Acta Oceanol. Sin., 2019, Vol. 38, No. 10, P. 9–19
https://doi.org/10.1007/s13131-019-1485-4
http://www.hyxb.org.cn
E-mail: hyxbe@263.net

Evaluation of the performance of alternative assessment configurations to account for the spatial heterogeneity in age-structure: a simulation study based on Indian Ocean albacore tuna

Wenjiang Guan[1, 2*], Jiawen Wu[1], Siquan Tian[1, 2]

[1] College of Marine Sciences, Shanghai Ocean University, Shanghai 201306, China

[2] Key Laboratory of Sustainable Exploitation of Oceanic Fisheries Resources, Ministry of Education, Shanghai 201306, China

Received 10 July 2018; accepted 28 February 2019

Abstract

Various population structures or spatial heterogeneities in population distribution have been an important source of model misspecification and have had an impact on estimation performance in fisheries stock assessment. In this study, we simulated the Indian Ocean albacore spatial heterogeneity in age-structure using Stock Synthesis according to the stage-dependent migration rate and region-dependent fishing mortality rate and generated the stock assessment data. Based on these data, we investigated the performances of different spatial configurations, selectivity curves and selections of CPUE (catch per unit effort) indices of the assessment models which were used to account for spatial heterogeneity. The results showed: (1) although the spatially explicit configurations, which exactly matched the operating model, provided unbiased and accurate estimates of relative spawning biomass, relative fishing mortality rate and maximum sustainable yield in all simulation scenarios, their performance may be very poor if there were mismatches between them and the operating model due to gaps in knowledge and data; (2) for spatially explicit assessment configuration, the correct boundary was required, but for non-spatially explicit assessment configuration, it seemed more important for analysts to partition the area to properly reflect the transition in field data and to effectively account for the impacts of ignoring the spatial structure by using the additional spatially referenced parameters; (3) although the areas-as-fleets methods and flexible time-varying selectivity curves could be used as better alternative approaches to account for spatial structure, these configurations could not completely eliminate the impacts of model misspecification and the quality of estimates of different quantities from the same assessment model may be inconsistent or the performance of the same assessment configuration may fluctuate significantly between simulation scenarios; (4) although the worst estimates could generally be avoided by using multiple CPUE indices, there were no best solutions to select or regenerate the CPUE indices to account for the impacts of the ignored spatial structure to obviously improve the quality of stock assessment. Compared with the results of assessment model configurations which are used to account for the spatial structure by different modelers, the performances of the configurations are always case-specific except for spatially explicit configurations which exactly match the operating model. In this sense, our study will not only provide some insights into the current Indian Ocean albacore stock assessment but also enrich existing knowledge regarding the performance of assessment configurations to account for spatial structure.

Key words: spatial structure, simulation, stock assessment, Indian Ocean, *Thunnus alalunga*

Citation: Guan Wenjiang, Wu Jiawen, Tian Siquan. 2019. Evaluation of the performance of alternative assessment configurations to account for the spatial heterogeneity in age-structure: a simulation study based on Indian Ocean albacore tuna. Acta Oceanologica Sinica, 38(10): 9–19, doi: 10.1007/s13131-019-1485-4

1 Introduction

Population structures can range from a single population in multiple areas, to overlapping populations with natal homing, to subpopulations with reproductive mixing (Stephenson et al., 2009; Ying et al., 2011; Guan et al., 2013; Hurtado-Ferro et al., 2014; Punt et al., 2015). Spatial distribution of fish stocks can be in spatially non-homogenous patterns of abundance, sex, age or length across their habitats in response to stage-dependent migrations, age-specific or sex-specific mortality, etc. (Chen et al., 2005; Berger et al., 2012; Farley et al., 2012). All of the above result in a complex spatial structure for the stock being assessed (Goethel et al., 2011). There are three common approaches to dealing with spatial structure in fisheries stock assessment. First, the spatial structure is ignored according to an assumption of spatially homogeneous distribution and the data are aggregated spatially in the stock assessment model (Guan et al., 2016a). Second, although the spatially homogeneous distribution is still assumed in the stock assessment model, spatially referenced parameters are applied to account for the impacts of spatial structure (Berger et al., 2012). Third, spatially explicit assessment

Foundation item: The National Key Research and Development Program of China under contract No. 2016YFC1400903; the NSFC-Zhejiang Joint Fund for the Integration of Industrialization and Informatization under contract No. U1609202.
*Corresponding author, E-mail: wjguan@shou.edu.cn

models, such as Stock Synthesis (Methot and Wetzel, 2013) or Multifan-CL (Fournier et al., 1998), are applied in stock assessment to take spatial structure into account (Punt et al., 2015).

Spatially explicit assessment models, which can represent more complicated spatial population and fishery structures with increasing biological and fishery realism, are still rare (Punt et al., 2015). This is because it is difficult for the models to estimate movement rates in the absence of tagging data that is costly and often unavailable (Berger et al., 2012; Punt et al., 2015) or to converge to the true minimum of the objective function given the number of parameters and potentially confounded process (Szuwalski and Punt, 2015). The correct delineation of the boundaries of the spatial structures poses another serious challenge to their application (Fonteneau, 2016; Langley and Hoyle, 2016). Although some studies (Berger et al., 2012; Guan et al., 2013) have shown that spatially-aggregated assessment approaches will produce biased results and have poor impacts on the performance of management strategies given that the assumption of homogenous distribution has been violated to some extent (Punt et al., 2015), these approaches are still widely used in modern stock assessment (Berger et al., 2012; Guan et al., 2016a). Under some situations, these approaches have even outperformed more complicated models, such as spatially explicit assessment models (Punt et al., 2015). Because the spatially implicit approaches, which allow for spatial structure by means of spatially referenced parameters, are more transparent, practical, and applicable to many fisheries in stock assessment (Berger et al., 2012), they are also increasingly commonly used to account for the effects of spatial structure. For example, the areas-as-fleets approach is quite widespread in statistical catch at age stock assessment models (Cope and Punt, 2011; Hurtado-Ferro et al., 2014; Waterhouse et al., 2014). At present, although some simulation studies have been used to identify the appropriate assessment configurations that can be used to address spatial structure due to specific processes such as age-specific movement or region-dependent growth, recruitment and fishing mortality (Punt et al., 2015; Lee et al., 2017), there seems to be no consensus on how to configure the assessment models to account for the effects of spatial structure and more research is still needed.

The Indian Ocean albacore tuna (*Thunnus alalunga*) is distributed throughout the Indian Ocean between 25°N and 45°S and is one of the main target species of the commercial tuna fishery (Chen et al., 2005). Although the Indian Ocean albacore is considered to constitute a single stock, its size or age composition varies with latitude (Chen et al., 2005; Nikolic et al., 2014). In general, the Indian Ocean albacore is classified by latitude with the large and mature group distributed northward of 10°S, the spawning group between 10°S and 30°S, and the small and immature group southward of 30°S (Chen et al., 2005). A spatially explicit assessment model had been developed to assess the Indian Ocean albacore, but it was considered that the model was not sufficiently reliable for stock assessment purposes and was not progressed (Langley and Hoyle, 2016). As a consequence, spatially-aggregated models (Nishida et al., 2016; Guan et al., 2016b; Matsumoto, 2016) or the areas-as-fleets method (Langley and Hoyle, 2016) was used for Indian Ocean albacore stock assessment. In 2014, using the areas-as-fleets method, the Indian Ocean was partitioned into two regions demarcated at 20°S to account for the spatial patterns of size or age structure for albacore stock assessment (Hoyle et al., 2014). However, in 2016, the Indian Ocean was partitioned into four regions, demarcated at 25°S and 75°E, for albacore stock assessment to allow for potential longitudinal variations in the major fishery data sets as well as the zonal variation of size or age (Langley and Hoyle, 2016). The boundaries of the geographical stratification used in the stock assessment appear to be widely questionable (Fonteneau, 2016), as there does not seem to be enough information to correctly delineate the boundaries, and the effects of the changes to the position of the boundaries are unclear. At the same time, single or multiple standardized CPUE (catch per unit effort) indices from different regions were used as indices of relative abundance in albacore stock assessment and their effects on the results of stock assessment were different (Matsumoto, 2016; Nishida et al., 2016; Langley and Hoyle, 2016). There does not appear to be enough information to choose the appropriate CPUE indices or their combinations to improve the performance of the stock assessment.

In this paper, we simulated the Indian Ocean albacore spatial heterogeneity in age to generate the data used in stock assessment which consists of catch, fishery-dependent indices of relative abundance (i.e., CPUE indices) and age composition of the catch. Based on these data, we investigated the performances of a range of assessment configurations that could be used to account for the spatial structure. More specifically: (1) the performance of spatially-aggregated, spatially-implicit and spatially-explicit methods with different selectivity curves and correct or incorrect boundaries are evaluated according to their ability to provide reliable estimates of the management quantities of interest; (2) the stability of the performance of these assessment configurations under different simulation scenarios which were created by using different movement rate, fishing intensity and distribution are explored; (3) in addition, the performance of several selection schemes for CPUE indices are also investigated. The results of this simulation study will not only provide some insights into the current Indian Ocean albacore stock assessment but also enrich existing knowledge regarding the performance of assessment configurations to account for spatial structure.

2 Materials and methods

2.1 *Operating model*

2.1.1 *Assumptions of population dynamics*

The operating model assumes that the Indian Ocean albacore has one sex and constitutes a single stock, but that its age composition changes with latitude. To simulate the spatial patterns in age, the Indian Ocean is partitioned into three regions demarcated at 10°S and 30°S which corresponds to the classification of albacore age composition by latitude (Chen et al., 2005) and the albacore is assumed to be homogenously distributed in each region with the same nature mortality, maturity, and growth across the regions. The annual recruits were calculated using the deterministic Beverton-Holt stock-recruitment function during the period from 1935 to 1949 and then annual recruitment deviates were added for 1950 to 2014 (Methot and Wetzel, 2013; Punt et al., 2015). The total spawning biomass is the total weight of sexually mature individuals over the three regions and the parameter values used to determine the recruits are defined in Table 1. All the recruits are assumed to enter Region 1 (i.e., southward of 30°S) first, and then the fish gradually disperse into Region 2 (i.e., between 10°S and 30°S), and finally into Region 3 (i.e., Northward of 10°S). The age-specific movement rates from Region 1 to Region 2 or from Region 2 to Region 3 are interpolated using a piecewise linear function in logarithmic space which is defined by two ages and their corresponding movement rates (Methot and Wetzel, 2013).

Table 1. Parameter values used in the simulation

Parameter	Value	Reference
Natural mortality rate at age/a^{-1}	0.4 (Age 0), 0.364 1 (Age 1), 0.328 3 (Age 2), 0.292 4 (Age 3), 0.256 6 (Age 4), 0.220 7 (Age 5–29)	Nishida et al. (2014)
Maturity rate at age/a^{-1}	0.0 (Age 0), 0.0 (Age 1), 0.0 (Age 2), 0.0 (Age 3), 0.09 (Age 4), 0.47 (Age 5), 0.75 (Age 6), 0.88 (Age 7), 0.94 (Age 8), 0.97 (Age 9), 0.99 (Age 10), 1.00 (Age 11–29)	Langley and Hoyle (2016)
Weight and length relationship	W=1.371 8×10^{-5} $L^{3.097\,3}$	Penney (1994)
Growth equation	L=124.1×(1–e$^{-0.164×(t+2.239\,0)}$)	Wells et al. (2013)
Spawner-recruit relationship	Beverton-Bolt	Langley and Hoyle (2016)
Spawner-recruit steepness	0.8	Langley and Hoyle (2016)
Log of recruitment at virgin biomass	9.589 685	Langley and Hoyle (2016)
Recruitment variability	0.6	Langley and Hoyle (2016)
Bias-adjustment factor for year	1.0 (1950–2014)	our assumption
Years for recruitment deviates added	1950–2014	our assumption
Age selectivity curve for the longline fleets	$S_t = \dfrac{1}{1+\mathrm{e}^{-2.944\,4\times(t-3)/2.355\,55}}$	our assumption
Catchability for the longline fleets	1.0	our assumption

Note: W is body weight (kg), L fork length (cm), and t age (a).

The operating model covers an 80-year period, i.e., from 1935 to 2014 using a yearly time step and the start year of catch is 1950. A plus-group is used as an accumulator for all ages greater than 29 years to reduce the age categories.

2.1.2 *Fisheries*

The Indian Ocean albacore is caught using longlines, purse seines, handline, trolling and other gears, but the catch from longline fisheries has accounted for more than 90% of the total albacore catch in the Indian Ocean apart from a few years in the early-1950s and the period from 1985 to 1992 when drift nets were employed. For the sake of simplicity, we assume only longline fisheries operate simultaneously in the three regions for albacore. The longline fleets keep the same selectivity and catchability (Table 1) across the regions, but the fishing effort is different for each region.

2.1.3 *Simulation scenarios*

We considered a series of six simulation scenarios in which the movement rate, fishing effort and its distribution were different (Table 2).

In Scenario 1, which is selected as the base-case scenario, the piecewise linear functions used to calculate the age-specific movement rate are defined by a vector (1, 5, 0.05, 0.20) for the movement from Region 1 to Region 2 or a vector (4, 8, 0.05, 0.20) for the movement from Region 2 to Region 3. The first two components in the vector are two ages for the young and old fish and the last two components are the movement rates for the two ages. The time series of annual fishing mortality rate (Fig. 1a) estimated by Guan et al. (2016b) is multiplied by 3 and then used as the time series of total fishing effort over the three regions. The annual proportion of fishing effort in each region is equal to the ratio of the annual catch in the region to the annual total catch over the three regions (Fig. 1b) and the albacore catch data from 1950 to 2014 were obtained from the Indian Ocean Tuna Commission (IOTC) website which were published for the 6th Working Party on Temperate Tunas in 2016 (http://www.iotc.org/meetings/6th-working-party-temperate-tunas-wptmt06).

Scenarios 2 and 3 are similar to Scenario 1 except that the movement vectors are different, namely, the ages of the two movement vectors remain unchanged in Scenarios 2 and 3, but each age-specific movement rate is increased 2 times in Scenario 2 and the movement rates for young and old fish of the two movement vectors are interchanged in Scenario 3 (Table 2).

Compared with Scenario 1, only the spatial distributions of the fishing effort are different in Scenarios 4 and 5. The fishing effort is uniformly distributed over the three regions in Scenario 4, but the proportions of fishing effort in Region 1 and Region 2 are interchanged in Scenario 5 (Table 2).

Scenario 6 is similar to Scenario 1 except that the total fishing effort over the three regions is twice the level of Scenario 1 (Table 2).

2.1.4 *Observation data*

The parametric bootstrap feature of Stock Synthesis (Version 3.24f) was used to generate 120 simulated data sets for each scenario and each simulated data set consisted of a 65-year (i.e., from 1950 to 2014) catch (in number), CPUE indices (i.e., fishery-dependent indices of relative abundance) and age-composition data of catch in each region. The coefficient of variation of the CPUE indices and catch is 0.1 and 0.01, respectively. There is no

Table 2. Characteristics of the six simulation scenarios

Scenarios	Movement vector from Region 1 to Region 2	Movement vector from Region 2 to Region 3	Total fishing effort	Proportion of fishing effort along regions
1 (base-case)	1, 5, 0.05, 0.20	4, 8, 0.05, 0.20	3×F_y	($P_{1,y}$, $P_{2,y}$, $P_{3,y}$)
2	1, 5, 0.10, 0.40	4, 8, 0.10, 0.40	3×F_y	($P_{1,y}$, $P_{2,y}$, $P_{3,y}$)
3	1, 5, 0.20, 0.05	4, 8, 0.20, 0.05	3×F_y	($P_{1,y}$, $P_{2,y}$, $P_{3,y}$)
4	1, 5, 0.05, 0.20	4, 8, 0.05, 0.20	3×F_y	(1/3, 1/3, 1/3)
5	1, 5, 0.05, 0.20	4, 8, 0.05, 0.20	3×F_y	($P_{2,y}$, $P_{1,y}$, $P_{3,y}$)
6	1, 5, 0.05, 0.20	4, 8, 0.05, 0.20	6×F_y	($P_{1,y}$, $P_{2,y}$, $P_{3,y}$)

Note: F_y is fishing mortality in year y and $P_{x,y}$ is the proportion of fishing effort in Region x and year y. F_y and $P_{x,y}$ are shown in Fig. 1. The first two components of the movement vector are two ages for the young and old fish and the last two components are movement rates for the two ages.

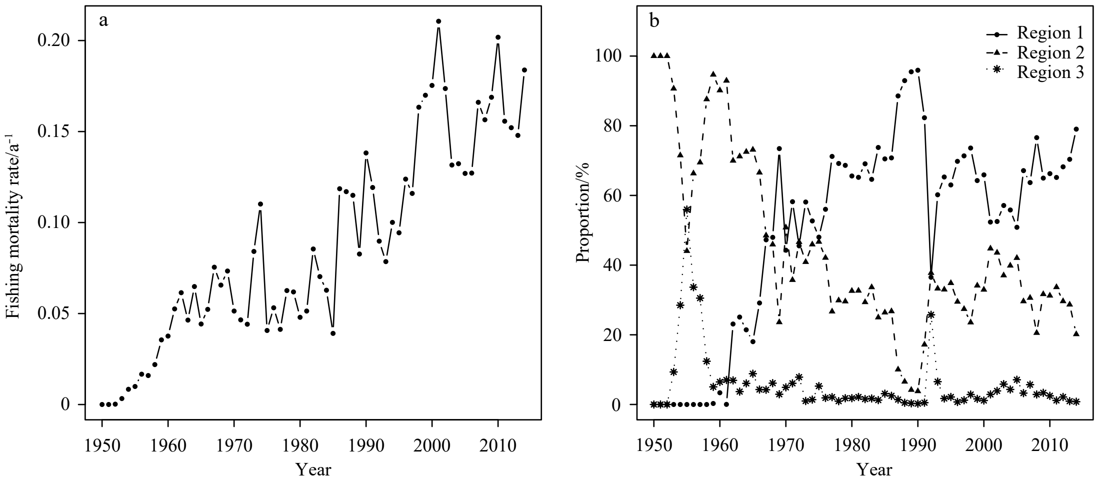

Fig. 1. The fishing mortality rate (a) and the annual proportion of fishing effort (b) in each region. The fishing mortality is estimated by Guan et al. (2016b) and the annual proportion of fishing effort in each region is calculated as the ratio of the catch in each region to total albacore catch in the year. The albacore catch data from 1950 to 2014 for the three regions was downloaded from the IOTC website (http://www.iotc.org/meetings/6th-working-party-temperate-tunas-wptmt06).

age-reading error in age-estimates and the effective sample size (ESS) for the age-composition data is 200.

2.2 *Configurations of the assessment models*

2.2.1 *Spatial configuration of the assessment models*

We considered seven different spatial configurations (Table 3) as follows:

(1) Spatially-explicit configuration with accurate boundaries (SCAB). This configuration is fully consistent with the spatial population dynamic of the operating model.

(2) Spatially-explicit configuration with inaccurate boundaries (SCIB). This configuration is consistent with the spatial population dynamic of the operating model except that the boundaries of the three regions are delineated incorrectly which causes 25% of the area of Regions 1 and 3 to be incorrectly allocated to Region 2.

(3) Simple spatially-aggregated configuration (SSC). This configuration assumes that fish and fishing effort are homogenously distributed across the three regions and the catch, CPUE indices and age-composition data are aggregated over all regions respectively. For this configuration, only one region and one fleet are assumed in the stock assessment models.

(4) By region spatially-aggregated configuration with accurate boundaries (BRSCAB). This configuration assumes fish are homogenously distributed across the three regions but each region contains a separate longline fleet. Although the catch, CPUE indices and age-composition data for each fleet are kept separate, the selectivity for the fleets is the same except that the catchability can be different among fleets.

(5) By region spatially-aggregated configuration with inaccurate boundaries (BRSCIB). This configuration is similar to BRSCAB except that 25% of the area of Regions 1 and 3 is incorrectly allocated to Region 2.

(6) Areas-as-fleets configuration with accurate boundaries (ACAB). This configuration is the same as the configuration BZSCAB except selectivity is estimated separately for each fleet.

(7) Areas-as-fleets configuration with inaccurate boundaries (ACIB). This configuration is similar with ACAB except that 25% of the area of Regions 1 and 3 is incorrectly allocated to Region 2.

2.2.2 *Selectivity curves of the assessment models*

(1) Logistic selectivity curve (SL).

(2) Double-normal selectivity curve (SD).

(3) Time-varying selectivity curve (ST). The time-varying selectivity curve is implemented by estimating the annual deviates for the age of the peak selectivity of the double-normal selectivity curve (Methot and Wetzel, 2013) and the deviates are assumed to be from a normal distribution with a standard deviation of three years which is similar to the time-varying selectivity curve used in the Indian Ocean albacore stock assessment (Langley and Hoyle, 2016).

2.2.3 *Using CPUE indices scheme*

(1) Using all CPUE indices scheme (IA). In this scheme, all three CPUE indices from the respective three regions are simultaneously used in the assessment model. For spatial configurations SCIB, SSC, BRSCIB and ACIB, the CPUE indices need to be

Table 3. Spatial configurations of the assessment models and their characteristics

Spatial configurations	Number of regions	Region boundary	Number of fleets	Selectivity curve for the fleets	Catchability for the fleets
SCAB	3	correct	1	same	same
SCIB	3	incorrect	1	same	same
SSC	1	incorrect	1	same	same
BRSCAB	1	correct	3	same	can be different
BRSCIB	1	incorrect	3	same	can be different
ACAB	1	correct	3	can be different	can be different
ACIB	1	incorrect	3	can be different	can be different

regenerated according to the new definition of the regional boundaries or fleets (See Sections 2.2.1 and 2.3).

(2) Using a single CPUE index or their pairwise combination schemes (IS). For this scheme, the three CPUE indices from three regions or their pairwise combinations are respectively used in the assessment model. For spatial configuration SSC, the pairwise combination of the three CPUE indices means a new CPUE index used in the assessment model is regenerated by amalgamating the two corresponding CPUE indices (See Section 2.3).

2.2.4 *Data weighting scheme*

The weights assigned to the data are the same as in the operating model except for those data which are regenerated due to region combination and whose weights are recalculated as in Section 2.3 (See below).

For each scenario, 21 combinations of assessment configurations related to seven spatial configurations, three selectivity curves and one CPUE indices scheme IA are explored to investigate the effects of spatial configuration, selectivity curve and simulation scenario. Another 18 combinations of assessment configurations related to three spatial configurations (i.e., ECAB, SSC, and ACAB), six CPUE indices usage scenarios (i.e., IS) and one selectivity curve SD are added in Scenario 1 to investigate the effects of the CPUE indices usage scenarios.

2.3 *Data reorganization and weight adjustment*

The simulation data (i.e., catch, CPUE indices and age-composition data) of each scenario need to be re-organized according to the spatial configuration of the assessment models. Because selectivity and catchability are the same over all regions and fish are assumed to be homogenously distributed in each region in the operating model, the data for the new regions defined in the spatial configuration of the assessment model can be regenerated in the following ways:

$$C_M = \sum_{r=1}^{n_M} p_{r,\,M} C_r, \tag{1}$$

$$CPUE_M = \sum_{r=1}^{n_M} p_{r,\,M} CPUE_r, \tag{2}$$

$$O_{y,\,a,\,M} = \frac{\sum\limits_{r=1}^{n_M} O_{y,\,a,\,r} p_{r,\,M} C_r}{\sum\limits_{a=0}^{A}\sum\limits_{r=1}^{n_M} O_{y,\,a,\,r} p_{r,\,M} C_r}, \tag{3}$$

where C_M or C_r is the catch (in number) in the new region M or the original region r, $p_{r,\,M}$ is the proportion of the original region r allocated to the new region M which is defined in each spatial configuration of the assessment model, $CPUE_M$ or $CPUE_r$ are the CPUE indices in the new region M or the original region r, $O_{y,\,a,\,M}$ is the observed proportion at age a for the new region M in year y, $O_{y,\,a,\,r}$ is the observed proportion at age a for the original region r in year y, A is maximum age, and n_M is the number of original region which will be amalgamated in the new region M.

The coefficients of variation or ESSs for the new region, which is only split, are left unchanged, but the other new regions are adjusted as follows (Cope and Punt, 2011; McAllister and Ianelli, 1997):

$$CV_M = CV_r\sqrt{n_M}, \tag{4}$$

$$ESS_M = \frac{\sum\limits_{y=1}^{N} \dfrac{\sum\limits_{a=0}^{A} P_{y,\,a,\,M}\left(1 - P_{y,\,a,\,M}\right)}{\sum\limits_{a=0}^{A}\left(O_{y,\,a,\,M} - P_{y,\,a,\,M}\right)\left(O_{y,a,M} - P_{y,\,a,\,M}\right)}}{N}, \tag{5}$$

where CV_r is the coefficient of variation of CPUE or catch for the original region r, CV_M is the coefficient of variation of CPUE or catch for the new region M, N is the number of years, $P_{y,\,a,\,M}$ is the true proportion at age a for the new region M in year y, and ESS_M is the ESS for the new region M.

2.4 *Parameter estimation*

To keep the inconsistencies between the operating model and the assessment model to a minimum, except for the intended model misspecifications, the SS3 is also used as an assessment model to estimate the recruitment at virgin biomass, the annual recruitment deviates, the parameters of selectivity curve, the catchability and the full-selected fishing mortality rates. The spatially explicit assessment models also estimate the movement rates between regions. The other parameters are assumed to be known exactly as in the operating model.

The convergence of the assessment model is verified by the ability of SS3 to invert the Hessian matrix, and simulated data sets are removed from the analysis if the Hessian matrix of the model fitting the data sets is not positive-definite (Hurtado-Ferro et al., 2014). The performance of each assessment configuration in each simulation scenario is evaluated based on the first 100 convergence estimates from the 120 simulated data sets.

2.5 *Performance measures*

The assessment model provides many outputs for each assessment configuration (Methot and Wetzel, 2013). However, the focus for this paper is the estimates of maximum sustainable yield (MSY), the relative spawning stock biomass (RS) which equals the ratio of SSB_{cur} (current spawning stock biomass) to SSB_{MSY} (spawning stock biomass at MSY) and the relative fishing mortality rate (RF) which equals the ratio of F_{cur} (current fishing mortality rate) to F_{MSY} (fishing mortality rate at MSY). These quantities are important biological benchmarks for fishery managers to define the current stock status and to set fishing regulations.

The median relative error (MRE) and the median absolute relative error (MARE) across iterations of each assessment configuration within a scenario are calculated to determine the bias and accuracy of the estimates of the quantities of interest as in Ono et al. (2015).

$$\mathrm{MRE} = \mathrm{Median}\left(\frac{E_1 - T_1}{T_1}, \ldots, \frac{E_{100} - T_{100}}{T_{100}}\right) \times 100, \tag{6}$$

$$\mathrm{MARE} = \mathrm{Median}\left(\left|\frac{E_1 - T_1}{T_1}\right|, \ldots, \left|\frac{E_{100} - T_{100}}{T_{100}}\right|\right) \times 100, \tag{7}$$

where the E is the estimate of the quantity of interest, T is the true value, and the subscript indicates the iteration number. The quantity is considered to be accurately estimated when MARE is equal to or below 16% or to have a low bias when MRE is between –4% and 4% (Carvalho et al., 2017).

3 Results

3.1 *Convergence problem of the assessment model*

When spatially explicit configurations (i.e., SCAB and SCIB) and the selectivity curve ST are used, there are not enough convergence models (less than 50) to draw conclusions. Apart from the cases mentioned above, the number of convergence models for each assessment configuration in each scenario is more than 100 and our analysis is based on the first 100 convergence estimates.

3.2 *Performance of assessment configurations in the base-case scenario*

For the two spatially explicit configurations, SCAB provides unbiased (i.e., MRE between 4% and –4%) and accurate (i.e., MARE less than 16%) estimates of the three quantities, but all the estimates from SCIB are biased and the RF or RS and RF are inaccurately estimated from SCIB with selectivity curve SL or SD (Fig. 2).

The five non-spatially explicit configurations provide biased estimates in most instances (Fig. 2). For spatially-aggregated configurations SSC, BRSCAB and BRSCIB, the estimates of the three quantities are very poor except for MSY from SSC or RS from BZ-SCIB. In contrast with the three spatially-aggregated configurations, the performance of the areas-as-fleets configurations, irrespective of correct or incorrect boundaries (i.e., ACAB and ACIB), are markedly improved and even better than SCIB when the selectivity curve SD is used.

3.3 *Impacts of simulation scenarios*

The change of movement rates (i.e., Scenario 2 and Scenario 3) has a relatively small impact on the estimates of the three quantities for spatially explicit configurations, especially for SCAB with selectivity curve SL. However, for the five non-spatially explicit configurations, the impacts are significant and there is a large fluctuation in the relative errors of the estimates of the three quantities (Table 4). Another important characteristic is that all but two median relative errors for the three quantities simultaneously go down or go up with the change of movement rates for the non-spatially explicit configurations SSC, ACAB and ACIB (Table 5). Specifically, compared with the base-case scenario, the median relative errors for RS and MSY go down, or for RF go up, simultaneously when the non-spatially explicit configurations SSC, ACAB and ACIB are used in Scenario 2, but the results are opposite in Scenario 3. However, for BRSCAB and BRSCIB, the median relative errors for RS and MSY go down, or for RS go up, in most instances in Scenarios 2 and 3 in comparison with the base-case scenario (Fig. 2 and Table 4).

The effects of the spatial distribution of fishing effort on the estimates of the three quantities are relatively small for SCAB, but remarkable for the other assessment configurations. When the fishing effort is uniformly distributed across the regions (i.e., Scenario 4), the quality of the parameter estimates is obviously improved in most instances, specifically for SCIB and SSC from which all the quantities are accurately estimated (Table 4). However, in Scenario 5, the effects are inconsistent for different assessment configurations. For example, in Scenario 5, the estimates of the three quantities are markedly worse for SCIB with selectivity curve SL or almost unaffected for ACAB with selectivity curve SL and SD, and obviously improved for ACAB with selectivity curve ST (Fig. 2, Table 4).

When fishing intensity is increased 2 times, i.e., in Scenario 6, the estimates of the three quantities are almost unaffected for SCAB or slightly improved for SCIB (Fig. 2, Table 4). The increase of fishing intensity also has different effects on the estimates of the three quantities for the five non-spatially explicit configurations. For example, in comparison with the base-case scenario, the estimates of RS and RF have deteriorated, but MSY improves when ACAB and selectivity SL are used or the performance of parameter estimates change little when ACIB and selectivity ST are used (Fig. 3, Table 4).

3.4 *Effect of the choice of CPUE indices*

The CPUE indices from different regions or their combinations used as indices of relative abundance in the stock assess-

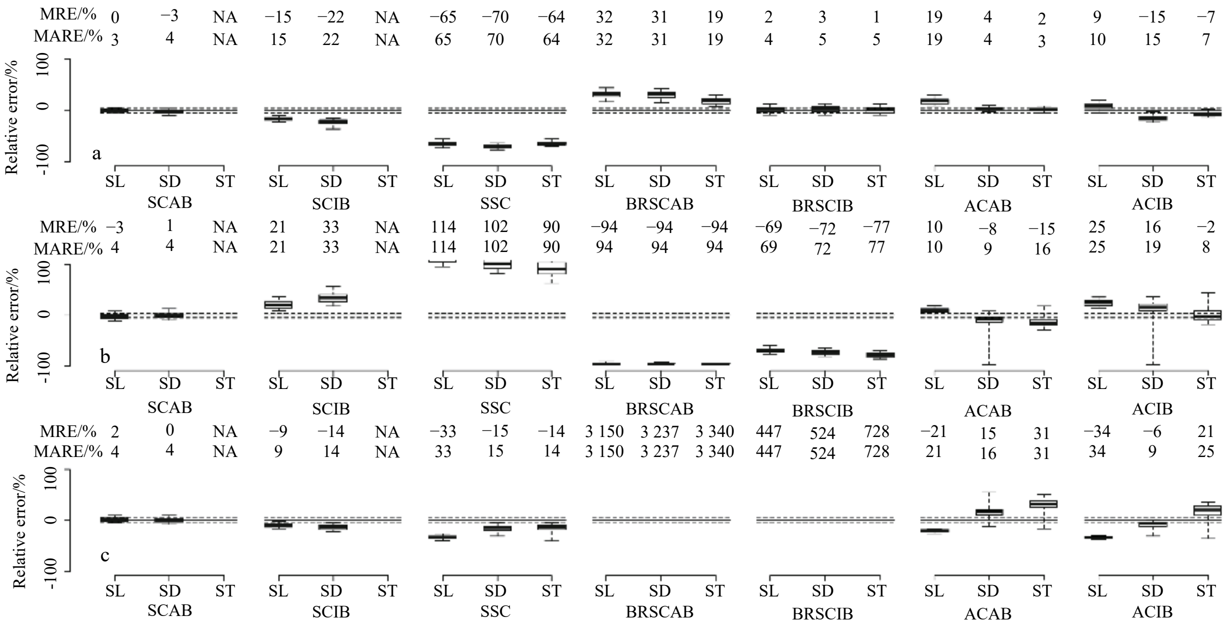

Fig. 2. The relative errors for the three quantities of interest. The results in this figure are for Scenario 1. a. The relative errors for relative spawning biomass, b. the relative errors for relative fishing mortality rate, and c. the relative errors for maximum sustainable yield. The band inside the box is the median, the boxes cover 50% of the relative errors and the line at the end of the whiskers represents the 95th (upper) or 5th (lower) percentile of the relative errors respectively.

Table 4. The median relative error (MRE) and the median absolute relative error (MARE) for maximum sustainable yield (MSY), the relative spawning stock biomass (RS) and the relative fishing mortality rate (RF)

Scenario	Quantities	Assessment configurations and selectivity curves																				
		SCAB			SCIB			SSC			BRSCAB			BRSCIB			ACAB			ACIB		
		SL	SD	ST	SL	SD	ST	SL	SD	ST	SL	SD	ST	SL	SD	ST	SL	SD	ST	SL	SD	ST
2	RS-MRE	1	–2	–2	–10	–22	–9	–75	–82	–71	8	12	11	–15	–11	–6	–7	–5	–4	–12	–25	–11
	RS-MARE	3	4	3	10	22	10	75	82	71	8	12	11	15	11	7	8	7	4	12	25	11
	RF-MRE	–4	3	6	16	44	14	278	235	165	–49	–62	–62	3	–30	–48	103	9	0	137	59	20
	RF-MARE	8	9	7	16	45	20	278	235	165	49	62	62	8	30	48	103	18	9	137	68	21
	MSY-MRE	3	–1	0.0	–12	–20	–5	–63	–31	–31	223	339	343	56	147	238	–60	–5	9	–67	–27	–2
	MSY-MARE	6	7	7	12	20	15	63	31	31	223	339	343	56	147	238	60.0	7.6	11	67	28	8
3	RS-MRE	0	–8	–8	–10	–20	–17	–49	–50	–48	23	26	21	3	4	3	19	24	26	3	1	5
	RS-MARE	4	8	8	10	20	17	49	50	48	23	26	21	4	4	4	19	24	26	4	4	5
	RF-MRE	0	13	15	22	40	34	94	86	85	–35	–41	–42	–11	–16	–19	–16	–38	–54	9	–3	–13
	RF-MARE	7	13	15	22	40	34	94	86	85	35	41	42	11	16	19	16	38	54	9	7	14
	MSY-MRE	0	–2	–3	–3	–6	–4	–12	–5	–9	151	177	200	81	94	107	56	130	219	19	51	54
	MSY-MARE	4	3	4	4	6	5	12	5	9	151	177	200	81	94	107	56	130	219	19	51	54
4	RS-MRE	0	3	4	0	2	3	0	1	1	15	15	12	3	4	2	11	–2	–3	10	–2	–5
	RS-MARE	2	4	4	2	3	4	3	3	3	15	15	12	6	6	5	11	3	4	10	4	5
	RF-MRE	0	–4	–8	4	–6	–9	–1	–8	–6	–63	–63	–65	–47	–51	–50	2	–9	–11	5	–11	–12
	RF-MARE	3	5	9	4	6	9	3	8	7	63	63	65	47	51	50	3	10	11	5	11	12
	MSY-MRE	1	5	5	–5	6	6	1	10	10	285	286	326	169	191	195	–14	17	23	–17	21	24
	MSY-MARE	3	5	5	5	6	6	2	10	10	285	286	326	169	191	195	14	17	23	17	21	24
5	RS-MRE	1	2	3	27	27	5	39	64	15	29	29	20	42	44	40	15	–1	–2	42	26	14
	RS-MARE	3	3	3	27	27	5	39	64	15	29	29	20	42	44	40	15	3	3	42	26	14
	RF-MRE	–1	–3	–2	–33	–33	2	106	–38	20	–93	–95	–95	–81	–85	–82	8	–7	–7	25	–24	–4
	RF-MARE	3	4	3	33	33	4	106	39	20	93	95	95	81	85	82	8	8	8	25	24	7
	MSY-MRE	1	2	5	41	43	–5	–71	20	–29	1 998	2 838	2 954	523	703	581	–22	14	15	–45	22	–6
	MSY-MARE	3	4	5	41	43	6	71	25	29	1 998	2 838	2 954	523	703	581	22	14	16	45	28	6
6	RS-MRE	0	–4	–3	–15	–22	–15	–69	–74	–67	61	64	51	16	17	16	43	11	7	31	–8	0
	RS-MARE	2	4	3	15	22	15	69	74	67	61	64	51	16	17	16	43	11	7	31	9	5
	RF-MRE	0	1	1	16	18	8	63	57	54	–81	–83	–85	–64	–66	–67	–17	–14	–15	–10	–6	–5
	RF-MARE	3	4	4	16	18	8	63	57	54	81	83	85	64	66	67	17	14	15	10	6	8
	MSY-MRE	0	0	3	0	2	7	–5	22	10	1 023	1 197	1 484	470	513	604	11	25	40	–5	15	29
	MSY-MARE	3	3	5	3	3	7	5	22	14	1 023	1 197	1 484	470	513	604	11	25	40	5	15	29

Note: The unit is %.

Table 5. The difference in median relative error for the three quantities between Scenarios 2 and 1 or between Scenarios 3 and 1

Scenario	Quantities	Spatial configurations and selectivity curves								
		SSC			ACAB			ACIB		
		SL	SD	ST	SL	SD	ST	SL	SD	ST
2 and 1	RS	–11	–12	–7	–26	–9	–6	–21	–10	–4
	RF	163	133	75	93	17	15	112	43	22
	MSY	–30	–17	–17	–40	–20	–23	–33	–21	–23
3 and 1	RS	15	20	16	**0**	21	24	**–6**	17	12
	RF	–20	–17	–5	–26	–30	–39	–15	–19	–11
	MSY	21	10	5	76	114	188	53	57	33

Note: The data in bold have distinct characteristics from the other figures. The unit is %.

ment model have significantly different effects on the parameter estimates, especially for the non-spatially explicit configurations, i.e., SSC and ACAB (Fig. 3). For the spatially-explicit configuration SCAB, the quality of estimates is generally improved when more CPUE indices from different regions are used. For the two non-spatially explicit configurations, the median relative errors of the estimates of RS or RF based on the combinations of multiple CPUE indices are usually within the interval estimated by using the single CPUE index, but this is not true for MSY because its median relative error may be outside the interval (Fig. 3). Figure 3 also shows for the non-spatially explicit configurations, the performance of the CPUE indices from the core area (i.e., Region 1 or combination of Region 1 and Region 2) is not markedly improved or even worse in comparison with the CPUE indices from non-core areas (i.e., Region 3) or their combinations.

4 Discussion

4.1 *The performance of spatial configurations*

Fish stock often engage in stage-dependent migration to satisfy the particular life history requirements of spatial habitat features (Berger et al., 2012), which usually lead to complex age structures in spatial distribution. In order to improve the quality of fisheries stock assessment, it is very important for analysts to

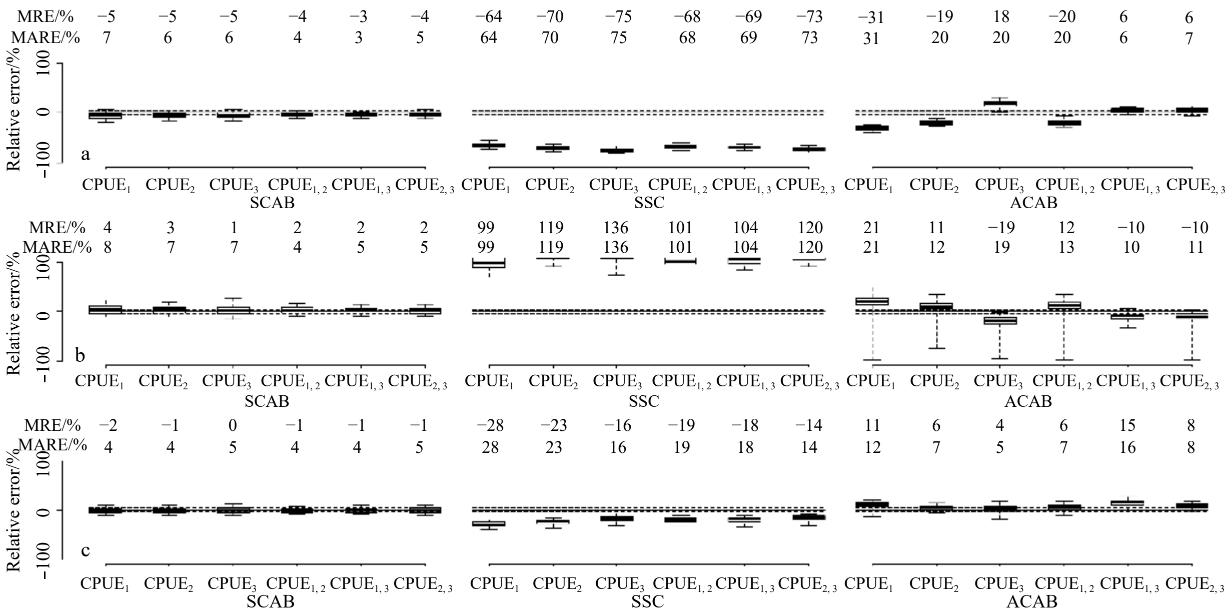

Fig. 3. The relative errors for the three quantities based on different CPUE indices or their combinations. a. The relative error for relative spawning biomass, b. the relative error for relative fishing mortality rate, and c. the relative error for maximum sustainable yield. The band inside the box is the median, the boxes cover 50% of the relative error and the line at the end of the whiskers represents the 95th (upper) or 5th (lower) percentile of the relative errors respectively. The subscripts to CPUE in the figure represent the number of the region. $\mathrm{CPUE}_{1,2}$ indicates that two CPUE indices from Region 1 and Region 2 are used in the stock assessment model, but for SSC, it means the two CPUE indices are amalgamated to generate a new CPUE index.

correctly account for this spatial structure. For example, only SCAB, which is fully consistent with the spatial population dynamic of the operating model, can provide an accurate estimate of the quantities in all scenarios. When the assessment configuration does not exactly match the spatial structure of the operating model, a large uncertainty will be involved in the parameter estimates, such that the quality of estimates of different quantities from the same assessment model cannot be guaranteed to be consistent and the performance of the same assessment configuration in different scenarios may also be different. To correctly account for the spatial structure, we not only need a reliable and spatially explicit stock assessment software package, but also need to accumulate a large number of observed data which can be used to accurately estimate migration and mixing rate (Goethel et al., 2011). Also a considerable body of knowledge from relevant ecological and biological studies is needed, which can be used to correctly delineate the boundaries between regions and simulate spatial population dynamics. At present, spatially explicit stock assessments are rare due to gaps in knowledge and data (Berger et al., 2012; Punt et al., 2015). Therefore, even though spatially explicit stock assessments are used, there are still many problems associated with model convergence (Szuwalski and Punt, 2015) or performance (Punt et al., 2015; Langley and Hoyle, 2016). For example, the performance of the spatially explicit assessment configuration SCIB in this paper is obviously degraded owing to incorrect boundaries, particularly for the selectivity curve SD. Consequently, the performance of the explicit-spatial assessment configuration in the face of making incorrect assumptions can be considerably poorer, such that a simpler assessment configurations may be preferred (Punt et al., 2015).

The performance of SSC, except for Scenario 4 where the fishing effort is homogenously distributed across regions, is generally poor, in particular for the estimates of relative spawning biomass and relative fishing mortality rate. Compared with SSC, the estimates of relative spawning biomass and relative fishing mortality rate are slightly improved for BRSCAB or BRSCIB, but the estimates of MSY are worse. The poor performance of BRSCAB in this paper is very different from that in Punt et al. (2015) where the configuration performed better. For the areas-as-fleets method (i.e., ABAC or ABIC), although the estimate of MSY is relatively inaccurate and the quality of its estimate varies considerably among scenarios, the estimates of relative spawning biomass and relative fishing mortality rate are relatively accurate, particularly when the selectivity curve ST is used. In terms of the performance of the five non-spatially explicit assessment configurations, the areas-as-fleets methods are best. Thus, to some extent, the areas-as-fleets approach can be used to reduce the impacts of spatial structure on the parameter estimates in stock assessment (Hurtado-Ferro et al., 2014). However, as shown in this paper, these approaches should be applied with caution because they are unable to completely eliminate bias and their performances are still case-specific and unstable (Hurtado-Ferro et al., 2014; Punt et al., 2015).

4.2 *Effects of incorrect boundaries*

Comparing the parameter estimates of the three quantities from SCAB and SCIB, BRSCAB and BRSCIB, or ACAB and ACIB, the poor performance of SCIB is directly caused by the incorrect boundaries, but for ACIB the incorrect boundaries have very slight negative effects on the performance of the parameter estimation, and for BRSCIB the effects are actually positive. Therefore, for the spatially explicit assessment configuration, the correct boundaries may be important, but for the non-spatially explicit assessment configuration, whether or not the boundary is correct is not the key to the performance of the assessment configuration. It may be more important for analysts to partition the area to properly reflect the transition in field data and to effectively account for the impacts of ignoring the spatial structure by using the additional spatially referenced parameters.

4.3 *Performance of the selectivity curves*

Selectivity is an important factor for estimation performance in stock assessment. The selectivity defined in stock assessments is generally determined by gear selectivity and availability of the fish (Punt et al., 2014). Consequently, as well as gear selectivity, the shape of the selectivity is also affected by the assumption of spatial population dynamics in the stock assessment. The selectivity in stock assessment may be stable and identical to the gear selectivity (e.g., for SCAB) or time-variant and different from that gear selectivity due to differences in the partitioning of the region, the regional fishing mortality rate, the age-specific movement rate, and the regional distribution of the recruits (Waterhouse et al., 2014). Time-varying selectivity curves are increasingly being used in stock assessment to account for the impacts of ignoring spatial structure in the population dynamics and to improve the quality of parameter estimates (Punt et al., 2014). However, except for the estimate of relative spawning biomass and relative fishing mortality rate from ACAB and ACIB, the advantage of a time-varying selectivity curve is not pronounced. Moreover, those estimates based on a time-varying selectivity curve are not always accurate or the best among the three selectivity curves. For example, the performance of the time-varying selectivity curve is the worst for ACAB and ACIB in Scenario 3. The reason for this may be because the time-varying selectivity curve has not provided sufficient flexibility to match the varying shape of a selectivity generated by the assessment configuration which differs from the operating model (Waterhouse et al., 2014). On the other hand, when a statistical catch-at-age model is used in stock assessment, three types of data, i.e., indices of the relative abundance, catch, and age or length-composition data, are usually needed and primacy will be given to the indices of relative abundance which means that more attention should be devoted to ensuring that the indices of relative abundance are well fitted by the stock assessment model (Francis, 2011). Accordingly, even though the selectivity curve is sufficiently flexible to keep the catch and age-composition data well fitted, as in Waterhouse et al. (2014), the difference between observed CPUE indices and that predicted by assessment models may also produce biased and imprecise results, due to different selectivity and availability of fish in the operating model and assessment model. Moreover, the regular change of median relative errors with movement rates for the three quantities may also imply that the performance of selectivity curves is case-specific (Table 5). Consequently, this is an alternative approach to using the time-varying selectivity curves to account for model misspecification causing the lack of fit, but it is also unable to completely eliminate bias as it ignores spatial structure.

In addition, the convergence problem of the spatially explicit stock assessment model with time-varying selectivity curve is likely to be influenced by the potentially confounded process where the age-specific movement rates and time-varying selectivity will simultaneously affect the estimates of age composition data and their inter-annual variability.

4.4 *Effects of the choice of CPUE indices*

If the assessment configurations exactly match the operating model, the performance of parameter estimation is better when more CPUE indices from different regions are used. Otherwise, it is difficult to determine which regional CPUE indices or their combinations can provide better parameter estimates, because the information of all the predicted CPUE indices in a stock assessment model are generally further distorted from the observed CPUE indices given the model misspecification caused by ignoring the spatial structure. At the same time, the lack of fit for CPUE indices resulted from model misspecification is hard to be evaluated and is always ignored. Nevertheless, as shown in this paper, the worst estimates can generally be avoided by using the combinations of different regional CPUE indices. Although it is common practice in stock assessment to use the CPUE indices from major fishing areas or exclude some CPUE indices to prevent a conflicting trend with the CPUE indices from major fishing area (Guan et al., 2016b; Langley and Holey, 2016), the results in this paper indicate their performance of the parameter estimation may not necessarily be improved. For example in recent years, although the CPUE indices from Region 3 have a different trend with the CPUE indices from Region 1 (i.e., major fishing area) (Fig. 4), the performance of the CPUE indices from Region 3 is better than that from Region 1 for spatial configuration ACAB (Fig. 3).

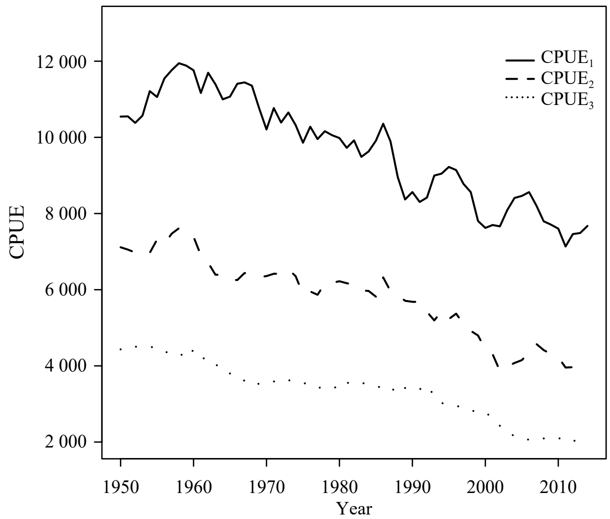

Fig. 4. The time-trajectory of the median CPUE indices for the 120 simulations in Scenario 1. Subscript of CPUE in the figure is the number of the region. The CPUE indices do not include observation errors.

4.5 *Model diagnosis and selection*

As shown in this paper, when a range of assessment configurations are used in stock assessment, some assessment configurations may have good performance. Consequently, model diagnosis methods such as the R_0 likelihood profile, residual analysis, and age-structured production model (Carvalho et al., 2017) or model selection methods such as the Akaike Information Criterion (AIC; Akaike, 1973), Bayes Information Criterion (BIC; Schwarz, 1978), and Deviance Information Criterion (DIC; Spiegelhalter et al., 2002), may provide a good opportunity to select the appropriate models (Helu, 2000; Punt et al., 2014). However, owing to computing complexity and time limitation, these methods are not applied in this paper. In practice, it may be easy to exclude some unreasonable assessment models by using model diagnosis or model selection methods, for example, the serious overestimation of MSY can be easily diagnosed as an unreasonable estimate when the non-spatially explicit configurations BRSCAB and BRSCIB are used. However, the efficacy of the model diagnosis or model selection depend on many factors such as data weighting (Punt et al., 2014) or model misspecification (Carvalho et al., 2017) and these methods may provide unreasonable results (Punt et al., 2014, 2015; Carvalho et al., 2017). On the other hand, even in the same assessment model, the quality of estimate of different quantities may be different as shown in this paper. Consequently, the accuracy of all parameter estimates based on the best model selected by using these methods may not be guaranteed. This is usually ignored in practice and may introduce risk in fishery management.

4.6 *Case-specific simulation studies*

In this paper, the performance of the areas-as-fleets methods is best among the five non-spatially explicit configurations, which is different from the results in Punt et al. (2015) where the performance of BZSA (i.e., BRSCAB in this paper) is better. The different conclusions from the two studies can be attributed to case-specific simulation studies. In this paper, the spatial structure relates to stage-dependent migration and region-dependent fishing mortality rate and differs from the spatial heterogeneity resulted from region-dependent fishing mortality rate, growth and recruitment in Punt et al. (2015). In addition, the number of simulation scenarios and assessment configurations are limited due to a limitation of time in this study. For example, only one kind of time-varying selectivity is used and evaluated, but the performance of other time-varying selectivities, such as the time-blocking selectivity, are not evaluated. Moreover, in this paper, some parameters such as steepness and nature mortality rate are assumed to be known exactly and better data, which may not be expected in real-world stock assessment, are generated to clearly present the difference in performance that can be ascribed to model misspecification induced by spatial structure (Lee et al., 2017). All of these may have some effects on performance estimation (Punt et al., 2015; Lee et al., 2017). As a result, simulation studies can provide important insights into specific stock assessment (Punt et al., 2014), but the conclusions may be case-specific and care should be taken when extending them to other situations.

5 Summary

Only a stock assessment model which can correctly simulate the spatial population dynamics can provide unbiased and precise estimates of quantities of interest, otherwise estimates are likely to be biased and imprecise. Although model diagnosis and model selection can be used to select an appropriate assessment model and improve the quality of parameter estimates, there are still large uncertainties in these methods (Punt et al., 2014, 2015; Carvalho et al., 2017). Therefore, management strategy evaluation may be an important approach to identify an appropriate management strategy to deal with the inevitable uncertainties associated with these fisheries assessments (Butterworth, 2007; Punt et al., 2016). In future, more powerful stock assessment models, for example, a coupled ocean-biogeochemical-populations dynamics model with high temporal and spatial resolution similar to SEAPODYM (Spatial Ecosystem and Population Dynamics Model), may be developed to provide a better way of effectively avoiding the impacts of spatial structure (Lehodey et al., 2008; Senina et al., 2008; Goethel et al., 2011).

References

Akaike H. 1973. Information theory and an extension of the maximum likelihood principle. In: Proceedings of the 2nd International Symposium on Information Theory. Budapest: Academiai Kiado, 267–281

Berger A M, Jones M L, Zhao Yingming, et al. 2012. Accounting for spatial population structure at scales relevant to life history improves stock assessment: The case for Lake Erie walleye *Sander vitreus*. Fisheries Research, 115–116: 44–59, doi: 10.1016/j.fishres.2011.11.006

Butterworth D S. 2007. Why a management procedure approach? Some positives and negatives. ICES Journal of Marine Science, 64(4): 613–617, doi: 10.1093/icesjms/fsm003

Carvalho F, Punt A E, Chang Y J, et al. 2017. Can diagnostic tests help identify model misspecification in integrated stock assessments?. Fisheries Research, 192: 28–40, doi: 10.1016/j.fishres.2016.09.018

Chen I C, Lee P F, Tzeng W N. 2005. Distribution of albacore (*Thunnus alalunga*) in the Indian Ocean and its relation to environmental factors. Fisheries Oceanography, 14(1): 71–80, doi: 10.1111/j.1365-2419.2004.00322.x

Cope J M, Punt A E. 2011. Reconciling stock assessment and management scales under conditions of spatially varying catch histories. Fisheries Research, 107(1-3): 22–38, doi: 10.1016/j.fishres.2010.10.002

Farley J H, Williams A J, Davies C R, et al. 2012. Population biology of albacore tuna in the Australian region. Hobar, Tasmania: CSIRO Marine and Atmospheric Research

Fornteneau A. 2016. Indian Ocean albacore stock: review of its fishery, biological data and results of its 2014 stock assessment. Shanghai, China: Working Party on Temperate Tuna

Fournier D A, Hampton J, Sibert J R. 1998. MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. Canadian Journal of Fisheries and Aquatic Sciences, 55(9): 2105–2116, doi: 10.1139/f98-100

Francis R I C C. 2011. Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences, 68(6): 1124–1138, doi: 10.1139/f2011-025

Goethel D R, Quinn T J, Cadrin S X. 2011. Incorporating spatial structure in stock assessment: Movement modeling in marine fish population dynamics. Reviews in Fisheries Science, 19(2): 119–136, doi: 10.1080/10641262.2011.557451

Guan Wenjiang, Cao Jie, Chen Yong, et al. 2013. Impacts of population and fishery spatial structures on fishery stock assessment. Canadian Journal of Fisheries and Aquatic Sciences, 70(8): 1178–1189, doi: 10.1139/cjfas-2012-0364

Guan Wenjiang, Tang Lin, Zhu Jiangfeng, et al. 2016a. Application of a Bayesian method to data-poor stock assessment by using Indian Ocean albacore (*Thunnus alalunga*) stock assessment as an example. Acta Oceanologica Sinica, 35(2): 117–125, doi: 10.1007/s13131-016-0814-0

Guan Wenjiang, Zhu Jiangfeng, Xu Liuxiong, et al. 2016b. Using a Bayesian biomass dynamics model to assess Indian Ocean albacore (*Thunnus alalunga*). Shanghai, China: Working Party Temperature Tuna

Helu S L, Sampson D B, Yin Yanshui. 2000. Application of statistical model selection criteria to the Stock Synthesis assessment program. Canadian Journal of Fisheries and Aquatic Sciences, 57(9): 1784–1793, doi: 10.1139/f00-137

Hoyle S D, Sharma R, Herrera M. 2014. Stock assessment of albacore tuna in the Indian Ocean for 2014 using Stock Synthesis. Busan, Korea: Working Party on Temperate Tuna

Hurtado-Ferro F, Punt A E, Hill K T. 2014. Use of multiple selectivity patterns as a proxy for spatial structure. Fisheries Research, 158: 102–115, doi: 10.1016/j.fishres.2013.10.001

Langley A, Hoyle S. 2016. Stock assessment of albacore tuna in the Indian Ocean using stock synthesis. Shanghai, China: Working Party on Temperate Tuna

Lee H H, Piner K R, Maunder M N, et al. 2017. Evaluation of alternative modelling approaches to account for spatial effects due to age-based movement. Canadian Journal of fisheries and Aquatic Sciences, 74(11): 1832–1844, doi: 10.1139/cjfas-2016-0294

Lehodey P, Senina I, Murtugudde R. 2008. A spatial ecosystem and populations dynamics model (SEAPODYM) -Modeling of tuna and tuna-like populations. Progress in Oceanography, 78(4): 304–318, doi: 10.1016/j.pocean.2008.06.004

Matsumoto T. 2016. Stock and risk assessments of albacore in the Indian Ocean based on ASPIC. Shanghai, China: Working Party on Temperature Tuna

McAllister M K, Ianelli J N. 1997. Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm. Canadian Journal of Fisheries and Aquatic Sciences, 54(2): 284–300

Methot R D Jr, Wetzel C R. 2013. Stock Synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fisheries Research, 142: 86–99, doi: 10.1016/j.fishres.2012.10.012

Nikolic N, Fonteneau A, Hoarau L, et al. 2014. Short review on biology, structure, and migration of *Thunnus alalunga* in the Indian Ocean. Busan, Korea: Working Party on Temperature Tuna

Nishida T, Kitakado T, Matsumoto T, et al. 2014. Consideration of albacore parameters for stock assessments in the Indian Ocean. Busan, Korea: Working Party on Temperature Tuna

Nishida T, Sato K, Chang Y, et al. 2016. Stock assessment of albacore *Thunnus alalunga* in the Indian Ocean using Statistical-Catch-At-Age (SCAA). Shanghai, China: Working Party on Temperature Tuna

Ono K, Licandeo R R, Muradian M L, et al. 2015. The importance of length and age composition data in statistical age-structured models for marine species. ICES Journal of Marine Science, 72(1): 31–43, doi: 10.1093/icesjms/fsu007

Penney A J. 1994. Morphometric relationships, annual catches and catch-at-size for South African caught south Atlantic albacore (*Thunnus alalunga*). Collective Volume of Scientific Papers, 42(1): 371–382

Punt A E, Butterworth D S, de Moor C L, et al. 2016. Management strategy evaluation: best practices. Fish and Fisheries, 17(2): 303–334, doi: 10.1111/faf.12104

Punt A E, Haddon M, Tuck G N. 2015. Which assessment configurations perform best in the face of spatial heterogeneity in fishing mortality, growth and recruitment? A case study based on pink ling in Australia. Fisheries Research, 168: 85–99, doi: 10.1016/j.fishres.2015.04.002

Punt A E, Hurtado-Ferro F, Whitten A R. 2014. Model selection for selectivity in fisheries stock assessments. Fisheries Research, 158: 124–134, doi: 10.1016/j.fishres.2013.06.003

Schwarz G. 1978. Estimating the dimension of a model. The Annals of Statistics, 6(2): 461–464, doi: 10.1214/aos/1176344136

Senina I, Sibert J, Lehodey P. 2008. Parameter estimation for basin-scale ecosystem-linked population models of large pelagic predators: application to skipjack tuna. Progress in Oceanography, 78(4): 319–335, doi: 10.1016/j.pocean.2008.06.003

Spiegelhalter D J, Best N G, Carlin B P, et al. 2002. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society Series B-Statistical Methodology, 64(4): 583–639, doi: 10.1111/1467-9868.00353

Stephenson R L, Melvin G D, Power M J. 2009. Population integrity and connectivity in Northwest Atlantic herring: a review of assumptions and evidence. ICES Journal of Marine Science, 66(8): 1733–1739, doi: 10.1093/icesjms/fsp189

Szuwalski C S, Punt A E. 2015. Can an aggregate assessment reflect the dynamics of a spatially structured stock? Snow crab in the eastern Bering Sea as a case study. Fisheries Research, 164: 135–142, doi: 10.1016/j.fishres.2014.10.020

Waterhouse L, Sampson D B, Maunder M, et al. 2014. Using areas-as-fleets selectivity to model spatial fishing: asymptotic curves are unlikely under equilibrium conditions. Fisheries Research, 158: 15–25, doi: 10.1016/j.fishres.2014.01.009

Wells R J D, Kohin S, Teo S L H, et al. 2013. Age and growth of North Pacific albacore (*Thunnus alalunga*): implications for stock assessment. Fisheries Research, 147: 55–62, doi: 10.1016/j.fishres.2013.05.001

Ying Yiping, Chen Yong, Lin Longshan, et al. 2011. Risks of ignoring fish population spatial structure in fisheries management. Canadian Journal of Fisheries and Aquatic Sciences, 68(12): 2101–2120, doi: 10.1139/f2011-116