-2.01</td>
<td>-0.20</td>
<td>-1.54</td>
<td>-3.31</td>
<td>-5.27</td>
<td>-3.51</td>
</tr>
<tr>
<td>XII</td>
<td>-3.12</td>
<td>-8.42</td>
<td>-2.76</td>
<td>0.24</td>
<td>-1.37</td>
<td>-4.08</td>
<td>-5.97</td>
<td>0.29</td>
<td>-1.95</td>
<td>-15.86</td>
</tr>
<tr>
<td>XIII</td>
<td>1.21</td>
<td>-0.69</td>
<td>4.03</td>
<td>0.24</td>
<td>-3.27</td>
<td>-1.71</td>
<td>-4.09</td>
<td>-5.53</td>
<td>-6.39</td>
<td>-3.63</td>
</tr>
<tr>
<td>XV</td>
<td>4.55</td>
<td>-3.66</td>
<td>4.16</td>
<td>-1.68</td>
<td>-3.20</td>
<td>-8.71</td>
<td>-5.03</td>
<td>-4.95</td>
<td>-2.95</td>
<td>-2.98</td>
</tr>
<tr>
<td>XVI</td>
<td>-1.40</td>
<td>1.02</td>
<td>3.63</td>
<td>0.23</td>
<td>-0.26</td>
<td>-0.61</td>
<td>0.51</td>
<td>0.45</td>
<td>-8.63</td>
<td>-5.78</td>
</tr>
<tr>
<td>XVII</td>
<td>-0.78</td>
<td>2.18</td>
<td>4.11</td>
<td>-0.18</td>
<td>-0.60</td>
<td>-0.96</td>
<td>-1.40</td>
<td>-1.92</td>
<td>-7.04</td>
<td>-5.75</td>
</tr>
<tr>
<td>XVIII</td>
<td>-1.30</td>
<td>-0.09</td>
<td>1.62</td>
<td>0.24</td>
<td>-5.92</td>
<td>-6.01</td>
<td>-4.86</td>
<td>-8.58</td>
<td>3.04</td>
<td>1.06</td>
</tr>
</tbody>
</table>

Table 3. The relative index (RI) calculated from Eqn 5 for the monthly standardised catch per unit effort (CPUE) for each scenario described in Table 1

<table>
<thead>
<tr>
<th>Scenario</th>
<th>June</th>
<th>July</th>
<th>August</th>
<th>September</th>
<th>October</th>
<th>November</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>1.60</td>
<td>2.43</td>
<td>2.89</td>
<td>1.23</td>
<td>1.98</td>
<td>0.46</td>
</tr>
<tr>
<td>II</td>
<td>2.34</td>
<td>2.20</td>
<td>2.17</td>
<td>2.40</td>
<td>1.98</td>
<td>-1.15</td>
</tr>
<tr>
<td>III</td>
<td>-0.74</td>
<td>2.20</td>
<td>2.76</td>
<td>4.44</td>
<td>2.00</td>
<td>-5.51</td>
</tr>
<tr>
<td>IV</td>
<td>-0.04</td>
<td>2.42</td>
<td>5.00</td>
<td>4.27</td>
<td>1.98</td>
<td>-3.29</td>
</tr>
<tr>
<td>V</td>
<td>5.86</td>
<td>0.28</td>
<td>-3.98</td>
<td>-3.79</td>
<td>2.01</td>
<td>2.24</td>
</tr>
<tr>
<td>VI</td>
<td>5.75</td>
<td>1.47</td>
<td>-3.88</td>
<td>-2.31</td>
<td>2.00</td>
<td>2.10</td>
</tr>
<tr>
<td>VII</td>
<td>-2.89</td>
<td>0.92</td>
<td>5.62</td>
<td>6.39</td>
<td>1.05</td>
<td>-9.34</td>
</tr>
<tr>
<td>VIII</td>
<td>-5.31</td>
<td>-1.18</td>
<td>1.56</td>
<td>6.49</td>
<td>0.87</td>
<td>-10.74</td>
</tr>
<tr>
<td>IX</td>
<td>3.00</td>
<td>-0.90</td>
<td>-5.83</td>
<td>-3.08</td>
<td>2.03</td>
<td>0.27</td>
</tr>
<tr>
<td>X</td>
<td>-3.62</td>
<td>2.11</td>
<td>6.08</td>
<td>6.36</td>
<td>0.58</td>
<td>-6.78</td>
</tr>
<tr>
<td>XI</td>
<td>-1.37</td>
<td>-3.98</td>
<td>-12.76</td>
<td>-15.01</td>
<td>-8.05</td>
<td>6.87</td>
</tr>
<tr>
<td>XII</td>
<td>-6.65</td>
<td>-8.55</td>
<td>8.76</td>
<td>6.54</td>
<td>-10.40</td>
<td>-2.38</td>
</tr>
<tr>
<td>XIII</td>
<td>-1.26</td>
<td>2.59</td>
<td>4.42</td>
<td>2.11</td>
<td>1.98</td>
<td>2.06</td>
</tr>
<tr>
<td>XIV</td>
<td>-2.93</td>
<td>-1.45</td>
<td>-2.83</td>
<td>-3.08</td>
<td>0.50</td>
<td>6.53</td>
</tr>
<tr>
<td>XV</td>
<td>-4.16</td>
<td>-4.37</td>
<td>-5.24</td>
<td>-4.64</td>
<td>2.05</td>
<td>4.79</td>
</tr>
<tr>
<td>XVI</td>
<td>2.68</td>
<td>3.88</td>
<td>0.95</td>
<td>-2.59</td>
<td>-1.61</td>
<td>6.39</td>
</tr>
<tr>
<td>XVII</td>
<td>3.51</td>
<td>1.23</td>
<td>-4.05</td>
<td>-6.54</td>
<td>-3.72</td>
<td>6.53</td>
</tr>
<tr>
<td>XVIII</td>
<td>3.75</td>
<td>-2.09</td>
<td>-2.63</td>
<td>-0.14</td>
<td>2.01</td>
<td>1.30</td>
</tr>
</tbody>
</table>

Discussion

The choice of the predictors considered for inclusion in the GAM modelling in this study was made by reviewing previous studies (Yatsu and Watanabe 1996; Chen and Chiu 1999; Fan 2004; Chen and Tian 2005; Wang and Chen 2005). The variables selected had been shown to influence the squid population dynamics in these studies. It is likely that some of the selected environmental variables are correlated to each other. The focus of this paper is to evaluate the impacts of data spatial scales. Thus, to simplify the comparison process among scenarios, we only considered the interaction of longitude and latitude and omitted other interactions in the GAM modelling. It is also common to include vessel factors when analysing catch and effort data, but in this squid fishery, Chinese mainland squid jigging vessels were all modified from the same inshore bottom-trawlers and equipped with the main engine power of 120 KW × 2, squid-attracting lamp power of 112 KW and 16 squid-jigging machines. Moreover, all fishing activities occurred at night, and thus all fishing scales had large differences in MRDI values for both annual and monthly CPUEs.
vessels were almost identical in fishing power (Chen et al. 2008) so the vessel factor was not included in the model. The choice of predictors in this study is sufficient to address its focus on evaluating the impacts of spatial scale on the standardisation of CPUE. As to variable selection, some (such as SSS and temperature at the depth of 35 m) were not included in the model with the most data points and the finest spatial resolution because they were not significant ($P > 0.05$). This might result from the large variability at the finest spatial resolution. At coarser spatial resolutions, the data were aggregated and the mean values were used, which smoothed the data and reduced the variability. This is likely to have resulted in some environmental variables that were found to be not significant at fine spatial scales to become significant in scenarios at coarse spatial scales.

Plotting the spatial distribution of observed CPUE and environmental data aids in the understanding of the spatial dynamics of CPUE and environmental variables. However, because of the existence of large amounts of information and time series, it was unfeasible for us to plot all the data grouped in the different spatial scales in the study. For this reason, we only selected 1 month (September 1998) and plotted its spatial distribution of CPUE and environmental data for Scenario I ($0.5^\circ \times 0.5^\circ$) and Scenario XV ($2^\circ \times 2^\circ$). Due to the multivariate nature of the data and nonlinear nature of relationships between CPUEs and environmental variables, a comparison between plots of observed data at different spatial scales might not reveal whether there was a difference in the standardisation of CPUE resulting from different spatial scales. The complexity of the
data and interactions between CPUEs and environmental variables call for a quantitative evaluation, as carried out in this study, of the impacts that spatial scale choice has on CPUE standardisation.

The large RI differences suggest that the choice of spatial scale can have a significant impact on annual or monthly standardised CPUEs for this squid fishery. The CVs of both annual and monthly standardised CPUEs in the GAM modelling differed among spatial scales. This suggests that different spatial scales used in grouping data may result in different annual and monthly patterns of standardised CPUEs, thus affecting the interpretation of temporal variability in the squid population.

The CVs of annual standardised CPUEs tended to have large variations along the longitude scales when the latitude scale was 0.5°, 1°, or 2°, but little variation was observed along the latitude scales when the longitude scales were fixed (Figs 5a, 6a). This may be related to the spatial gradients of biological processes. Neon flying squid is a migrating and schooling species (Yatsu et al. 1997). Certain spatial scales of data collection by month could cover different spatial structures of the squid. In the North Pacific Ocean, there is a transition zone (TZ) situated between the sub-tropical front and the sub-arctic front, and the neon flying squid is considered to inhabit this zone and not beyond the fronts (Roden 1991). Neon squid distribution, determined by the TZ, is narrow (<10°) latitudinally, but much larger longitudinally (Chen and Tian 2005). In addition, while relatively uniform latitudinally, neon squid distribution exhibits much more variation in the longitudinal direction. Moreover, the CVs of monthly CPUEs showed the same increasing trend along the longitudinal direction as the CVs of annual CPUEs, but showed a different trend along the latitudinal direction when the longitude scales were fixed (Figs 5b, 6b). This could result from significant environmental variation between months, especially in the latitudinal direction. The trend of the CVs of annual standardised CPUEs along the sizes of spatial square could reveal that the variation of interannual standardised CPUEs was larger when a larger spatial scale was used in data grouping for the squid fishery. Thus, when standardised CPUEs derived from different spatial scales are used in stock assessment, they will result in different estimates of interannual stock abundance.

Commercial fisheries are often limited to particular areas where higher yields from target resources can be obtained, in particular when targeted species tend to aggregate in small areas (Booth 2000; Pitcher et al. 2000; Marrs et al. 2002; Verdoit et al. 2003). Spatial scale choice can influence the estimates of catch and fishing effort and their ratio (i.e. CPUE). Thus, different spatial scales can result in different conclusions about stock status in data analyses. Environmental factors are usually considered in analysing commercial fisheries data and these values are also determined by spatial scales in data collection and grouping (Thrush et al. 1998). Different ecological processes also operate at different spatial scales, so the choice of spatial

**Fig. 4.** Coefficient of variation (CV) of the (a) annual and (b) monthly standardised CPUEs for all the scenarios with different sizes of squares.

**Fig. 5.** Coefficient of variation (CV) of the (a) annual and (b) monthly standardised catch per unit effort (CPUE) for all the scenarios with different longitude scales.
The appropriate scale of spatial distribution of a species is generally unknown (Maynou 1998), so it is often difficult to determine an optimal spatial scale that matches the spatial scale of biological/ecological processes and fisheries processes in a fishery-dependent or fishery-independent sampling program. Large uncertainties always exist in our understanding of fisheries and biological processes. Fish populations tend to distribute neither uniformly nor randomly, and their abundance and structure are likely to vary with space and time (Legendre and Fortin 1989; Booth 2004). The fine-scale structure of fish distribution is difficult to measure using CPUE data although such information is important to fisheries stock assessment and management (Vignaux et al. 1998).

An optimal spatial scale for a sampling program should be consistent with spatial scales of the targeted species’ ecological processes and follow statistical sampling theory (Legendre and Fortin 1989; Nancy et al. 2000; Booth 2004). Economic, managerial and social factors should be also considered. The collection of data on a fine scale, however, tends to be costly, and may not be affordable for many fisheries. Even though the economic cost of spatial scales with the same size of square for data collection may be similar, different longitudinal and latitudinal scales can result in large disparities because of different environmental gradients in both longitudinal and latitudinal directions. Fine-scale data collection also broaches the sensitive issue of revealing exact fishing locations of commercial fishers. We must be careful in determining a spatial scale for data collection. If an optimal spatial scale cannot be clearly defined, it should be kept as fine as the budget allows. This makes a study similar to the present study possible, which can be used to identify the possible impacts of spatial scales and identify an optimal scale for future data collection and analysis. In practice, we often need to find a balance between scientific needs and cost limitations or personal commercial fishing information in determining the spatial scale for data collection.

We hope to identify a cost-effective spatial scale for fisheries data collection and analysis. However, different studies require data of certain precisions, and thus appropriate spatial scales of data collection. For some large scale fisheries, coarse data (large spatial scales) may yield similar results to those calculated at a finer spatial scale. For these types of fisheries, we can apply large spatial scales to data collection and analysis. For example, some studies (Pitcher et al. 2000; Marrs et al. 2002) identified CPUE calculated by the ICES statistical rectangles (1° × 0.5°), which made the identification of fisheries resource concentration areas with higher productivity difficult and had limited uses as a reliable index of abundance. However, for longline tuna fisheries, 1° × 0.5° might be too small to reflect the dynamics of fishing effort because the length of longlines usually exceeds 1°.

Fig. 6. Coefficient of variation (CV) of the (a) annual and (b) monthly standardised catch per unit effort (CPUE) for all the scenarios with different latitude scales.

Fig. 7. The mean relative variance (MRDI) calculated based on Eqn 6 for (a) annual and (b) monthly standardised catch per unit effort (CPUE) with different spatial scales.
The finest spatial scale considered in this study was 0.5° × 0.5° and only one level of temporal scale (monthly) was used. We did not evaluate finer spatial scales and any other temporal scales because of data limitations. Unlike a fishery-independent survey program, the CPUE data in this study were collected from Chinese commercial jigging vessels that used different spatio-temporal scales in data recording and reporting. For example, some vessels recorded and reported their catch daily with locations and timing; some vessels recorded daily, but only reported the records monthly in a ‘small fishing grid (0.5° × 0.5°)’ because they wanted to keep exact locations and timing of fishing activities confidential. To make the scales comparable between fisheries data and environmental data, we aggregated the data over the grid cells and over the month. This is sufficient regarding the objective of this study: evaluating the impact of different spatial scales on the GAM-based CPUE standardisation. Although limited in scope and data, the present study indicates that the choice of spatial scale in data collection and grouping can greatly influence fisheries stock assessment and management, and emphasises the importance of evaluating such impacts.

Acknowledgements

We thank the Chinese squid-jigging technology group for providing fisheries data and the IRI/LEDEO Climate Data Library of Columbia University for providing the environmental data. We thank the three anonymous referees and Editor for their comments, which greatly improved the quality of the paper. Comments from Sam Truesdell on editorial changes are also greatly appreciated. This study is supported by the Program for New Century Excellent Talents in University (NCET-06-0437), Shanghai Leading Academic Discipline Project (Project # S30702) and Excellent Young Teacher Fund of Shanghai Education Commission.

References


Manuscript received 15 April 2009, accepted 13 May 2009

http://www.publish.csiro.au/journals/mfr