



# Parameter estimation for basin-scale ecosystem-linked population models of large pelagic predators: Application to skipjack tuna

Inna Senina<sup>b,\*</sup>, John Sibert<sup>a</sup>, Patrick Lehodey<sup>b</sup>

<sup>a</sup> Pelagic Fisheries Research Program, University of Hawaii at Manoa, 1000 Pope Road, Honolulu, HI 96822, United States

<sup>b</sup> Marine Ecosystems Modelling and Monitoring by Satellites, CLS, Satellite Oceanography Division, 8-10 rue Hermes, 31520 Ramonville, France

## ARTICLE INFO

### Article history:

Received 26 October 2007

Received in revised form 2 June 2008

Accepted 29 June 2008

Available online 22 July 2008

### Keywords:

Advection–diffusion–reaction equations

Optimization

Adjoint

Parameter estimation

Skipjack tuna

## ABSTRACT

A Spatial Ecosystem and Population Dynamic Model (SEAPODYM) is used in a data assimilation study aiming to estimate model parameters that describe dynamics of Pacific skipjack tuna population on ocean-based scale. The model based on advection–diffusion–reaction equations explicitly predicts spatial dynamics of large pelagic predators, while taking into account data on several mid-trophic level components, oceanic primary productivity and physical environment. In order to improve its quantitative ability, the model was parameterized through assimilation with commercial fisheries data, and optimization was carried out using maximum likelihood estimation approach. To address the optimization task we implemented an adjoint technique to obtain an exact, analytical evaluation of the likelihood gradient. We conducted a series of computer experiments in order to (i) determine model sensitivity with respect to variable parameters and, hence, investigate their observability; (ii) estimate observable parameters and their errors; and (iii) justify the reliability of the computed solution. Parameters describing recruitment, movement, habitat preferences, natural and fishing mortality of skipjack population were analysed and estimated. Results of the study suggest that SEAPODYM with achieved parameterization scheme can help to investigate the impact of fishing under various management scenarios, and also conduct forecasts of a given species stock and spatial dynamics in a context of environmental and climate changes.

© 2008 Elsevier Ltd. All rights reserved.

## 1. Introduction

High rates of exploitation of tuna populations during the last 20 years have led to widespread concern over the status of tuna populations (Sibert et al., 2006). Tunas are very mobile animals that are classified as “highly-migratory” under international law. They typically occupy entire ocean basins, but their populations are not uniformly distributed nor are tuna fisheries uniform in space and time. Movements on all scales are mediated by environmental conditions. Therefore, models of tuna populations must include population temporal and spatial dynamics, and dependency dynamics on environmental forcing for application to fishery management (Sibert and Hampton, 2003).

Continuous advection–diffusion–reaction equations (hereafter ADRs) provide excellent tools for studying the influence of spatial structure of the environment and spatial behavior of individuals on overall population dynamics (Berezovskaya et al., 1999; Govorukhin et al., 2000; Petrovskii and Li, 2001). They have been successfully used for explicit descriptions of population spatial dynamics since early 1950s (Skellam, 1951; Okubo, 1980; Edelstein-Keshet, 1988; Murray, 1989; Czaran, 1998; Turchin,

1998). ADRs enable investigation of a range of biological phenomena on different scales, such as occurrence of spatial patchiness due to chemotaxis (Keller and Segel, 1971; Berezovskaya and Karev, 1999), biological invasions (Petrovskii et al., 2002), schooling and shoaling (Grunbaum, 1994; Tyutyunov et al., 2004) based on attraction and repulsion between organisms. When applied to trophic systems ADRs can be used to take into account individual interactions, which produce non-linear functional response (Arditi and Ginzburg, 1989) on a macroscopic scale (Arditi et al., 2001).

Describing the spatial dynamics of fish population at large scales (i.e., ocean basin) is of paramount importance for fisheries management, so as to understand and predict the consequences of fishing, climate change and changes in fishery management regulations. However, before applying the model to solve fishery management problems we need to be confident in the reliability of model predictions. Hence, we first need to work on improving the model dynamics and initial conditions under the tight coupling of the model predictions to observations. This is referred to as “data assimilation” in the field of quantitative ecosystem modelling, the purpose of which is “to provide estimates of nature which are better estimates than can be obtained by using only the observational data or the dynamical model” (Robinson and Lermusiaux, 2002).

\* Corresponding author.

E-mail address: [isenina@cls.fr](mailto:isenina@cls.fr) (I. Senina).

The SEAPODYM model was specifically developed for investigating tuna spatial dynamics linked to the ocean ecosystem (see Bertignac et al., 1998; Lehodey, 2001, 2004a,b; Lehodey et al., 2003). The main features of this model are (i) taking into account the climate variability, which is known to have a strong influence on fish population dynamics and (ii) predicting both temporal and spatial distribution of age-structured populations. Setting the value of each parameter in SEAPODYM has, however, been performed mostly by *ad hoc* manual “tuning”, by using independent models estimates, and by application of parameter values gleaned from the scientific literature. The resultant lack of confidence in parameter values has been a serious barrier to the application of SEAPODYM to practical fishery management problems, such as estimation of fishing impact, testing various restrictions on effort, area and seasons of fishing (Senina et al., 1999).

The purpose of the current study is to test whether model hypotheses and predictions are consistent with observations. This step is prerequisite to application of SEAPODYM as a tool for estimating the broad suite of anthropogenic effects on tuna populations. As a first step, we estimate parameters governing dynamics of skipjack tuna (*Katsuwonus pelamis*). Skipjack tuna is the most abundant tuna species in the Pacific and its contribution to the total tuna catch is very significant (about 80% in Western and Central Pacific Ocean (WCPO) and 72% of tuna catches on en-

tire Pacific during last 15 years). A short life cycle, and the availability of spatially-distributed catch data make skipjack a very convenient species for the initial model validation. Similar model with optimization approach is being developed by Faugeras and Maury (2005), with application to skipjack population in the Indian Ocean.

Although SEAPODYM is a coupled predator–prey model, such coupling is “off-line” in that forage biomass and physical forcing variables are treated as an input data by the predator sub-model. However, despite such simplification the uncoupled model optimization study allowed us to achieve reliable fit between model predictions and observations. The description of the forage sub-model is presented in Lehodey (2004a) and will be therefore omitted from this paper. Herein, we focus on the description of the predator population sub-model and the optimization method. A series of numerical experiments is presented, along with the estimated parameters and their errors.

## 2. The model

The major model compartments are schematically shown in Fig. 1. SEAPODYM incorporates multiple categories of forcing. Predator dynamics are forced by the physical environment, as predicted by an independent Ocean General Circulation Model (see

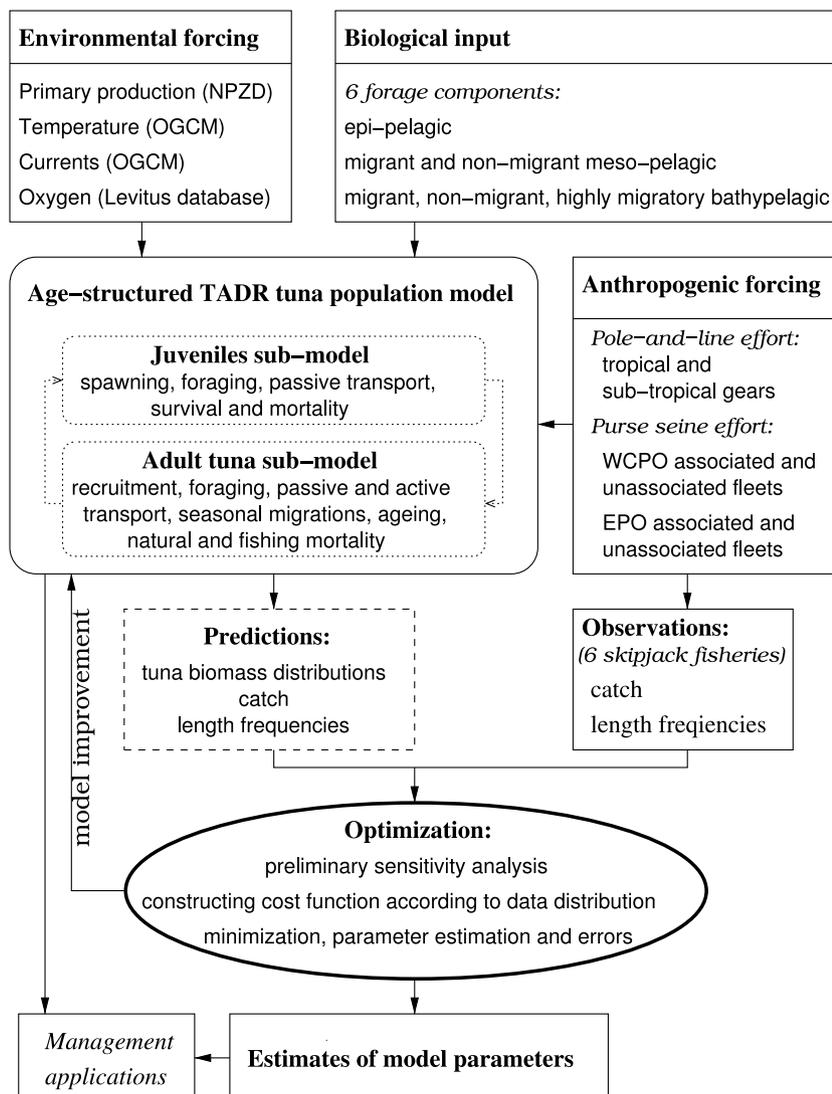


Fig. 1. General scheme of the model with optimization approach.

Chen et al., 1994; Murtugudde et al., 1996). Biogeochemical forcing is the output from the Nutrients–Phytoplankton–Zooplankton–Detritus model (Christian et al., 2002), and biomass–tuna forage components are predicted by SEAPODYM forage sub-model (Lehodey, 2004a).

Several studies on tropical ocean variability and coupled ecosystem variability have reported the model's ability to capture the ocean dynamics and biogeochemical fields at seasonal to inter-annual time scales (Murtugudde et al., 1996; Christian et al., 2002; Christian and Murtugudde, 2003; Wang et al., 2005, 2006).

Predicted fields of physical environmental data are averaged by month over three depth layers: 0–100 m, 100–400 m and 400–1000 m. Let us denote  $\mathbf{v}_z = (u_z, v_z)$ , the vector of oceanic horizontal currents,  $T_z$  is the temperature, and  $O_z$  is the dissolved oxygen at layer  $z = 0, 1, 2$ . Primary production  $P$  is integrated over the depth range 0–400 m and is given in units of  $\text{mmol C m}^{-2} \text{d}^{-1}$ . Anthropogenic forcing is represented by fisheries effort data, grouped by gear type. Effort data are used to parameterize fishing mortality and therefore predicted catch.

The top predator population in SEAPODYM can be structured either by age or life stage. Considering age structure let  $J_k(t, x, y)$ ,  $k = 0, 1, 2$  denote juvenile density so that  $J_0$  is the density of larvae of age 0–1 month, and let  $N_a(t, x, y)$ ,  $a = 1, \dots, K$  be the density of adults tuna of age  $a$  at time  $t \in (0, T)$  and position in two-dimensional space  $(x, y) \in \Omega$ . The maximal index  $K$  depends on the step used for age discretization.

We construct a discrete-continuous system in two-dimensional space, consisting of discrete ageing equations and continuous advection–diffusion–reaction (ADR) equations for describing transport of tuna population. The state variables  $J_k$  and  $N_a$  as well as environmental variables are determined at point  $(x, y)$  and time  $t$  (hereafter we will omit the notations of space and time). For brevity, we use gradient operator  $\nabla = (\partial_x, \partial_y)^T$ , divergence operator of a vector field  $\text{div}(\mathbf{v}) = \partial_x u + \partial_y v$  and  $\Delta = \text{div grad}$  for Laplacian of scalar field of population density.

The ADR system describing dynamics of age-structured tuna population is

$$\partial_t J_k = -\text{div}(J_k \mathbf{v}_0) + \delta \Delta J_k - m_k J_k + S_{J_k}, \quad k = 0, 1, 2; \quad (1)$$

$$\partial_t N_a = -\text{div}(N_a \tilde{\mathbf{v}} + N_a \mathbf{V}_a) + \text{div}(D_a \nabla N_a) - M_a N_a + S_{N_a} \quad a = 1, \dots, K; \quad (2)$$

where  $\delta$  is constant diffusion coefficient of larvae and juveniles;  $m_0 = m_0(x, y) = f(T_0, P, F)$  is larval (age zero) natural mortality rate, a function of water temperature at surface layer  $T_0$ , primary production  $P$  and forage density  $F$ ;  $m_{1,2} = m_{1,2}(x, y) = f(T_0, N)$  are juvenile (ages 1, 2) natural mortality rates dependent on surface layer temperature and total adult tuna density. In Eq. (2)  $\tilde{\mathbf{v}}$  denotes weighted average (by the accessibility to depth layer, see Appendix A for details) of oceanic currents through all layers,  $\mathbf{V}_a$  is vector of directed velocity of adult tuna at age  $a$ , which is proportional to the gradient of the habitat index (Appendix A), diffusion rates  $D_a = D_a(x, y) = D(a, T, O, F)$  are functions of age and environmental factors, and  $M_a = M_a(x, y) = M(a, F, N_a, E_f)$  are total (natural and fishing) mortality of adults. Terms  $S_{J_k}$  and  $S_{N_a}$  represent sources of new population density to corresponding variable and include both survival from younger age classes as well as the effects of spawning and recruitment.

The system ((1) and (2)) is completed by Neumann boundary conditions describing impermeability of the domain bounds  $\partial\Omega$ :

$$\mathbf{n} \cdot \mathbf{v}|_{\mathbf{x} \in \partial\Omega} = \mathbf{n} \cdot \nabla J_k|_{\mathbf{x} \in \partial\Omega} = \mathbf{n} \cdot \nabla N_a|_{\mathbf{x} \in \partial\Omega} = 0. \quad (3)$$

These conditions mean no additional source of biomass and no loss are possible when recruitment and mortality are absent.

Since age discretization in time can be different from the time step chosen to numerically solve the system (1)–(3), ageing equa-

tions were constructed in order to smooth transition from one age group to another. In the present study, we consider 3 monthly juvenile and 16 quarterly adult groups. The system (1)–(3) is supplemented by the discrete equations implying that tuna survivors of each age class are computed as a number of individuals remaining in the age class at a current time step plus recruits from younger age, minus the number of individuals which pass to the older age group. Thus, we have simple relationships:

$$J_k^{t+1} = q_{n,k-1} J_{k-1}^t + (1 - q_{n,k}) J_k^t, \quad k = 1, 2; \quad (4)$$

$$N_1^{t+1} = q_{n,2} J_2^t + (1 - p_{n,1}) N_1^t; \quad (5)$$

$$N_a^{t+1} = p_{n,a-1} N_{a-1}^t + (1 - p_{n,a}) N_a^t, \quad a = 2, \dots, K. \quad (6)$$

The survival coefficients  $q_{n,\cdot}$  and  $p_{n,\cdot}$  determine the rates of decay of the density due to natural, predation and fishing mortality depending on the time spent in corresponding age. They are relative values between 0 and 1, such that

$$q_{n,k} = \frac{e^{-nm_k}}{\sum_{i=1}^n e^{-im_k}}; \quad (7)$$

$$p_{n,a} = \frac{e^{-nM_a}}{\sum_{i=1}^n e^{-iM_a}};$$

where  $n$  is the ratio between time step in population age structure and the time step of discretization in numerical approximation of (1)–(3), i.e.,  $\Delta t = n \Delta t$  (month). Note that if  $n = 1$ , i.e., age discretization coincides with time discretization, then  $q_{1,k} = 1$  and the corresponding equation, i.e., Eq. (4) simplifies to  $J_k^{t+1} = J_{k-1}^t$ .

A detailed description of functional links between fish population dynamics and physical environmental processes is presented in Appendix A, along with the biological meaning of each estimated parameter. A list of mathematical symbols can be found in Appendix B.

### 3. Numerical simulations

Numerical approximations of system (1)–(3) and discrete ageing relationships ((4)–(6)) define the simulation model of tuna population dynamics. The partial derivatives of ADRs ((1) and (2)) are approximated by second order finite differences with upwind differencing of advective terms (see discretization scheme in Sibert et al., 1999). Although a non-optimized version of SEAPODYM runs on the spatial domain of the entire Pacific ocean, in the optimization study we restricted the computational domain to  $\Omega = \{x \in (99^\circ\text{E}, 69^\circ\text{W}), y \in (45^\circ\text{N}, 39^\circ\text{S})\}$  since this is the region where skipjack catches have been recorded during the 1950–2005 period (see Fig. 2). We assume that skipjack abundances are very low outside the area. Boundary conditions (Eq. (3)) are implemented in the discretization scheme. The complex boundary of the domain is presented by the land mask generated from ocean floor topography (ETOPO2 – 2-min worldwide bathymetric/topographic data). A regular square grid is used, with  $\Delta x = \Delta y = 120$  Nmi (nautical miles,  $120 \text{ Nmi} \approx 222.24 \text{ km}$ ). The resulting algebraic problem is solved using the alternate direction implicit (ADI) method, with a time step of  $\Delta t = 1$  (month) for (1)–(3).

Initial conditions are generated by the following “spin-up” process: starting from uniform zero spatial distribution, the population density is modelled by Eqs. (1)–(6) which are forced by the climatological environment (generated over the period 1948–2005) during first  $3 \Delta \tau_0 + K \Delta \tau_a$  time steps. Every month, a new larval source is computed using the temperature only (i.e.,  $S_{J_0} = R\Phi(T_0)$ , see A.2). The duration of the spin-up was set up to match the total lifespan of skipjack population, so that the density of each age class can be computed at the end of the climatological run. After spin-up, the

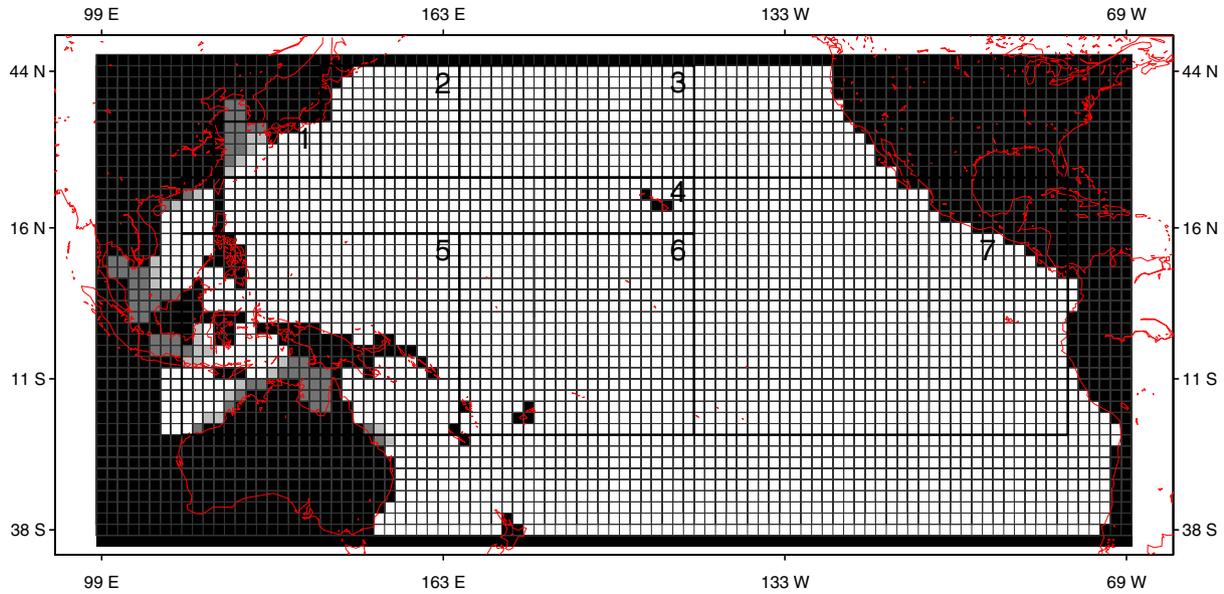


Fig. 2. Computational domain with 3-layer mask (black color – land; dark gray – upper layer with depth < 100 m; light gray – layer 100–400 m) and 2° grid. Numbered regions 1–7 are the regions used in the MULTIFAN-CL stock assessment.

simulation continues with actual forcing fields for another 2 years in order to reduce the influence of initial climatological forcing. Final distributions were saved for later use as initial conditions for optimization experiments. As the initial distributions play an important role in parameter estimation process, we then repeated the same procedure several times re-generating initial state of the model using the optimized parameters.

4. The optimization approach

SEAPODYM explicitly describes spatial dynamics of pelagic fish populations influenced not only by intrinsic population dynamics processes, but also by extrinsic environmental variability. Predictions of the model strongly depend on environmental forcing which is taken by the model as the input. Use of oceanographic data allows the model to reflect the impact of ocean temperature, currents and primary production anomalies associated with El Niño or La Niña events, which results in changes of population abundance and distribution - from juveniles to adults (see Lehodey et al., 1997). However, in order to have confidence in the model predictions, particularly the adequacy of the population dynamical responses to environmental variability, we need to combine simulations with quantitative optimization (see the lower part of Fig. 1) and to express how well the model describes the observational data.

4.1. Fisheries data

Historical data collected from multiple fisheries operating in the Western and Central Pacific ocean were provided by the Secretariat of the Pacific Community (SPC), and data available for Eastern Pacific ocean are supplied by Inter-American Tropical Tuna Commission. Monthly spatially distributed data on fishing effort  $E_{t,ij}$  (in days) and catch  $C_{t,ij}^{obs}$  (in tonnes) are aggregated into six generalized gear types or “fisheries” defined by unique values of the “catchability coefficient”,  $q_f$  with  $f = 1, 2, \dots, 6$  corresponding to four WCPO (PLSUB, PLTRO, WPSASS, WPSUNA) and two EPO fisheries (EPSASS and EPSUNA). Summarized distribution of catch by these six fisheries is shown on Fig. 3. Seasonal size composition of the catch is available for each fishery aggregated over seven spatial regions in Fig. 2.

4.2. Model predictions

The predicted catch,  $C_{t,ij}^{pred}$ , at time  $t$  for fishery  $f$  is computed in the model using observed fishing effort  $E_{t,ij}$  at location  $(i, j)$  by

$$C_{t,ij}^{pred} = q_f E_{t,ij} \sum_{a=1}^K s_{fa} w_a N_{a,ij} \Delta x \Delta y,$$

where  $w_a$  is the mean weight of fish in the  $a$ th cohort.

The predicted proportion at age  $a$  in the catch at time  $t$  for fishery  $f$  in region  $r$  is

$$Q_{t,fa}^{pred} = \frac{s_{fa} \sum_{i,j \in r} E_{t,ij} N_{a,ij} \Delta x \Delta y}{\sum_{a=1}^K s_{fa} \sum_{i,j \in r} E_{t,ij} N_{a,ij} \Delta x \Delta y}.$$

4.3. Likelihood function

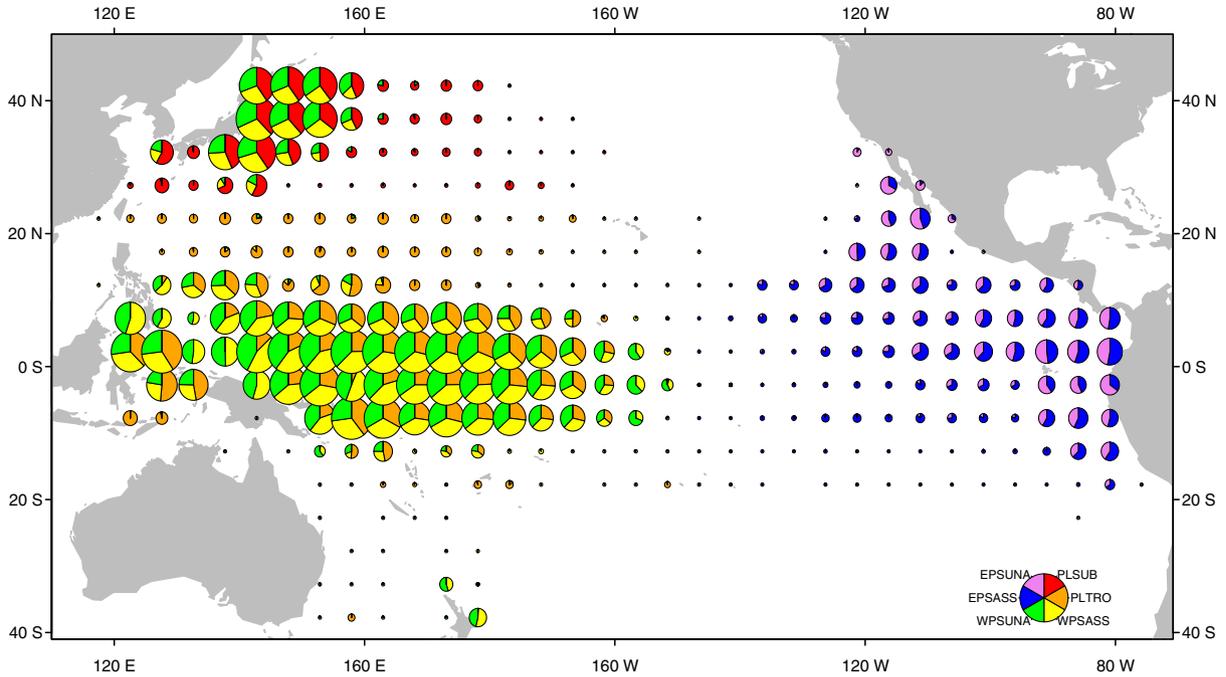
We use the maximum likelihood method to estimate model parameters  $\theta_k$  that would allow the model predictions to approach observations. In the WCPO most of the skipjack catch is made by purse-seine fleets targeting skipjack, and we assume that spatially scattered catch data have a Poisson distribution. Skipjack catch data from Eastern Pacific Ocean (EPO) fisheries contain many zeros since fleets target mostly yellowfin tuna, and we assume these data follow a negative binomial distribution with zero inflation. This yields the following likelihood components for WCPO fisheries:

$$L_1(\theta | C^{obs}) = \prod_{t,ij} \frac{C_{t,ij}^{pred} C_{t,ij}^{obs} e^{-C_{t,ij}^{pred}}}{C_{t,ij}^{obs}!}, \quad f = 1, 2, 3, 4 \tag{8}$$

and for EPO they are correspondingly

$$L_2(\theta | C^{obs}) = \begin{cases} \prod_{t,ij} \left( p_f + (1 - p_f) \left( \frac{\beta_f}{1 + \beta_f} \right)^{\frac{\beta_f C_{t,ij}^{pred}}{1 - p_f}} \right) & \text{if } C_{t,ij}^{obs} = 0, \\ \prod_{t,ij} \left( (1 - p_f) \frac{\Gamma \left( \frac{C_{t,ij}^{obs} + \beta_f C_{t,ij}^{pred}}{1 - p_f} \right)}{\Gamma \left( \frac{\beta_f C_{t,ij}^{pred}}{1 - p_f} \right) C_{t,ij}^{obs}!} \left( \frac{\beta_f}{1 + \beta_f} \right)^{\frac{\beta_f C_{t,ij}^{pred}}{1 - p_f}} \left( \frac{1}{1 + \beta_f} \right)^{C_{t,ij}^{obs}} \right) & \text{if } C_{t,ij}^{obs} > 0, \end{cases} \tag{9}$$

$f = 5, 6;$



**Fig. 3.** Summarized over simulation period, distribution of catch data for six fisheries used in SEAPODYM (PLSUB and PLTRO are pole-and-line sub-tropical and tropical fleets correspondingly, WPSASS and WPSUNA are western purse-seine fisheries associated or not with any drifting objects and analogously the eastern purse-seine fleets are abbreviated as EPSASS and EPSUNA). Note that each pie centered in 5° square shows proportions of catch by fishery operating in corresponding area. Size of pies correspond to log-scaled cumulative catch.

where the parameters  $\beta_f$  and  $p_f$  are the negative binomial parameters (showing how much variance exceeds expected value) and probability of getting a null observation, respectively. Both are estimated in the optimization process.

We assume that fish lengths at catch are normally distributed, which gives the following contribution from length frequency data to the negative log-likelihood:

$$-L_3(\theta|Q^{obs}) = \frac{1}{2\sigma_Q^2} \sum_{t=1}^T \sum_{f=1}^6 \sum_{a=1}^K \sum_{r=1}^7 (Q_{tfar}^{pred} - Q_{tfar}^{obs})^2, \quad (10)$$

where the proportion at age  $a$  in the catch computed from the observed length-frequency data is  $Q_{tfar}^{obs} = \frac{N_{tfr}^{obs}}{\sum_r N_{tfr}^{obs}}$ ,  $l \in [l_a, \bar{l}_a]$ .  $N_{tfr}^{obs}$  is the number of fish of length that belongs to the cohort of age  $a$ . The variance  $\sigma_Q^2$  is fixed at the constant value of 1.5 cm.

The negative log-likelihood function,  $L^- = -\ln(L)$ , to be minimized is thus the sum of the three components given in (Eqs. (8)–(10)), i.e.,

$$L^- = -\ln L_1(\theta|C^{obs}) - \ln L_2(\theta|C^{obs}) - L_3(\theta|Q^{obs}). \quad (11)$$

Instead of setting penalties to the boundaries of  $\theta$ , we chose to perform a constrained minimization through parameter scaling (see e.g., Bard, 1974; Vallino, 2000). The latter implies that the optimization routine operates in the unbounded parametric space that is mapped to the bounded one with the transformation  $\theta_k = \frac{\theta_k + (\bar{\theta}_k - \theta_k)(1 + \sin \frac{\pi \theta'_k}{2})}{2}$ , i.e., variable  $\theta'_k$  can vary from  $-\infty$  to  $\infty$  while  $\theta_k$  remains within the imposed bounds.

#### 4.4. Implementation of adjoint method

Solving the problem of function minimization with gradient methods, requires evaluation of the derivatives of objective function with respect to each control parameter. We evaluated these derivatives using the method of integration of the adjoint model.

First, we attempted to use automatic differentiation for deriving adjoint computer code, however for SEAPODYM it gave unrealistic demands of computer memory needed for storing the model variables during forward computation (see e.g., Griewank and Corliss, 1991). As result all computer codes for the adjoint model were manually written and then tested using utilities of automatic code differentiation library AUTODIF (Otter Research Ltd., 1994). This library also provided quasi-Newton numerical function minimizer and parameter scaling algorithms (see above).

The resulting analytic derivatives were verified in two ways. First, they were simply compared to derivatives computed by AUTODIF. Second, it was verified that

$$\frac{L^-(\theta_k + h) - L^-(\theta_k - h)}{2h} - \nabla_k L^- = O(h^2),$$

i.e., that the discrepancy between each gradient component  $\nabla_k L^-$  obtained by analytic differentiation and its finite difference approximation changes parabolically with step  $h$  (which was varied from  $10^{-6}$  to  $10^{-1}$ ).

#### 4.5. Optimization experiments

Three optimization experiments spanned three different time periods: 1980–1990 (E1), 1990–2005 (E2), and the total 25-year range from 1980 to 2005 (E3) to determine the extent to which the estimated parameters depend on data from different periods, and consequently, how model predictions change as a result of assimilating data over different time periods.

In order to reduce the influence of initial conditions (spatial distributions of cohorts produced by the model with initial guess parameters) on the results of minimization procedure, the first 6-month predictions in each optimization experiment were excluded from the likelihood function.

As initial guess parameters of the mortality and recruitment functions we used estimates from recently published skipjack stock assessments for the WCPO using MULTIFAN-CL (Langley

et al., 2005). See the pre-specified values of model parameters in Table 1.

Preliminary optimization experiments revealed several problems. As expected, the simultaneous estimation of the parameters of the recruitment function (Eq. A.11) and the natural and fishing mortality functions (Eqs. (A.12)–(A.14)) lead to biased estimates of the total population size. Namely, if we try to estimate simultaneously parameters  $R$ ,  $b$ ,  $\beta_p$  and  $\beta_s$ , the minimization procedure tends to increase the total stock and also the computational time (i.e., the number of iterations) increases by a factor of more than three compared to experiments in which  $R$  is fixed. Such drastic increase of the number of iterations due to the release of one more parameter is an indication of strong correlation with other parameters. Considering the importance of the estimation the mortality rate for each cohort given its functional form over ages, and taking into account availability of more information for the adult cohort in the fishing data than for larvae stage, we have chosen to fix the recruitment parameters (see Table 1).

We further noticed that the coarse spatial resolution ( $2^\circ$ ) of the grid being used and a bias in the predicted seasonal peak of temperature in the forcing field for the Kuroshio region do not allow the model to resolve the fine scale habitat variability that determines seasonal population migrations in this area. Therefore with the current grid resolution SEAPODYM cannot adequately describe the seasonal skipjack migration through the Kuroshio extension (Nihira, 1996) and attempts to assimilate data within this geographic area would bias parameters of habitat indices. Consequently, the sub-tropical pole-and-line Japan fleet catch data were excluded from the likelihood calculation, and the catchability coefficient (fishing mortality) for this fleet was fixed.

Finally, to avoid the parameterization problem caused by the interplay between parameters, we reduced the number of mortality function parameters (Eqs. (A.14) and (A.15)) by fixing the coefficient of variability with habitat index  $\varepsilon$  at its guessed value 0.5. Sensitivity analysis (see below) indicate that the model is barely sensitive to this parameter.

## 5. Results

### 5.1. Sensitivity analyses

We applied sensitivity analyses to reveal which parameters can be estimated from available data and which can not. If model predictions are insensitive to some parameters, it is unlikely that they will be determined uniquely from available observations and should, therefore, be removed from the optimization. The parameters of the model  $\theta \in \mathfrak{R}^n$ , where  $n = 35$  (see Table 1).

Two types of sensitivity analyses were performed. The first, SA-1, examines whether the predictions of a given model are sensitive to its parameters. For this purpose we simply need to construct a function of the model solution, which represents model predictions (see, e.g., Worley, 1991). Then, the measures of sensitivity can be computed using precise gradients obtained from adjoint calculations. Since two types of data are assimilated within the SEAPODYM-APE model, i.e., catch and length frequencies, we construct the following functions:

$$R_1 = \sum_{tfij} (C_{tfij}^{\text{pred}})^2, \quad R_2 = \sum_{tfar} (Q_{tfar}^{\text{pred}})^2, \quad r = 1, \dots, 7. \quad (12)$$

**Table 1**  
Control parameters of the constrained optimization problem, imposed lower ( $\underline{\theta}$ ) and upper ( $\bar{\theta}$ ) boundaries and initial guess values ( $\theta^0$ )

$\theta$	Description		$\underline{\theta}$	$\bar{\theta}$	$\theta^0$
$\bar{m}_p$	Maximal mortality rate due to predation, Eq. (A.13)		0	1	0.5*
$\beta_p$	Slope coefficient in predation mortality, Eq. (A.13)		0	0.5	0.057
$\bar{m}_s$	Maximal mortality rate due to senescence, Eq. (A.12)		0	1	0.5*
$\beta_s$	Slope coefficient in senescence mortality, Eq. (A.12)		-0.5	0	-0.167
$A$	Threshold age (in month) of tuna for senescence mortality		20	40	31.29*
$\varepsilon$	Variability of tuna mortality with habitat quality, Eqs. (A.14) and (A.15)		0	1	0.5*
$\sigma_0$	Standard deviation in the temperature function of $I_0$ , Eq. (A.2)		2	4	3.5*
$T_0^c$	Optimal surface layer temperature for juveniles, Eqs. A.2 and A.3		28.5	31.5	30
$\alpha$	Half saturation constant for the food to predator ratio in the spawning index, Eq. (A.1)		0	5.0	0.1*
$\sigma_T$	Standard deviation in temperature function of $I_{2,m}$ , Eqs. (A.4) and (A.6)		1	3	2
$T_K^c$	Optimal temperature for oldest tuna, Eqs. (A.4) and (A.5)		25	28	26.0*
$\gamma$	Slope coefficient in the function of oxygen, defining adult habitat index, Eq. (A.4)		-10	0	-8*
$\bar{O}$	Threshold value of dissolved oxygen, defining adult habitat index, Eq. (A.4)		0.1	3.0	1.0
$c$	Coefficient of diffusion variability with habitat index, Eq. (A.10)		0	1	0.1
$V_m$	Maximal sustainable speed (in body length) of tuna, Eq. (A.9)		0	2	1.0
$R$	Maximal number of larvae at large spawning biomass of adults, Eq. (A.11)		0	2	0.5*
$b$	Slope coefficient in Beverton–Holt function, Eq. (A.11)		0	2	1.5*
$q_1$	Catchability of the fishery		0	0.1	0.00144*
$\zeta_1$	Steepness of selectivity function, type I, fishery	PLSUB	0	2.0	0.41*
$l_f$	Threshold fish length (see Eq. (A.17)) for the fishery		20	70	42*
$q_2$	Catchability of the fishery		0	0.1	0.003
$\zeta_2$	Steepness of sigmoid selectivity function, fishery	PLTRO	0	2.0	0.2
$l_2$	Threshold fish length (see Eq. (A.17)) for fishery		20	70	50
$q_3$	Catchability of the fishery		0	0.1	0.005
$\sigma_{s,3}$	Coefficient of selectivity function, type II, fishery	WPSASS	2	20	7.5
$l_3$	Target fish length (see Eq. (A.17)) for fishery		20	70	50
$q_4$	Catchability of the fishery		0	0.1	0.005
$\sigma_{s,4}$	Coefficient of selectivity function, type II, fishery	WPSUNA	2	20	7.5
$l_4$	Target fish length (see Eq. (A.17)) for the fishery		20	70	50
$q_5$	Catchability of the fishery		0	0.5	0.0018*
$\zeta_5$	Steepness of sigmoid selectivity function, fishery	EPSASS	0	2.0	0.22*
$l_5$	Threshold fish length (see Eq. (A.17)) for fishery		20	70	44.2*
$q_6$	Catchability of the fishery		0	0.5	0.0022*
$\zeta_6$	Steepness of sigmoid selectivity function, fishery	EPSUNA	0	2.0	0.29*
$l_6$	Threshold fish length (see Eq. (A.17)) for fishery		20	70	44.1*

Parameters marked by asterisks were fixed at their specified values in all experiments.

Then we define two measures of relative sensitivity  $\xi_1(\theta_k^0)$  and  $\xi_2(\theta_k^0)$  for corresponding model predictions and each initial guess parameter  $\theta_k^0$  as follows:

$$\xi_1(\theta_k^0) = \frac{1}{R_1} \frac{\partial R_1}{\partial \theta_k^0}, \quad \xi_2(\theta_k^0) = \frac{1}{R_2} \frac{\partial R_2}{\partial \theta_k^0}. \tag{13}$$

The second sensitivity analysis, SA-2, examines whether the objective function (which incorporates both predicted and observed data) is sensitive to model parameters. We compare values of likelihood at some found minimum  $\theta^\dagger$  to those evaluated at boundaries of parameter space (Vallino, 2000). We define two further measures of relative sensitivity:

$$\xi_3(\theta_k^\dagger) = \frac{L^-(\theta^\dagger + \delta\bar{\theta}_k \cdot \mathbf{e}_k) - L^-(\theta^\dagger)}{L^-(\theta^\dagger)}, \tag{14}$$

$$\xi_4(\theta_k^\dagger) = \frac{L^-(\theta^\dagger - \delta\underline{\theta}_k \cdot \mathbf{e}_k) - L^-(\theta^\dagger)}{L^-(\theta^\dagger)},$$

where  $\delta\bar{\theta}_k = \bar{\theta}_k - \theta_k^\dagger$ ,  $\delta\underline{\theta}_k = \theta_k^\dagger - \underline{\theta}_k$  and  $\mathbf{e}_k$  is a standard basis vector with 1 in the  $k$ th element and 0 elsewhere.

Both sensitivity tests applied for three optimization experiments showed similar qualitative results for most of the parameters (see Fig. 4). Sensitivity measures for parameters  $\alpha$  are persistently low and for parameter  $\bar{m}_p$  are ambiguous, showing however weak response of the cost function to this parameter's variation in the experiments with larger data sets (E2 and E3 experiments). Note that both parameters define mortality of youngest cohorts ( $\alpha$  introduces variability into mortality of larvae depending on forage and primary production ratio  $\lambda$  (Eq. A.1) and  $\bar{m}_p$  is the maximal mortality rate of larvae due to predation). Low sensitivities to these parameters are hence not too surprising, because explicit information about larvae and juveniles is not presented either in the observed catch or in size data. Sensitivities for parameters  $T_K^*$ ,  $\mu_4$ ,  $\zeta_5$  and  $\zeta_6$  were also low by both measures for all three experiments, and as a consequence these parameters were fixed, i.e., held constant, in the optimization.

Performing exhaustive sensitivity analysis, i.e., exploring entire likelihood hyper-surface in  $n$ -dimensional parametric space is practically impossible. Consequently, some non-estimable parameters are likely unrecognized, but large uncertainties obtained from the error analysis will be the good indicators of poorly determined parameters. Sensitivity analysis only gives a tentative representation of observable and non-observable parameters unless the entire parameter space is explored thoroughly. For example, it was found that decreasing one of the mortality parameters  $A$  increases

sensitivity to parameter  $\bar{m}_p$ . For this reason, the number of mortality parameters was reduced by fixing  $\bar{m}_p$  and  $\bar{m}_s$  and controlling the mortality-at-age function (Eqs. A.12 and A.13) with slope coefficients  $\beta_p$  and  $\beta_s$  only.

In contrast, despite the high model sensitivity to  $\sigma_0$  which is responsible for the width of the preferred temperature range for larvae and thus can extend or shrink the area of spawning grounds, the optimization tends to force it to its upper bound. Relaxing this bound in turn leads to searching solutions with unrealistic high population densities in the EPO and reducing catchability coefficients for Eastern fleets. It seems impossible to overcome these difficulties with the current state of the model and available biophysical and fishing data, hence coefficients  $\sigma_0$  and catchabilities for EPSASS and EPSUNA fisheries which appear to balance each other were fixed during the final estimation stage. There are several possible reasons causing the problem. However other than the errors associated with environmental forcing or presence of many zeros in the catch data, the most probable cause is the improper consideration of the epipelagic layer depth EPO which is known to be much shallower. Since we used the constant depth (=100 m) for aggregating environmental data everywhere, it could lead to lower temperatures in the East and hence to the wider temperature range. In future studies we envisage the use of dynamic in time and space euphotic depth to define the epipelagic layer.

### 5.2. Identical twin experiments

In order to verify that both the model and the method allow us to estimate chosen parameters using the available amount of observations we conducted so-called "identical twin experiments". These tests consist of estimating parameters from artificial data series constructed from predictions given by deterministic models. If optimization works well with our model and experiment set-up, then after sufficient perturbation of optimal parameters we should be able to retrieve them, because they determine known a priori solution represented in the artificial data series.

Thus, we constructed three artificial fishing data sets, using deterministic model outputs, i.e., without adding the noise and performed several minimizations for each experiment (E1–E3) starting with perturbed parameters (but not changing initial conditions). Control parameters were successfully recovered for all three experiments with small relative errors  $\varepsilon < 0.001$  due to the computer round-off error. The example plot of parameter evolution during minimization process is given in Fig. 5 for the simulated twin experiment with the E2 parameter set. All control

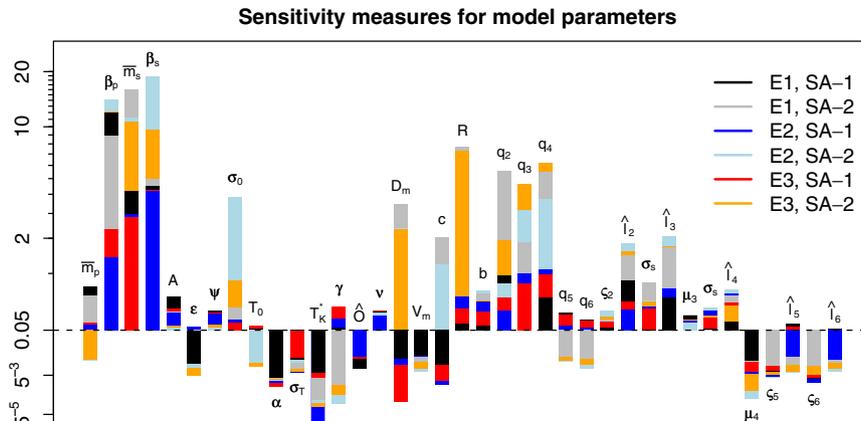
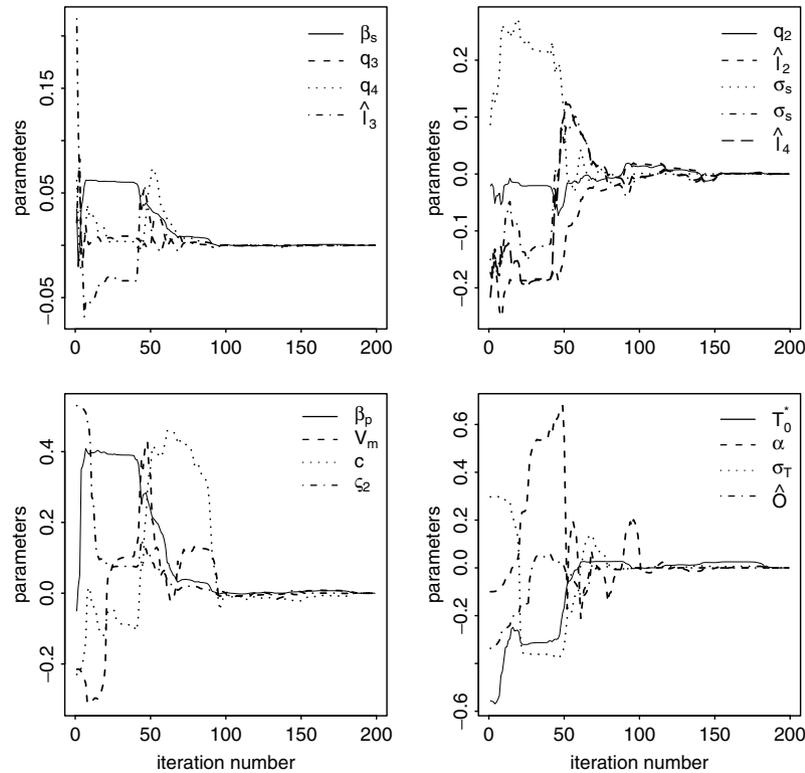


Fig. 4. Log-scaled measures of sensitivity obtained for each parameter within estimation experiments E1–E3. The values below the dashed line correspond to less than 5% sensitivity of either predictions (SA-1, i.e.,  $\max|\xi_1, \xi_2|$ ) or cost function (SA-2,  $\max|\xi_3, \xi_4|$ ) to corresponding parameter.



**Fig. 5.** Evolution of control parameters during twin data experiment conducted for the artificial data simulated with E2 parameter set. Parameters are grouped by their sensitivities in descending order, except  $\beta_p$  for which high sensitivity was calculated by both SA-1 and SA-2 analysis.

parameters (except  $\beta_p$ ) are grouped according to the sensitivity measures described above – from highest sensitivities to lowest. The parameters to which the model is most sensitive were recovered after fewer iterations of the quasi-Newton method. Despite high sensitivity of the model to  $\beta_p$  and small relative error (<1%), it was one of the last parameters to stabilize at its optimum value. This is likely due to its high correlation with the senescence  $\beta_s$  mortality coefficient (see further in Table 5), which stabilized around 100th iteration but continued to vary within a tiny range (see Fig. 5, upper left plot) until stabilization of  $\beta_p$  at its optimal value.

### 5.3. Parameter estimates and errors

The resulting parameter estimates for experiments E1–E3 are given in Table 2. The values differ among experiments depending on time period. The E3 estimates should be considered to be the most representative as describing the population life cycle because they were obtained by assimilation of the longest time series of fishing data.

Unfortunately, there is no simple way to evaluate the uniqueness of the estimated parameters in non-linear problems especially for ecosystem models due to their extreme complexity, high dimension of the objective functional and scarcity of available observations (Robinson and Lermusiaux, 2002; Vallino, 2000; Matear, 1995). Thus, we used the common approach of perturbing the model parameters and restarting the experiment but were not able to determine if the found solutions are local minima. However, considering the dimension of the minimization problem and the computer time to perform one experiment, it seemed unrealistic to make an exhaustive study that would allow us to conclude that computed solutions are global. We can, however, determine whether the estimated parameters were well determined at the minima detected by minimization routine.

**Table 2**

Estimated parameters and their standard deviation uncertainties, obtained from Hessians approximated by finite differences which utilize the exact derivatives

N	$\theta$	E1	E2	E3
1	$\beta_p$	$0.124 \pm 0.0026$	$0.388 \pm 0.002$	$0.296 \pm 0.0018$
2	$\beta_s$	$-0.087 \pm 0.0067$	$-0.0347 \pm 0.001$	$-0.044 \pm 0.0015$
3	$T_0^*$	$29.9 \pm 0.0063$	$31.43 \pm 0.017$	$30.47 \pm 0.0047$
4	$\alpha'$	$1.39 \pm 0.011$	$3.178 \pm 0.015$	$3.67 \pm 0.016$
5	$\sigma_T$	$1.26 \pm 0.0047$	$2.116 \pm 0.002$	$1.62 \pm 0.0015$
6	$\hat{O}$	$3.65 \pm 0.002$	$3.854 \pm 0.0009$	$3.86 \pm 0.0009$
7	$c$	$0.47 \pm 0.0098$	$0.504 \pm 0.01$	$0.4 \pm 0.005$
8	$V_m$	$1.72 \pm 0.0097$	$1.526 \pm 0.008$	$1.3 \pm 0.006$
9	$q_2$	$0.004 \pm 0.0018$	$0.0088 \pm 0.0055$	$0.0045 \pm 0.0016$
10	$q_3$	$0.0085 \pm 0.0027$	$0.0063 \pm 0.0014$	$0.0044 \pm 0.0018$
11	$q_4$	$0.0023 \pm 0.0013$	$0.0045 \pm 0.0015$	$0.0024 \pm 0.001$
12	$\zeta_2$	$0.24 \pm 0.0028$	$0.185 \pm 0.0026$	$0.192 \pm 0.001$
13	$l_2$	54	$63.33 \pm 0.007$	60.33
14	$\sigma_{s,3}$	$3.55 \pm 0.0215$	$4.98 \pm 0.004$	$4.96 \pm 0.0036$
15	$l_3$	54	$48.73 \pm 0.0008$	$48.76 \pm 0.0007$
16	$\sigma_{s,4}$	$9.48 \pm 0.0089$	$13.98 \pm 0.005$	$13.92 \pm 0.004$
17	$l_4$	$47.29 \pm 0.0042$	$61.01 \pm 0.004$	$59.3 \pm 0.003$
18	$\beta_5$	$0.004 \pm 0.0009$	$0.0075 \pm 0.0009$	$0.004 \pm 0.0001$
19	$\beta_6$	$0.007 \pm 0.0012$	$0.005 \pm 0.0009$	$0.003 \pm 0.0001$
20	$p_{0.5}$	$0.25 \pm 0.014$	$0.29 \pm 0.0097$	$0.24 \pm 0.007$
21	$p_{0.6}$	$0.05 \pm 0.022$	$0.075 \pm 0.017$	$0.05 \pm 0.012$

Parameters, for which uncertainties are not given, were fixed in the current experiment.

Note:  $\alpha'$  here is the argument of  $I_{SE,a}$  function (Eq. (A.7)), while it was fixed (=0.1) in spawning function (Eq. (A.11)) and natural mortality for larvae (Eq. (14)), see text for details.

We compute the variance of the estimated parameters from the inverse of the Hessian matrix, i.e.,  $C = \mathbf{H}^{-1}$  (Bard, 1974), where  $\mathbf{H} = \frac{\partial^2 L}{\partial \theta_i \partial \theta_j}$ ,  $i, j = 1, 2, \dots, n$  is the Hessian matrix evaluated at the minimum of the negative log-likelihood function. The diagonal elements of  $C$  provide estimates of the variance of the optimal

**Table 3**

Spatial monthly average correlation between predicted and observed catch by fishery, computed in SEAPODYM simulation with non-optimized (N/O) parameter set (Lehodey et al., 2003, see Table 1, column for S4 simulation) and simulation with estimated E3 parameters

Fishery	N/O	E3
PL Japan	0.37	0.8
PL tropical	0.73	0.88
WCPO PS associated	0.7	0.84
WCPO PS unassociated	0.51	0.81
EPO PS associated	–	0.67
EPO PS unassociated	–	0.71

parameters. The Hessian matrix was approximated with central finite difference using first derivatives exactly evaluated by adjoint calculations.

The estimated parameters and their calculated uncertainty (one standard deviation) are shown in Table 2. All population parameters ((1)–(8)) and coefficients of selectivity functions are estimated most accurately by the optimization given their small uncertainties, while catchabilities of WCPO fisheries and probability of zero catch in unassociated EPO fisheries (see Eq. 9) have higher errors in all experiments.

Correlation coefficients between pairs of estimated parameters can also be calculated from the error-covariance matrix (see Tables 4–6). These characteristics provide additional information

to the question of identifiability of the model parameters that we appealed to earlier. Eventually, the high correlations were observed between selectivity and catchability parameters (Tables 5 and 6). High values for the pair  $\sigma_T$  and  $T_0^*$  (Table 4) suggests that predicted water temperature data for 1980–1990 period (E1) did not provide a clear signal for simultaneous estimation of the optimal temperature for spawning and the extension of the adult thermal habitat. Note, that a high correlation between these parameters appears again in the E3 experiment, which covers E1 period as well. Also, the high correlation between  $\beta_p$  and  $\beta_s$  in all experiments suggests combining these mortality parameters.

Assimilating the longest time series (E3) showed dependency between the senescence mortality coefficient  $\beta_s$  (i.e., mortality rate of oldest cohorts) and catchability (fishing mortality) for pole-and-line tropical fishery, while this dependency does not exist within the solutions found in E1 and E2. Such a result most likely is due to the poor choice of the selectivity parameter for PLTRO fishery, which was fixed in the E3 experiment (see Table 2) because of low sensitivity.

**6. Discussion and conclusion**

The goal of this study was to find the “best” model parameters which would give us confidence in the ability of the model to reasonably describe real ecosystem. The definition of the best solu-

**Table 4**

Correlation coefficients between optimal parameters obtained for E1

	$\beta_p$	$\beta_s$	$T_0^*$	$\alpha$	$\sigma_T$	$\hat{O}$	$V_m$	$c$	$q_2$	$q_3$	$q_4$	$\zeta_2$	$\sigma_{s,3}$	$\sigma_{s,4}$	$\hat{l}_4$
$\beta_p$	1	<b>0.99</b>	-0.06	0.07	-0.06	-0.05	-0.24	-0.26	-0.11	-0.4	-0.4	0.23	0.06	0.02	0.01
$\beta_s$	<b>0.99</b>	1	-0.08	0.09	-0.07	-0.05	-0.19	-0.21	0.03	-0.3	-0.31	0.26	0.05	0.03	0
$T_0^*$	-0.06	-0.08	1	-0.24	<b>0.97</b>	-0.02	-0.03	-0.28	-0.01	-0.05	-0.05	0.04	0	-0.02	-0.01
$\alpha$	0.07	0.09	-0.24	1	-0.24	-0.01	0.43	0.06	0.05	0.03	0.01	0.02	-0.04	0.01	0.01
$\sigma_T$	-0.06	-0.07	<b>0.97</b>	-0.24	1	-0.02	-0.03	-0.26	-0.01	-0.04	-0.05	0.04	0	-0.02	-0.01
$\hat{O}$	-0.05	-0.05	-0.02	-0.01	-0.02	1	0.07	-0.02	-0.01	0.02	0.02	-0.01	0.01	0	-0.01
$V_m$	-0.24	-0.19	-0.03	0.43	-0.03	0.07	1	0.74	0.34	0.28	0.35	0.02	-0.09	0.04	0.09
$c$	-0.26	-0.21	-0.28	0.06	-0.26	-0.02	0.74	1	0.34	0.29	0.39	-0.01	-0.05	0.03	0.07
$q_2$	-0.11	0.03	-0.01	0.05	-0.01	-0.01	0.34	0.34	1	0.7	0.69	0.35	-0.08	0.04	-0.04
$q_3$	-0.4	-0.3	-0.05	0.03	-0.04	0.02	0.28	0.29	0.7	1	0.6	0.07	-0.68	0.02	-0.05
$q_4$	-0.4	-0.31	-0.05	0.01	-0.05	0.02	0.35	0.39	0.69	0.6	1	0.06	-0.06	-0.05	0.27
$\zeta_2$	0.23	0.26	0.04	0.02	0.04	-0.01	0.02	-0.01	0.35	0.07	0.06	1	0	0.01	-0.02
$\sigma_{s,3}$	0.06	0.05	0	-0.04	0	0.01	-0.09	-0.05	-0.08	-0.68	-0.06	0	1	0	0.02
$\sigma_{s,4}$	0.02	0.03	-0.02	0.01	-0.02	0	0.04	0.03	0.04	0.02	-0.05	0.01	0	1	0.78
$\hat{l}_4$	0.01	0	-0.01	0.01	-0.01	-0.01	0.09	0.07	-0.04	-0.05	0.27	-0.02	0.02	0.78	1

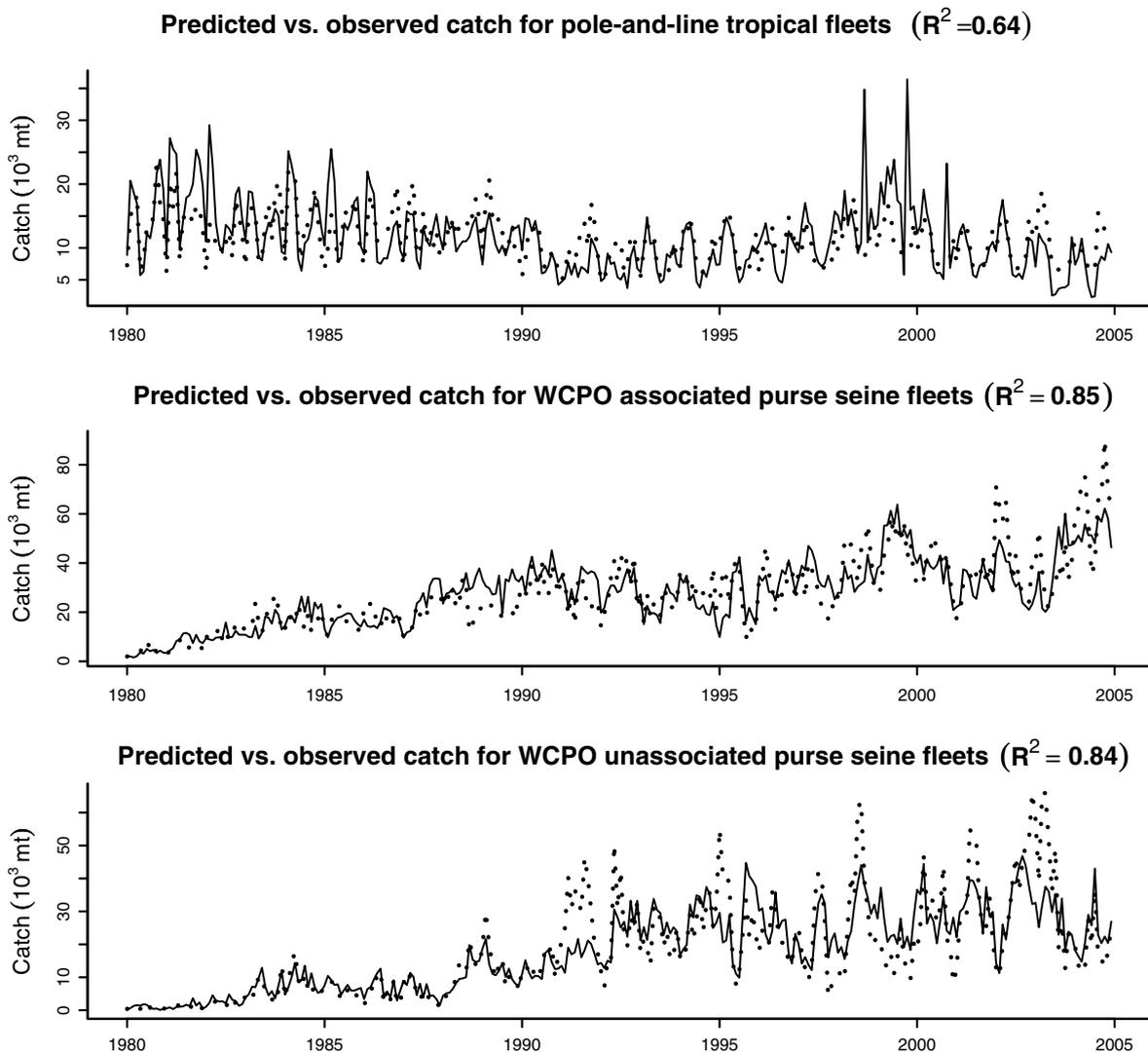
**Table 5**

Correlation coefficients between optimal parameters obtained for E2

	$\beta_p$	$\beta_s$	$T_0^*$	$\alpha$	$\sigma_T$	$\hat{O}$	$V_m$	$c$	$q_2$	$q_3$	$q_4$	$\zeta_2$	$\hat{l}_2$	$\sigma_{s,3}$	$\hat{l}_3$	$\sigma_{s,4}$	$\hat{l}_4$
$\beta_p$	1	<b>0.88</b>	0.6	-0.03	0.38	-0.08	0.54	0.5	-0.1	0.32	0.4	0.29	-0.27	-0.04	0.06	-0.02	0.06
$\beta_s$	<b>0.88</b>	1	0.27	-0.02	0.15	-0.1	0.58	0.67	-0.19	0.56	0.63	0.41	-0.41	-0.1	0.01	0.03	0.15
$T_0^*$	0.6	0.27	1	-0.05	0.72	-0.07	0.31	-0.08	0.01	-0.02	0.01	0.02	-0.02	0.01	0.07	-0.03	-0.03
$\alpha$	-0.03	-0.02	-0.05	1	-0.04	0.06	0.18	-0.13	0.09	-0.03	-0.04	-0.05	0.09	-0.02	-0.06	-0.06	-0.05
$\sigma_T$	0.38	0.15	0.72	-0.04	1	0.09	0.14	-0.05	0.01	-0.01	0.03	0.01	-0.01	0.01	0.05	0	0.01
$\hat{O}$	-0.08	-0.1	-0.07	0.06	0.09	1	0.06	-0.07	0.09	-0.07	-0.1	-0.08	0.11	0	-0.03	-0.03	-0.05
$V_m$	0.54	0.58	0.31	0.18	0.14	0.06	1	0.68	0.05	0.27	0.25	0.16	-0.1	-0.07	0.01	-0.16	-0.07
$c$	0.5	0.67	-0.08	-0.13	-0.05	-0.07	0.68	1	-0.03	0.43	0.39	0.24	-0.19	-0.14	-0.05	-0.13	-0.02
$q_2$	-0.1	-0.19	0.01	0.09	0.01	0.09	0.05	-0.03	1	-0.15	-0.31	-0.77	<b>0.96</b>	-0.02	-0.06	-0.18	-0.23
$q_3$	0.32	0.56	-0.02	-0.03	-0.01	-0.07	0.27	0.43	-0.15	1	0.45	0.3	-0.31	-0.79	-0.58	-0.01	0.06
$q_4$	0.4	0.63	0.01	-0.04	0.03	-0.1	0.25	0.39	-0.31	0.45	1	0.41	-0.47	-0.03	0.06	0.47	0.73
$\zeta_2$	0.29	0.41	0.02	-0.05	0.01	-0.08	0.16	0.24	-0.77	0.3	0.41	1	<b>-0.87</b>	-0.05	0.01	0.13	0.2
$\hat{l}_2$	-0.27	-0.41	-0.02	0.09	-0.01	0.11	-0.1	-0.19	<b>0.96</b>	-0.31	-0.47	<b>-0.87</b>	1	0.02	-0.05	-0.18	-0.26
$\sigma_{s,3}$	-0.04	-0.1	0.01	-0.02	0.01	0	-0.07	-0.14	-0.02	-0.79	-0.03	-0.05	0.02	1	<b>0.88</b>	0.12	0.09
$\hat{l}_3$	0.06	0.01	0.07	-0.06	0.05	-0.03	0.01	-0.05	-0.06	-0.58	0.06	0.01	-0.05	<b>0.88</b>	1	0.15	0.14
$\sigma_{s,4}$	-0.02	0.03	-0.03	-0.06	0	-0.03	-0.16	-0.13	-0.18	-0.01	0.47	0.13	-0.18	0.12	0.15	1	<b>0.92</b>
$\hat{l}_4$	0.06	0.15	-0.03	-0.05	0.01	-0.05	-0.07	-0.02	-0.23	0.06	0.73	0.2	-0.26	0.09	0.14	<b>0.92</b>	1

**Table 6**  
Correlation coefficients between optimal parameters obtained for E3

	$\beta_P$	$\beta_S$	$T_0^{\sigma}$	$\alpha$	$\sigma_T$	$\hat{O}$	$V_m$	$c$	$q_2$	$q_3$	$q_4$	$\zeta_2$	$\sigma_{s,3}$	$\hat{l}_3$	$\sigma_{s,4}$	$\hat{l}_4$
$\beta_P$	1	<b>0.93</b>	0.32	0.02	0.21	0.13	0.51	0.4	0.73	0.43	0.5	0.35	-0.15	-0.07	0	0.03
$\beta_S$	<b>0.93</b>	1	0.06	0.07	0	0.04	0.61	0.59	<b>0.91</b>	0.62	0.68	0.42	-0.22	-0.14	-0.01	0.06
$T_0^{\sigma}$	0.32	0.06	1	-0.21	<b>0.94</b>	0.27	-0.07	-0.44	-0.12	-0.18	-0.23	-0.04	0.05	0.07	0.01	-0.07
$\alpha$	0.02	0.07	-0.21	1	-0.21	-0.04	0.26	-0.02	0.1	0.06	0.08	0.06	-0.03	-0.07	-0.04	0
$\sigma_T$	0.21	0	<b>0.94</b>	-0.21	1	0.23	-0.13	-0.45	-0.13	-0.16	-0.22	-0.05	0.04	0.05	0.02	-0.07
$\hat{O}$	0.13	0.04	0.27	-0.04	0.23	1	0.11	-0.17	-0.06	-0.09	-0.09	-0.02	0.05	0.04	-0.01	-0.04
$V_m$	0.51	0.61	-0.07	0.26	-0.13	0.11	1	0.72	0.59	0.38	0.39	0.31	-0.14	-0.09	-0.15	-0.07
$c$	0.4	0.59	-0.44	-0.02	-0.45	-0.17	0.72	1	0.65	0.5	0.51	0.31	-0.2	-0.14	-0.14	-0.03
$q_2$	0.73	<b>0.91</b>	-0.12	0.1	-0.13	-0.06	0.59	0.65	1	0.71	0.73	0.61	-0.25	-0.18	0	0.07
$q_3$	0.43	0.62	-0.18	0.06	-0.16	-0.09	0.38	0.5	0.71	1	0.56	0.32	<b>-0.82</b>	-0.64	-0.05	0.01
$q_4$	0.5	0.68	-0.23	0.08	-0.22	-0.09	0.39	0.51	0.73	0.56	1	0.3	-0.19	-0.13	0.28	0.57
$\zeta_2$	0.35	0.42	-0.04	0.06	-0.05	-0.02	0.31	0.31	0.61	0.32	0.3	1	-0.14	-0.11	-0.02	0
$\sigma_{s,3}$	-0.15	-0.22	0.05	-0.03	0.04	0.05	-0.14	-0.2	-0.25	<b>-0.82</b>	-0.19	-0.14	1	<b>0.86</b>	0.08	0.05
$\hat{l}_3$	-0.07	-0.14	0.07	-0.07	0.05	0.04	-0.09	-0.14	-0.18	-0.64	-0.13	-0.11	<b>0.86</b>	1	0.11	0.08
$\sigma_{s,4}$	0	-0.01	0.01	-0.04	0.02	-0.01	-0.15	-0.14	0	-0.05	0.28	-0.02	0.08	0.11	1	<b>0.89</b>
$\hat{l}_4$	0.03	0.06	-0.07	0	-0.07	-0.04	-0.07	-0.03	0.07	0.01	0.57	0	0.05	0.08	<b>0.89</b>	1



**Fig. 6.** Domain-aggregated catches for Western Pacific fisheries taken into account in the function minimization procedure. Solid line denotes predicted catch, dotted – observed data.

tion here is circumscribed by the constraints posed on the parameters, the form of cost function being chosen, the accuracy of forcing, the observational errors, and the model itself. The best fit

could conceivably be located outside the bounds of parameter space, but the parameters which are beyond the scope of their biological meaning would probably yield an unrealistic solution

that does not correspond to our present knowledge of the ecosystem.

The agreement between model predictions and observations is presented as domain-aggregated time series of WCPO catch in Fig. 6. Also, the average spatial monthly correlation between predicted and observed catch by fishery are compared to values computed by SEAPODYM without optimization in Table 3. Fig. 7 shows examples of predicted spatial distributions of adult skipjack with CPUE data that was incorporated into the likelihood function. Environmental conditions clearly have a strong influence on population distributions as well as on CPUE indices and it is encouraging that the model is able to describe these effects with fairly small number of control variables.

The dynamics predicted by the habitat-based spatially explicit model, SEAPODYM-APE, are generally in agreement with the qualitatively different statistical length-based stock assessment model, MULTIFAN-CL, (Fig. 8), with the correlation  $R^2 = 0.46$ . However, SEAPODYM-APE predictions suggest much more moderate (in

amplitude) variations in skipjack stock. Note that range of variability predicted by the biophysical coupled model that drives SEAPODYM-APE is also too low compared to actual variability (McKinley et al., 2006).

Dynamics of the population biomass predicted by the two models differ substantially during 1978–1982 and 1992–1997 periods. These two periods correspond to post-El Niño ecosystem conditions which are known to be favorable for skipjack recruitment (Lehodey et al., 2003) through expansion of the skipjack spawning grounds and then, bringing more accessible forage to the western–central Pacific region. Comparison of predicted biomass time series of young tuna and the Southern Oscillation Index (SOI) shows direct relationship between ENSO events and changes in the population dynamics. The maximum correlation between the two series ( $-0.63$ ) is obtained with a SOI series lagged by 8 months, a time lag matching with the age of recruits, and thus suggesting that the ENSO impact occurs directly on the early life history of the species (i.e., spawning index). This results

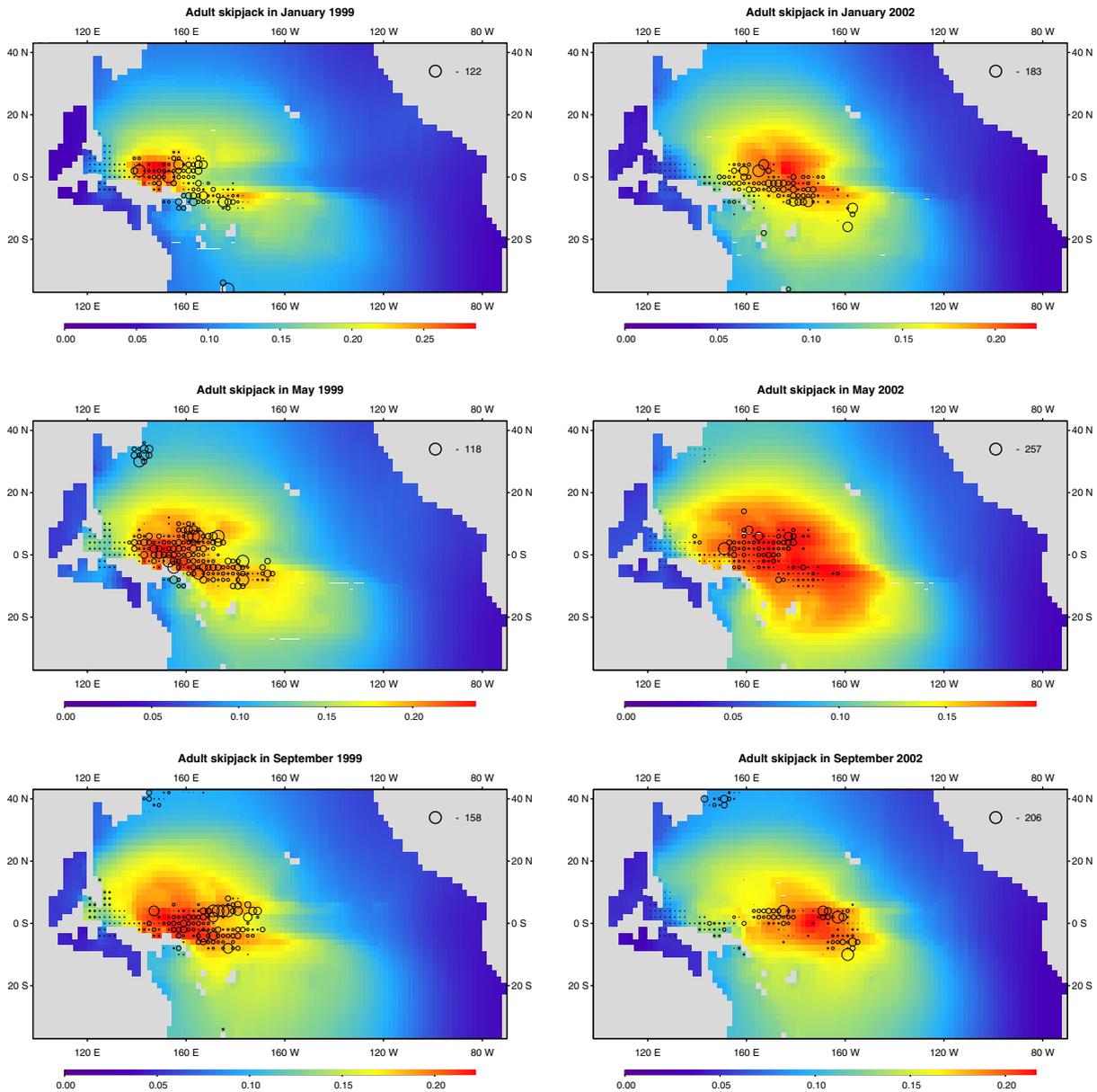
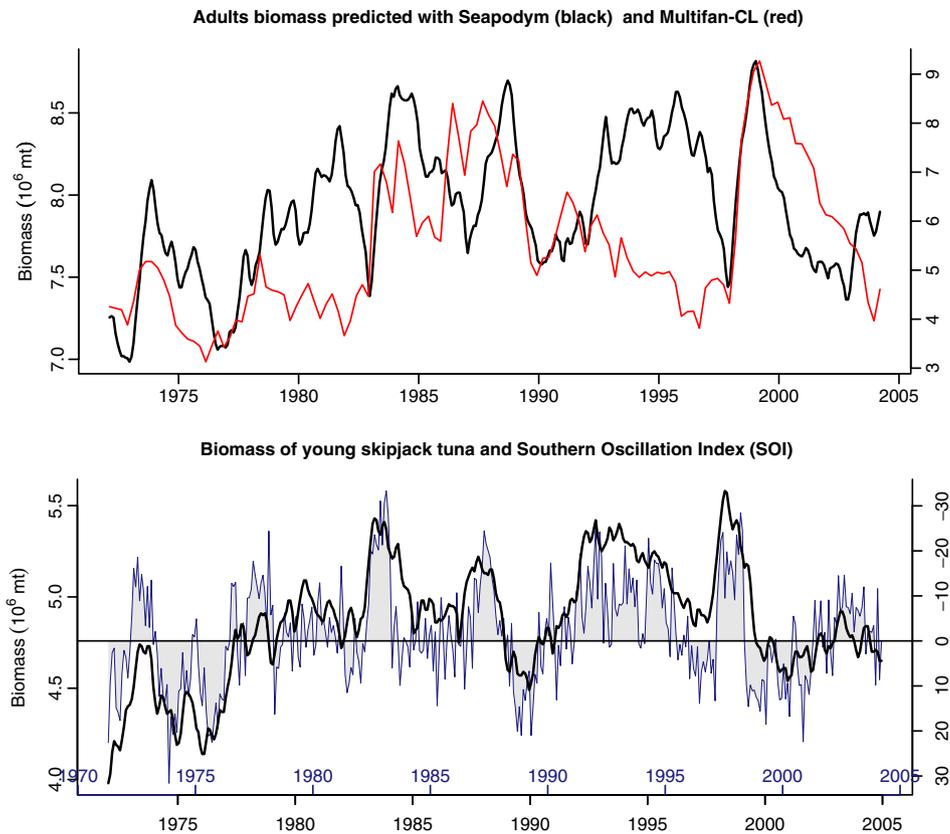


Fig. 7. Observed catch per unit of effort for purse seine fisheries plotted over predicted skipjack distribution (adults of size 24–71 cm) during La Niña conditions (left column) and development of El Niño at the second half of year 2002 (right column).



**Fig. 8.** Upper plot shows comparison of predicted biomass of skipjack in WCPO area (regions 1–6) by SEAPODYM (left axis) and MULTIFAN-CL (right axis, units are the same) models,  $R^2 = 0.46$ . Parameterization achieved in E3 experiment was used to simulate population dynamics for wider time period, starting from 1972. Lower plot shows biomass of young skipjack tuna (sum of ages from 3 months to 3 quarter) and 8-month lagged Southern Oscillation Index (notice that y-axis is inverted) as an indicator of El-Niño event (see text for more details).

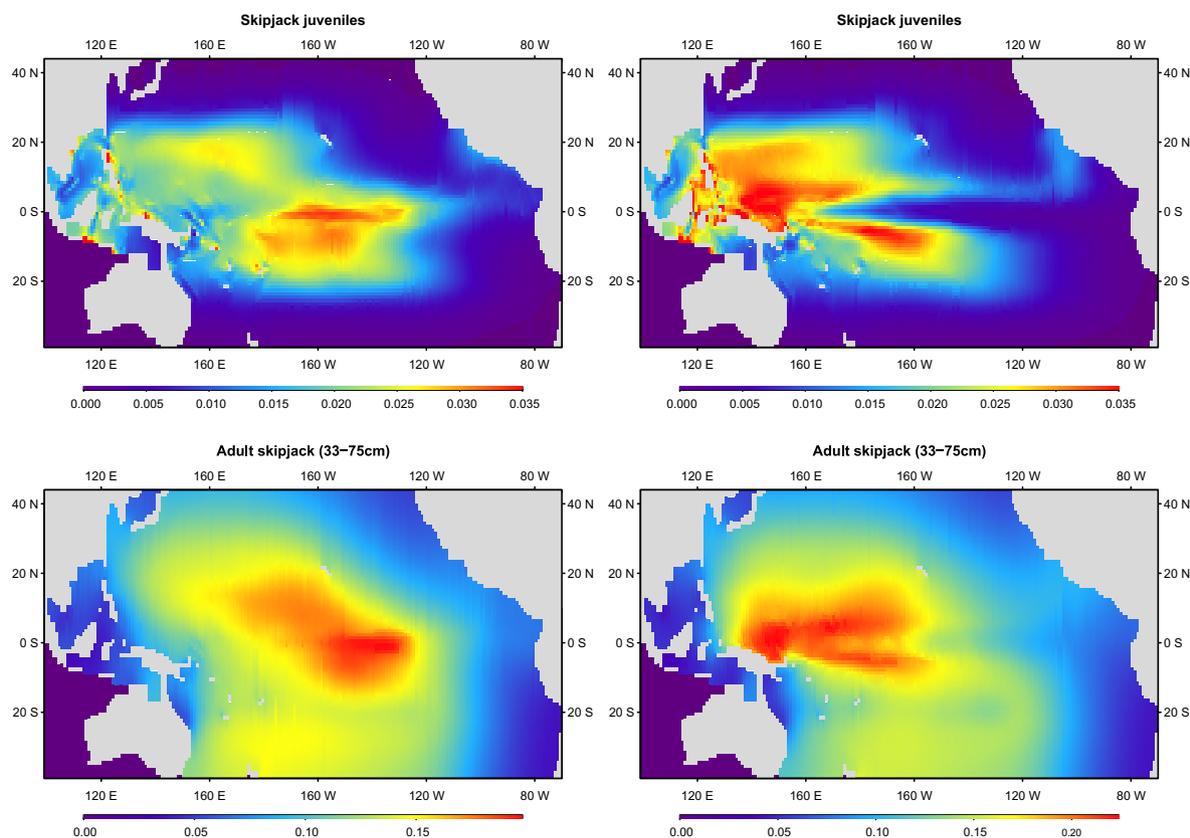
confirm previous analyses (e.g., [Lehodey et al., 2003](#)), but since it emerges from a rigorous statistical approach, it provides higher confidence in a finding that can be relevant to the economic management of the fishery, namely, the general trend in abundance of the adult stock being predictable 8 months in advance simply using the SOI.

The impact of ENSO variability on skipjack movement has been also demonstrated. During El Niño, the skipjack population moves eastward (see [Fig. 9](#)), and the biomass in the Central and Eastern Pacific increases, while it decreases in the Western Pacific correspondingly. These spatial changes very likely affect the catch and could explain the discrepancy between SEAPODYM-APE and MULTIFAN-CL biomass estimates. The latter would “interpret” a sudden drop of catches in the WCPO due to an eastward displacement of a large fraction of the stock by a decrease in the stock abundance. Conversely, the environmental spatial model SEAPODYM agrees with such catch reduction because it considers the entire Pacific domain and, more importantly, it explicitly predicts catch declines due to the population’s eastward migration, when application of current fishing effort gave lower catches.

Overall parameter values estimated by optimization procedures are biologically reasonable. The threshold value of ambient oxygen level in adult skipjack habitat was estimated to be  $\bar{O} = 3.86$  ml/l. This value coincides with the current knowledge of skipjack physiology stating that skipjack exhibit highest oxygen demands ([Brill, 1994](#)) among tuna species. For example, during a sonic tracking study of skipjack tuna of sizes 41–52 cm fork length, [Cayre \(1991\)](#) observed that skipjack spent most of the

time at depths with ambient oxygen level higher 3.8 ml/l. Note however, that at the time of the study only seasonal climatology of  $O_2$  was available.

Summarizing results of field studies, [Lehodey \(2001\)](#) previously suggested that the parameter  $\alpha$ , defining spawning habitat spatial structure (Eq. (A.1)) as well as adult seasonal migrations to spawning grounds (i.e.,  $I_0$  in Eq. (A.7)), should be small for skipjack and gradually increase for yellowfin, bigeye, albacore and bluefin tuna respectively. We intended to verify this in optimization experiments, however the low sensitivity of the model to  $\alpha$  (see [Fig. 5](#)) led us to differentiate the effects of the ratio of primary production to forage biomass (the strength of which is defined by  $\alpha$ ) on resulting larval distributions, and on the movement of adults toward the spawning zones respectively. Namely, when seasonal effects are not considered in the adult habitat definition, the parameter estimate always tended to be 0 leading to more smooth larval distributions defined only by temperature function. We fixed  $\alpha$  to a small value (0.1) in the spawning habitat and released it to estimate a value representing the effect on the movement of adult fish. In this case the model estimates a non-null positive value for  $\alpha$  with relatively small uncertainty (see [Table 2](#)). These results suggest that spawning migration of adult skipjack is important to include in sub-tropical areas where the seasonal threshold is efficient, but that either spawning conditions for adults are different from actual preferences of larvae, or more likely, the model does not predict correctly the redistribution of larvae and juveniles, e.g., because of insufficient data or underestimated currents, or too low spatial resolution with no mesoscale representation. Clearly, the model needs better resolu-



**Fig. 9.** Skipjack biomass predicted by SEAPODYM with optimized parameters during a strong El Niño event (November 1997, left column) and during La Niña conditions (November 1998, right column).

tion and data on early life history for a better understanding of this mechanism. Application to other species with a more distinct seasonal spawning behavior and habitat should also provide useful information.

The results of this study show that a new generation of models integrating the progress in physical and biogeochemical oceanography modelling, a detailed spatially explicit modelling of population dynamics and up-to-date data assimilation techniques can provide a new powerful tool for an ecosystem-based management of exploited species, allowing to investigate impacts due to both fishing and environmental changes. Despite the increased level of detail (both in space and time) in comparison to standard population dynamic models, this approach does not necessarily mean more parameters to estimate. Indeed, because environment is such a strong constraint, it allows reducing the number of parameters in the population dynamics model itself. Spatially-explicit models for stock assessment based on fishing data offer also the advantage of using gear catchability – a critical parameter in stock assessment – in a sense closer of its true definition, since environment-related variability (e.g., migration and recruitment) is explicit in the model and all changes increasing space-related fleet efficiency is directly included with the use of spatially-disaggregated fishing data.

But in parallel, environmental forcing fields need to be accurate, and environmentally-constrained mechanisms need to be robust to avoid introducing other biases in the model. In particular, due to the general dearth of observations, predicted outputs of large-scale micronekton biomass distributions lack a strict evaluation. It would be beneficial to apply the current coupled model for assimilating available forage data (e.g., acoustic profiles) together

with tuna catch data to optimize the parameterization of the mid-trophic sub-model.

Since the mechanisms in the model are linked to the environmental conditions, the optimization necessarily produces parameter estimates that depend on the forcing field used. It is therefore essential to run optimization experiments with multiple forcing data sets. They will highlight the most sensitive parameters and provide an envelope (or ensemble) of predictions. We can also expect that predicted physical–biogeochemical forcing fields will improve toward more and more realistic conditions, allowing approaching actual values of the biological parameters. These parameters can be also evaluated independently, for example using electronic tagging data. Independent measures or estimates of these parameters, for example using electronic tagging data, should assist in the evaluation of model predictions.

Finally, future efforts to optimize the model parameters for other tuna species in the same ocean, then in the Indian and Atlantic oceans should bring helpful complementary information on the capability of the model to produce coherent estimates between species. Given both conservation and economical concerns on big-eye and yellowfin tunas, this task is urgent.

#### Acknowledgements

We are grateful to Raghu Murtugudde from the Earth System Science Interdisciplinary Center (University of Maryland), who provided the outputs of the general ocean circulation model and NPZD modelled primary productivity data. We would like

to acknowledge Peter Williams (SPC) and Michael Hinton (IATTC) for preparing and supplying catch and size composition data. We also want to thank John Hampton (SPC) for useful comments and practical advice regarding the model parameterization and data structure. We offer sincere gratitude to Francois Royer (CLS) for the thoughtful review of the manuscript before its submission. This work was funded in part by Cooperative Agreement NA17RJ1230 between the Joint Institute for Marine and Atmospheric Research (JIMAR) and the National Oceanic and Atmospheric Administration (NOAA). The views expressed herein are those of the authors and do not necessarily reflect the views of NOAA of any of its sub-divisions.

## Appendix A. SEAPODYM parameterization

### A.1. Habitat indices

Physical and biogeochemical conditions influence fish population dynamics through changes in spawning conditions, habitat suitability, and distributions of food resources, thus inducing changes in fish movement behavior, reproduction and mortality. Environmental data are used in SEAPODYM to build habitat suitability indices, and the values of these indices control dynamical processes in the simulated populations. Three types of habitat indices are defined which link environmental variables to the dynamics of different life stages of tuna: spawning habitat index  $I_0$  describing favorability of the habitat for individuals less than one month in age, juvenile habitat index  $I_1$  defined for individuals aged 1–2 months, and adult habitat index  $I_{2,a}$ , which describes influence of environment on adult tunas of age  $a$ . Two special cases of the adult habitat index take into account seasonality of migrations and food requirements of adult tuna.

#### A.1.1. Spawning habitat index

The habitat index  $I_0$  used to constrain larval production and mortality of age 0 individuals (see below) is a function of surface layer temperature  $T_0$ , tuna forage biomass in the surface layer  $F_0$  and primary production converted to the wet weight of zooplankton  $P_{ww}$ . So, if ocean primary production  $P$  is given in  $\text{mmol C m}^{-2}$ , after conversion to the wet weight of zooplankton species the units of  $P_{ww}$  become  $\text{g/m}^2$ . Let  $\mathcal{A} = EP_{ww}/F_0$  denote the ratio between food for larvae and the tuna forage that is considered as the potential predator for larvae. The constant  $E$  is the energy transfer coefficient.

Spawning habitat index is defined as the following:

$$I_0 = \phi(\mathcal{A}) \cdot \Phi(T_0),$$

where  $\phi(\mathcal{A})$  is the non-linear saturation function determined in  $[0,1]$  interval:

$$\phi(\mathcal{A}) = \frac{\mathcal{A}}{\alpha + \mathcal{A}}. \quad (\text{A.1})$$

The curvature parameter  $\alpha$  is unknown and included in the list of parameters to be estimated from the data. Dependence on sea surface temperature is described by a Gaussian function:

$$\Phi(T_0) = \frac{g}{\sigma_0 \sqrt{2\pi}} e^{-\frac{(T_0 - T_0^*)^2}{2\sigma_0^2}}, \quad (\text{A.2})$$

where the parameters  $T_0^*$  and  $\sigma_0$  are the optimal temperature and width of tolerance interval (standard deviation) in the Gaussian.

### A.1.2. Juvenile habitat index

The juvenile habitat index  $I_1$  is a function of temperature in surface layer and adult tuna density, which accounts for cannibalism by adults:

$$I_1 = \left(1 - \frac{\langle N \rangle}{h + \langle N \rangle}\right) \cdot \Phi(T_0), \quad (\text{A.3})$$

where  $\langle N \rangle = \sum_{a=1}^K N_a$ , i.e., is the total size of adult portion of the population, and the unknown parameter  $h$  determines the cannibalism intensity in the habitat depending on the total number of adults tuna being present locally. Juvenile indices are used to compute juvenile mortalities, which therefore become variable in time and space (see below).

### A.1.3. Adult habitat index

The adult habitat index  $I_{2,a}$  is an indicator of suitability of the habitat for feeding fish. It is proportional to the local forage densities  $F_n$  weighted by the accessibility coefficients  $\Theta_{a,n}$  as the functions of environmental conditions:

$$I_{2,a} = \sum_n \Theta_{a,n} F_n. \quad (\text{A.4})$$

The more favorable environmental conditions are for tuna of age  $a$  at given depth layer to access  $n$ th forage component ( $n = 1, \dots, 6$ ), the more likely this habitat will be preferred by tuna for foraging. Two factors, temperature and oxygen, are considered to be critical for tuna during feeding. Their importance is described by a Gaussian function of temperature and a sigmoidal function of dissolved oxygen:

$$\Phi_a(T_z) = \frac{g}{\sigma_a \sqrt{2\pi}} e^{-\frac{(T_z - T_a^*)^2}{2\sigma_a^2}}, \quad \Psi(O_z) = \frac{1}{1 + e^{\gamma(O_z - \widehat{O})}},$$

where  $z$  denotes the depth layer. Since some forage species perform daily vertical migrations (see [Lehodey, 2001](#)), the resulting function depends on conditions in each layer where forage is present during both day and night, i.e.,

$$\Phi_{a,n}(T) = d \cdot \Phi(T_{z^*}) + (1 - d) \cdot \Phi(T_{z^{**}}), \\ \Psi_n(O) = d \cdot \Psi(O_{z^*}) + (1 - d) \cdot \Psi(O_{z^{**}}),$$

where  $d$  is the fraction of the daylight in a day,  $z^*$  and  $z^{**}$  are depth levels at which  $F_n > 0$  at daylight and night correspondingly (see [Lehodey, 2001](#)). Finally, accessibility functions are the products:  $\Theta_{a,n} = \Phi_{a,n}(T) \Psi_n(O)$ .

The temperature function is age-dependent with different optimal temperature  $T_a^*$  and tolerance interval  $\sigma_a$  for each age. They are determined according to existing knowledge about size-dependence of tuna body temperature and tuna heat budget (see [Lehodey, 2001](#)):

$$T_a^* = T_0^* + (T_K^* - T_0^*) \frac{l_a}{l_K}, \quad (\text{A.5})$$

$$\sigma_a = \sigma_T + \frac{w_a}{w_K}, \quad (\text{A.6})$$

where  $l$  and  $w$  are fish fork-length and weight.

### A.1.4. Seasonality in adult habitat index

The seasonal nature of environmental variability has a strong effect on fish reproduction and associated migrations. Changes of daylight length, i.e., the gradient  $\partial_t d$ , can work as a trigger switching tuna behavior from foraging to searching for spawning grounds. One of the hypothesis of how this search occurs assumes that adult tuna tend to direct their movements to find a place with environmental conditions as those occurring during their birth (see

e.g., Cury, 1994). Based on such assumption, a special case of the adult habitat index is

$$I_{SE,a} = \frac{I_{2,a}}{1 + e^{\kappa(\hat{G}_t d - \hat{G})}} + \frac{I_0}{1 + e^{\kappa(\hat{G} - \hat{G}_t d)}}, \quad (\text{A.7})$$

where  $\hat{G}$  is a fixed triggering value ( $=0.035$ ) of the daylight gradient and  $\kappa$  is large constant ( $=1000$ ) producing abrupt but continuous shift between feeding and spawning indexes.

#### A.1.5. Food requirement index for adult tuna

Another special case of adult habitat index considers how food requirements of adult tuna are satisfied in the habitat. Such an index does not govern tuna movement but influences mortality rate imposing “starvation” penalty (Eq. (A.16)). In the previous version of SEAPODYM the same habitat index (Eq. (A.4)) was used to constrain movement and to introduce variability of mortality coefficient, i.e., we assumed that both physical environmental conditions (temperature, oxygen) and food resources have equal impact on fish mortality rate. For simplicity in this study we tested the influence of only food factor on the population mortality. We define the adult food requirement index as the ratio between available forage in the habitat and food required by adult tuna at age  $a$ :

$$I_{FR,a} = \frac{\sum_n F_n}{\psi \sum_n r w_a N_a \vartheta_{a,n}},$$

where  $r$  is the food ration of an individual, i.e., proportion of tuna weight  $w_a$ ,  $\psi$  is a parameter responsible for consumption of forage by other predators and  $\vartheta$  is the relative accessibility coefficient, i.e.,

$$\vartheta_{a,n} = \frac{\Theta_{a,n}}{\sum_n \Theta_{a,n}}.$$

Finally, in order to scale this index between 0 and 1, we use the transformation

$$I_{FR,a} = \frac{1}{1 + I_{FR,a}^v}. \quad (\text{A.8})$$

#### A.2. Movement

Movements of adult tuna consist of two components, namely, random dispersal and directed migrations, described by diffusion and advective term in Eq. (2), respectively. Additionally, migrations can be directed by oceanic currents (passive transport) or by environmental stimuli. In the latter case, as in conventional chemotaxis models (see Keller and Segel, 1971; Czaran, 1998; Turchin, 1998) we determine velocity field of tuna ( $\mathbf{V}_a$ ) as being proportional to the gradient of external stimuli, incorporated into adult habitat index  $I_{2,a}$ :

$$\mathbf{V}_a = \chi_a \left( \frac{\partial I_{2,a}}{\partial x}, \frac{\partial I_{2,a}}{\partial y} \right)^T, \quad (\text{A.9})$$

where the taxis activity constant  $\chi_a$  is proportional to maximal sustainable speed of the fish  $V_{\max,a}$  expressed in the units of body length, which is, in turn, inversely related to the average size at age (Malte et al., 2004) following  $V_{\max,a} = V_m (1 - \eta \frac{1}{l_a})$ , where the parameter  $\eta = 0.1$  implies small negative slope.

Local diffusion coefficients are also linked to the adult habitat index. We define maximal diffusion coefficient in the null (extremely unfavorable) habitat according to the formula of two-dimensional mean square dispersal (see, e.g., Turchin, 1998), namely  $D_{\max} = R^2/4t$ , or if we assume that during time  $t$  individual will cov-

er the maximal distance moving with its maximal sustainable speed  $V_{\max}$ , we have as upper estimate of diffusion coefficient  $D_{\max} = V_{\max}^2 t/4$ . Thus, in each habitat, a given upper value is reduced according to non-linear relationship with the habitat index  $I_{2,a}$  and linear relationship with its gradient,  $\nabla I_{2,a}$ :

$$D_a = D_{\max} \left( 1 - \frac{I_{2,a}}{c + I_{2,a}} \right) (1 - \rho |\nabla I_{2,a}|), \quad (\text{A.10})$$

where  $c$  is the coefficient of variability of fish diffusion rate with habitat index. The expression  $(1 - \rho |\nabla I_{2,a}|)$  with  $\rho < 1$  balances diffusive and advective movements to ensure that maximal displacement due to both diffusion and taxis does not exceed the distance which fish can cover with its maximal sustainable speed.

#### A.3. Spawning

The density of new recruits to the tuna population is given by the product of two functions, the Beverton–Holt relationship giving the dependence on the density of mature adult tuna and  $I_0$ , the spawning habitat index being the function of food to predator ratio and surface layer temperature (Eqs. (A.1) and (A.2)):

$$S_0 = \frac{RN}{1 + bN} \cdot I_0. \quad (\text{A.11})$$

Setting parameter  $b$  to 0 gives us Malthusian growth of population density although still restricted by the habitat conditions.

#### A.4. Mortality

Tuna senescence and predation mortality are functions of age in months,  $\tau$ :

$$m_s(\tau) = \bar{m}_s (1 + e^{\beta_s(\tau-A)})^{-1}, \quad (\text{A.12})$$

$$m_p(\tau) = \bar{m}_p e^{-\tau\beta_p}, \quad (\text{A.13})$$

where  $\bar{m}_s$  and  $\bar{m}_p$  are maximal senescence and predation mortalities,  $\beta_s$  and  $\beta_p$  are slope coefficients, and  $A$  is the age at which  $m_s(A) = \bar{m}_s/2$ . The sum of (A.12) and (A.13) expresses total natural mortality-at-age rate:

$$M(\tau) = m_s(\tau) + m_p(\tau).$$

With added effects of fishing and environmental variability expressed through habitat index functions, the local mortality rates of each cohort are

$$m_0 = M(\tau_0)(1 - I_0 + \varepsilon); \quad (\text{A.14})$$

$$m_k = M(\tau_k)(1 - I_1 + \varepsilon), \quad k = 1, 2; \quad (\text{A.15})$$

$$M_a = M(\tau_a)(1 + e^{2I_{FR,a}^{-1}}) + \sum_f s_{f,a} q_f E_f, \quad a = 1, \dots, K. \quad (\text{A.16})$$

Mortalities of larvae and juveniles can vary in both directions, i.e., if  $I < 0.5$  mortality rate increases and the opposite is true for  $I > 0.5$ . Adult mortality, in contrast, can only increase depending on the food requirement index  $I_{FR,a}$  that determines the level of food deficit for each age group. Such penalty leads to highest local mortality rates for young tunas. The coefficient  $q_f$  is catchability of fishery  $f$ ,  $E_f$  is observed fishing effort and  $s_{f,a}$  is fleet-specific selectivity, which is specified as either sigmoid function (type I selectivity function) of age or asymmetric Gaussian (type II):

$$s_{f,a} = \begin{cases} (1 + e^{-s_f(l_a - \hat{l}_f)})^{-1}, & \text{type I} \\ e^{-\frac{(l_a - \hat{l}_f)^2}{\sigma_{s_f}^2}} & \text{if } l_a \leq \hat{l}, \text{ type II,} \\ \mu_f + (1 - \mu_f) e^{-\frac{(l_a - \hat{l}_f)^2}{\sigma_{s_f}^2}} & \text{if } l_a > \hat{l}, \text{ type II.} \end{cases} \quad (\text{A.17})$$

## Appendix B. Selected notation

N	Symbol	Description	Units
1	$\Omega, x, y$	Two-dimensional model domain with complex boundary and its coordinates	degrees
2	$z$	Vertical layers: (1) 0–100 m, (2) 100–400 m and (3) 400–1000 m	m
<i>Environmental data</i>			
3	$\mathbf{v}_z$	Vector ( $u, v$ ) of horizontal currents, averaged through each vertical layer (GCM modelled data)	Nmi/mo
4	$T_z$	Temperature, averaged through each layer $z$ (GCM data)	°C
5	$O_z$	Concentration of dissolved oxygen, averaged through each vertical layer (Levitus database)	ml/l
6	$P$	Primary production, averaged through 0–400 m depth (obtained from GCM–NPZD coupled model)	mmol C m <sup>-2</sup> mo <sup>-1</sup>
<i>ADR coupled model variables</i>			
7	$F_n$	Density of $n$ th forage component (food for tunas)	g/m <sup>2</sup>
8	$J_k$	Density of juvenile age class $k = 0, 1, 2$ of tuna population	g/m <sup>2</sup>
9	$N_a$	Density of adult age class $a = 1, \dots, K$ of tuna population	g/m <sup>2</sup>
<i>Environmental (habitat) indices</i>			
10	$\Theta_{a,n}$	Accessibility of tuna cohort $a$ to $n$ th forage vertical habitat	
11	$I_0$	Spawning or larvae's habitat index	
12	$I_1$	Juvenile's habitat index	
13	$I_{2,a}$	Adult's (feeding, movement and seasonal migrations) habitat index	
<i>Advection–diffusion–reaction parameters</i>			
14	$V_a$	Vector of velocity of each tuna cohort density	Nmi/mo
15	$D_a$	Diffusion coefficient for each tuna cohort	Nmi <sup>2</sup> /mo
16	$m_S$	Tuna senescence mortality	mo <sup>-1</sup>
17	$m_P$	Tuna predation mortality	mo <sup>-1</sup>
18	$m_F$	Tuna fishing mortality	mo <sup>-1</sup>
19	$s_{f,a}$	Selectivity functions for fishery $f$ and age of tuna $a$	
<i>Optimization variables</i>			
20	$C_f$	Total monthly tuna catch by fishery $f$	10 <sup>3</sup> tonnes
21	$Q_f, r$	Proportion of length frequencies for fishery $f$ and region $r$	
22	$L^-$	Total negative likelihood function	

## References

Arditi, R., Ginzburg, L.R., 1989. Coupling in predator–prey dynamics: ratio-dependence. *J. Theor. Biol.* 139, 311–326.

- Arditi, R., Tyutyunov, Yu., Morgulis, A., Govorukhin, V., Senina, I., 2001. Directed movement of predators and the emergence of density-dependence in predator–prey models. *Theor. Popul. Biol.* 59 (3), 207–221.
- Bard, Y., 1974. *Nonlinear Parameter Estimation*. Academic Press, New York.
- Berezovskaya, F.S., Karev, G.P., 1999. Bifurcations of running waves in population models with taxis. *Usp. Fiz. Nauk.* 9, 1011–1024 (in Russian).
- Berezovskaya, F.S., Davydova, N.V., Isaev, A.S., Karev, G.P., Khlebopros, R.G., 1999. Migration waves and the spatial dynamics of phytophagous insects. *Siber. Ecol. J.* 4, 347–357 (in Russian).
- Bertignac, M., Lehodey, P., Hampton, J., 1998. A spatial population dynamics simulation model of tropical tunas using a habitat index based on environmental parameters. *Fish. Oceanogr.* 7 (3/4), 326–334.
- Brill, R., 1994. A review of temperature and oxygen tolerance studies of tunas pertinent to fisheries oceanography, movement models and stock assessments. *Fish. Oceanogr.* 3 (3), 204–216.
- Cayre, P., 1991. Behavior of yellowfin tuna (*Thunnus albacares*) and skipjack tuna (*Katsuwonus pelamis*) around fish aggregating devices (FADs) in the Comoros Islands as determined by ultrasonic tagging. *Aquat. Living Resour.* 4, 1–12.
- Chen, D., Rothstein, L.M., Busalacchi, A.J., 1994. A hybrid vertical mixing scheme and its application to tropical ocean models. *J. Phys. Oceanogr.* 24, 2156–2179.
- Christian, J., Murtugudde, R., 2003. Tropical Atlantic variability in a coupled physical–biogeochemical ocean model. *Deep Sea Res. II* 50 (22–26), 2947–2969.
- Christian, J., Verschell, M., Murtugudde, R., Busalacchi, A., McClain, C., 2002. Biogeochemical modeling of the tropical Pacific Ocean I: seasonal and interannual variability. *Deep Sea Res. II* 49, 509–543.
- Cury, P., 1994. Obstinate nature: an ecology of individuals: thoughts on reproductive biology and biodiversity. *Can. J. Fish. Aquat. Sci.* 51, 1664–1673.
- Czaran, T., 1998. *Spatiotemporal Models of Population and Community Dynamics*. Chapman and Hall, London.
- Edelstein-Keshet, L., 1988. *Mathematical Models in Biology*. McGraw Hill, New York.
- Faugeras, B., Maury, O., 2005. An advection–diffusion–reaction population dynamics model combined with a statistical parameter estimation procedure: application to the Indian skipjack tuna fishery. *Math. Biosci. Eng.* 2 (4), 1–23.
- Govorukhin, V.N., Morgulis, A.B., Tyutyunov, Yu.V., 2000. Slow taxis in a predator–prey model. *Dokl. Math.* 61, 420–422. Translated from *Dokl. Akad. Nauk.* 372, 730–732.
- Griewank, A., Corliss, G.F., 1991. *Automatic Differentiation of Algorithms: Theory, Implementation, and Application*. SIAM, Philadelphia.
- Grunbaum, D., 1994. *Translating Stochastic Density-dependent Individual Behavior with Sensory Constraints*. Mathematical Models in Biology. McGraw Hill, New York.
- Keller, E., Segel, L.A., 1971. Travelling bands of chemotactic bacteria: a theoretical analysis. *J. Theor. Biol.* 30, 235–248.
- Langley, A., Hampton, J., Ogura, M., 2005. Stock assessment of skipjack tuna in the western and central Pacific Ocean. Western and Central Pacific Fisheries Commission, Scientific Committee, SA WP4. <[http://wcpfc.org/sc1/pdf/SC1\\_SA\\_WP\\_4.pdf](http://wcpfc.org/sc1/pdf/SC1_SA_WP_4.pdf)>.
- Lehodey, P., 2001. The pelagic ecosystem of the tropical Pacific Ocean: dynamic spatial modeling and biological consequences of ENSO. *Progr. Oceanogr.* 49, 439–468.
- Lehodey, P., 2004a. A Spatial Ecosystem and Populations Dynamics Model (SEAPODYM) for tuna and associated oceanic top-predator species: Part I. Lower and intermediate trophic components. In: 17th Meeting of the Standing Committee on Tuna and Billfish, Majuro, Republic of Marshall Islands, 9–18 August 2004. Oceanic Fisheries Programme, Secretariat of the Pacific Community, Noumea, New Caledonia. Working Paper: ECO-1, 26 pp. <<http://www.spc.int/OceanFish/Html/SCTB/SCTB17/ECO-1.pdf>>.
- Lehodey, P., 2004b. A Spatial Ecosystem and Populations Dynamics Model (SEAPODYM) for tuna and associated oceanic top-predator species: Part II. Tuna populations and fisheries. In: 17th Meeting of the Standing Committee on Tuna and Billfish, Majuro, Republic of Marshall Islands, 9–18 August 2004. Oceanic Fisheries Programme, Secretariat of the Pacific Community, Noumea, New Caledonia. Working Paper: ECO-2, 36 pp. <<http://www.spc.int/OceanFish/Html/SCTB/SCTB17/ECO-2.pdf>>.
- Lehodey, P., Bertignac, M., Hampton, J., Lewis, A., Picaut, J., 1997. El Niño Southern Oscillation and tuna in the western Pacific. *Lett. Nature* 389, 715–718.
- Lehodey, P., Chai, F., Hampton, J., 2003. Modelling climate-related variability of tuna populations from a coupled ocean biogeochemical–populations dynamics model. *Fish. Oceanogr.* 12 (4/5), 483–494.
- Malte, H., Larsen, C., Musyl, M.K., Brill, R., 2004. Differential heating and cooling rates in bigeye tuna (*Thunnus obesus*): options for cardiovascular adjustments. *Abst. Comp. Biochem. Physiol. A* 137, S51.
- Matear, R.J., 1995. Parameter optimization and analysis of ecosystem models using simulated annealing: a case study at Station P. *J. Mar. Res.* 53, 571–607.
- McKinley, G. et al., 2006. North Pacific carbon cycle response to climate variability on seasonal to decadal timescales. *J. Geophys. Res.* C 7, C07S06.
- Murray, J.D., 1989. *Mathematical Biology*. Springer-Verlag, New York. 767 p.
- Murtugudde, R., Seager, R., Busalacchi, A.J., 1996. Simulation of the tropical oceans with an ocean GCM coupled to an atmospheric mixed-layer model. *J. Climate* 9, 1795–1815.
- Nihira, A., 1996. Studies on the behavioral ecology and physiology of migratory fish schools of skipjack tuna (*Katsuwonus pelamis*) in the oceanic frontal area. *Bull. Tohoku Natl. Fish. Res. Inst.* 58, 137–233.
- Okubo, A., 1980. *Diffusion and Ecological Problems: Mathematical Models*. Springer-Verlag, New York.

- Otter Research Ltd., 1994. Autodif: a C++ array extension with automatic differentiation for use in nonlinear modeling and statistics. Otter Research Ltd., Nanaimo, Canada.
- Petrovskii, S., Li, B., 2001. Increased coupling between subpopulations in a spatially structured environment can lead to populational outbreaks. *J. Theor. Biol.* 212, 549–562.
- Petrovskii, S., Morozov, A., Venturino, E., 2002. Allee effect makes possible patchy invasion in a predator–prey system. *Ecol. Lett.* 5 (3), 345–352.
- Robinson, A.R., Lermusiaux, P.F.J., 2002. Data assimilation for modeling and predicting coupled physical biological interactions in the sea. In: Robinson, A.R., McCarthy, J.J., Rothschild, B.J. (Eds.), *The Sea*, vol. 12. John Wiley, New York, pp. 475–536.
- Senina, I., Tyutyunov, Yu., Arditì, R., 1999. Extinction risk assessment and optimal harvesting of anchovy and sprat in the Azov Sea. *J. Appl. Ecol.* 36, 297–306.
- Sibert, J., Hampton, J., 2003. Mobility of tropical tunas and the implications for fishery management. *Mar. Policy* 27, 87–95.
- Sibert, J.R., Hampton, J., Fournier, D.A., Bills, P.J., 1999. An advection–diffusion–reaction model for the estimation of fish movement parameters from tagging data, with application to skipjack tuna (*Katsuwonus pelamis*). *Can. J. Fish. Aquat. Sci.* 56, 925–938.
- Sibert, J., Hampton, J., Kleiber, P., Maunder, M., 2006. Biomass, size, and trophic status of top predators in the Pacific Ocean. *Science* 314, 1773–1776.
- Skellam, J.G., 1951. Random dispersal in theoretical populations. *Biometrika* 38, 196–218.
- Turchin, P., 1998. *Quantitative Analysis of Movement*. Sinauer, Sunderland, MS.
- Tyutyunov, Yu., Senina, I., Arditì, R., 2004. Clustering due to acceleration in the response to population gradient: a simple self-organization model. *The Am. Nat.* 164 (6), 722–735.
- Vallino, J.J., 2000. Improving marine ecosystem models: use of data assimilation and mesocosm experiments. *J. Mar. Res.* 58, 117–164.
- Wang, X., Christian, J.R., Murtugudde, R., Busalacchi, A.J., 2005. Ecosystem dynamics and export production in the central and eastern equatorial Pacific: a modeling study of impact of ENSO. *Geophys. Res. Lett.* 32, L2608.
- Wang, X., Christian, J.R., Murtugudde, R., Busalacchi, A.J., 2006. Spatial and temporal variability in new production in the equatorial Pacific during 1980–2003: physical and biogeochemical controls. *Deep Sea Res. II* 53, 677–697.
- Worley, B., 1991. Experience with the forward and reverse mode of GRESS in contaminant transport modeling and other applications. In: Griewank, A., Corliss, G.F. (Eds.), *Automatic Differentiation of Algorithms: Theory, Practice and Application*. SIAM, Philadelphia.