APPENDIX E. MODEL EQUATIONS

INTRODUCTION

Stock Assessment modelling was conducted using the Integrated Statistical Catch Age Model (iSCAM), developed by Steven J.D. Martell (Martell et al. 2011). iSCAM is written in AD Model Builder and the source code and documentation for the original iSCAM is available at https://github.com/smartell/iSCAM. iSCAM uses a statistical catch-at-age model implemented in a Bayesian estimation framework.

The version of iSCAM used for this assessment is based on a more recent IPHC-developers version, also developed by S. Martell, that models males and females separately and includes a length-based option for modelling selectivity (also available at https://github.com/smartell/iSCAM). Further modifications to the IPHC-developers branch have been made by the first author of this assessment (K.R. Holt) to allow density-dependent natural mortality and non-linear catchability options for Strait of Georgia Lincgod (see https://github.com/krholt/iSCAM-pbs-lingcod).

Running of iSCAM and compiliation of results figures was streamlined using the iscam-gui software package developed by Chris Grandin (pers. comm, Pacific Biological Station, Fisheries and Oceans Canada). iscam-gui is written in R (R Development Core Team, 2012), and provides an R gui interface that allows users to run and show output of multiple iSCAM model runs next to each other.

This appendix contains the documentation in mathematical form of the underlying iSCAM age-structured model, and its steady state version that is used to calculate reference points, the observation models used in predicting observations, and the components of the objective function that formulate the the objective function that is used to estimate model parameters. All of the model equations are laid out in tables and are intended to represent the order of operations, or pseudocode, in which to implement the model. A documented list of symbols used in model equations is given in Table E.1. The documentation presented here is a revised version of the iscam user-guide written by S. Martell and available at https://github.com/smartell/iSCAM. Much of the text and equations have been taken directly from the original users guide, with modifications by K.R. Holt to describe the expansion to a two-sex model and the additional lingcod options.

ANALYTIC METHODS: EQUILIBRIUM CONSIDERATIONS

A Steady-State Age-Structured Model

For the steady-state conditions represented in Table E.2, we assume the parameter vector Θ in (E.1) is unknown and would eventually be estimated by fitting iSCAM to data. For a given set of sex-specific growth parameters and maturity-at-age parameters defined by (E.2), growth is assumed to follow von Bertalanffy (E.3), mean weight-at-age is given by the allometric relationship in (E.4), and the age- and sex-specific vulnerability is given by a length-based logistic function (E.5). Note, there are alternative selectivity functions implemented in iSCAM; the length-based logistic function is shown here because it is used to model selectivity to the only current lingcod fishery in the Strait of Georgia (the recreational fishery with a 65cm minimum size limit) . Mean fecundity-at-age by sex is assumed to be proportional to the mean weight-at-age of mature fish, where maturity at age is specified by the sex-specific parameters $\dot{a}_s$ and $\dot{\gamma}_s$ for the logistic function.

*Table E.1 A list of symbols, constants and description for variables used in iSCAM. The symbol * after a value i*

Symbol	Value	Description
Indexes		
s		Index for sex
a		Index for age
t		Index for year
k		Index for gear
Model dimensions		
S	2	Number of sexes
$\acute{a}, A$	1, 20	Youngest and oldest age class (A is a plus group)
$\acute{t}, T$	1927, 2013	First and last year of catch data
K	5	Number of gears including survey gears
Observations (data)		
$C_{k,t}$		catch in weight by gear k in year t
$I_{k,t}$		relative abundance index for gear k in year t
Fixed parameters		
M	0.20*	Instantaneous natural mortality rate
ρ	0.98*	Fraction of the total variance associated with observation error
$\hat{a}_k, \hat{\gamma}_k$	See Table E.4	Selectivity parameters for gear k
Estimated parameters		
R_o		Age-$\acute{a}$ recruits in unfished conditions
κ		Recruitment compensation
$\bar{R}$		Average age-$\acute{a}$ recruitment from year $\acute{t}$ to T
ϑ		Total precision (inverse of variance) of the total error
$\Gamma_{k,t}$		Logarithm of the instantaneous fishing mortality for gear k in year t
ω_t		Age-$\acute{a}$ deviates from $\bar{R}$ for years $\acute{t}$ to T
Standard deviations		
σ		Standard deviation for observation errors in survey index
τ		Standard deviation in process errors (recruitment deviations)
σ_C	0.25	Standard deviation in observed catch by gear
Residuals		
δ_t		Annual recruitment residual
η_t		Residual error in predicted catch
Growth & maturity parameters		
$l_{\infty s}$	900 / 1040	Asymptotic length in mm for males / females
$\acute{k}_s$	0.20 / 0.20	Brody growth coefficient for males / females
t_{os}	-0.001 / -0.001	Theoretical age at zero length for males / females
$\acute{a}_s$	1.13e-11 / 1.13e-11	Scaler in length- weight allometry (mm to 100's of kg)
$\acute{b}_s$	3.329 / 3.329	Power parameter in length-weight allometry for males / females
$\dot{a}_s$	2.0 / 5.0	Age at 50% maturity for males / females
$\acute{\gamma}_s$	0.001 / 0.20	Standard deviation at 50% maturity for males / females

Table E.2. Steady-state age-structured model assuming unequal vulnerability-at-age, age-specific fecundity and Ricker type recruitment.

Parameters

$$\Theta = (B_o, \kappa)\,; \quad B_o > 0; \kappa > 1 \tag{E.1}$$

$$\Phi = (l_{\infty s}, \acute{k}_s, t_{os}, \acute{a}_s, \acute{b}_s, \dot{a}_s, \dot{\gamma}_s, \hat{a}, \hat{\gamma}, M) \tag{E.2}$$

Age-schedule information

$$l_{a,s} = l(1 - \exp(-k_s(a - t_{os}))) \tag{E.3}$$

$$w_{a,s} = \acute{a}_s(l_{a,s})^{\acute{b}_s} \tag{E.4}$$

$$v_{a,s} = (1 + \exp(-(\hat{a} - l_{a,s})/\hat{\gamma}))^{-1} \tag{E.5}$$

$$f_{a,s} = w_{a,s}(1 + \exp(-(\dot{a}_s - a)/\dot{\gamma}_s))^{-1} \tag{E.6}$$

Survivorship

$$\iota_a = \begin{cases} 1/S, & a = 1 \\ \iota_{a-1}e^{-M}, & a > 1 \\ \iota_{a-1}/(1 - e^{-M}), & a = A \end{cases} \tag{E.7}$$

.

$$\hat{\iota}_{a,s} = \begin{cases} 1/S, & a = 1 \\ \hat{\iota}_{a-1,s}e^{-M-F_e v_{a-1,s}}, & a > 1 \\ \hat{\iota}_{a-1,s}e^{-M-F_e v_{a-1,s}}/(1 - e^{M-F_e v_{a,s}}), & a = A \end{cases} \tag{E.8}$$

Incidence functions

$$\phi_E = \sum_{s=1}^{S}\sum_{a=1}^{\infty} \iota_a f_{a,s}, \quad \phi_e = \sum_{s=1}^{S}\sum_{a=1}^{\infty} \hat{\iota}_{a,s} f_{a,s} \tag{E.9}$$

$$\phi_B = \sum_{s=1}^{S}\sum_{a=1}^{\infty} \iota_a w_{a,s} v_{a,s}, \quad \phi_b = \sum_{s=1}^{S}\sum_{a=1}^{\infty} \hat{\iota}_{as} w_{a,s} v_{a,s} \tag{E.10}$$

$$\phi_q = \sum_{s=1}^{S}\sum_{a=1}^{\infty} \frac{\hat{\iota}_{a,s} w_{a,s} v_{a,s}}{M + F_e v_{a,s}} \left(1 - e^{(-M-F_e v_{a,s})}\right) \tag{E.11}$$

Steady-state conditions

$$R_o = B_o/\phi_B \tag{E.12}$$

$$R_e = R_o \frac{\ln(\kappa) - \ln(\phi_E/\phi_e)}{\ln(\kappa)} \tag{E.13}$$

$$C_e = F_e R_e \phi_q \tag{E.14}$$

Survivorship for unfished and fished populations is defined by (E.7) and (E.8), respectively. Note that fished survivorship is sex-specifric to allow for sex-specific $v_{a,s}$ when the length-based logistic function is used to model vulnerability. It is assumed that all individuals ages A and older (i.e., the plus group) have the same total mortality rate. The incidence functions refer to the life-time or per-recruit quantities such as spawning biomass per recruit (ϕ_E) or vulnerable biomass per recruit (ϕ_b). Note that upper and lower case subscripts denote unfished and fished conditions, respectively. Spawning biomass per recruit is given by (E.9), the vulnerable biomass per recruit is given by (E.10) and the per recruit yield to the fishery is given by (E.11). Unfished recruitment is given by (E.12) and the steady-state equilibrium recruitment for a given fishing mortality rate F_e is given by (E.13). Note that in (E.13) we assume that recruitment follows a Ricker stock recruitment model of the form:

$$R_e = s_o R_e \phi_e \exp(-\beta R_e \phi_e)$$

where the maximum juvenile survival rate is given by:

$$s_o = \kappa / \phi_E,$$

and the density-dependent term is given by:

$$\beta = \frac{\ln(\kappa)}{R_o \phi_E},$$

which simplifies to (E.13).

The equilibrium yield for a given fishing mortality rate is (E.14). These steady-state conditions are critical for determining various reference points such as F_{MSY} and B_{MSY}.

MSY-based Reference Points

When defining reference points for this assessment, only the current recreational fishery (with a 65cm size limit) was used to calculate MSY quantities. In the case of a single fishery such as this, iSCAM calculates F_{MSY} by finding the value of F_e that results in the zero derivative of (E.14). This is accomplished numerically using a Newton-Raphson method where an initial guess for F_{MSY} is set equal to 1.0M. Given an estimate of F_{MSY}, other reference points such as MSY are calculated using the equations in Table E.2.

A special class library to implement the MSY-based reference point calculations was developed specifically for iSCAM. Details of this alogorithim, including partial derivatives for the numerical calculation of F_{MSY}, are available in the original iscam documentation available at: https://github.com/smartell/iSCAM.

ANALYTIC METHODS: STATE DYNAMICS

The estimated parameter vector in iSCAM is defined in (E.15) of Table E.3. The unknown parameters R_0 and κ, as well as the fixed parameter M, are the leading population parameters that define the overall population scale. The total variance ϑ^2 is estimated, while the proportion of the total variance that is associated with observation errors ρ is assumed fixed. The total variance is partitioned into observation errors (σ^2) and process errors (τ^2) using (E.16).

The unobserved state variables (E.17) include the numbers-at-age in year t of sex s ($N_{t,a,s}$), the spawning stock biomass in year t of sex s (B_t), and the total age- and sex-specific total mortality rate ($Z_{t,a,s}$).

The initial numbers-at-age in the first year (E.18) and the annual recruits (E.19) are treated as estimated parameters and used to initialize the numbers-at-age array. Recruitment at age-1 is assumed to be 50% males and 50

State variables in each year are updated using equations E.22–E.25, where the spawning biomass is the product of the numbers-at-age and the mature biomass-at-age (E.22). The total mortality rate is given by (E.23), and the total catch (in weight) for each gear is given by (E.24), assuming that both natural and fishing mortality occur simultaneously throughout the year. Lingcod catch was not differentiated by sex, so both sexes were combined to calculate a total catch in (E.24). The sex-specific numbers-at-age are propagated over time using (E.25), where members of the plus group (age A) are all assumed to have the same total mortality rate.

Recruitment to age k is assumed to follow a Ricker model for Strait of Gerogia Lingcod (E.26) where the maximum juvenile survival rate (s_o) is defined by $s_o = \kappa/\phi_E$. For the Ricker model, β is derived by solving (E.26) for β conditional on estimates of κ and R_o:

$$\beta = \frac{\ln(\kappa)}{R_o \phi_E}$$

Option for Density-Dependent Natural Mortality

The option to specify density-dependent mortality was added to the version of iSCAM used for this assessment to allow replication of the 2005 assessment for the bridging analysis. Density-dependent mortality was implemented in the same way as the 2005 assessment (Logan et al. 2005):

$$M_t = M_0 + (M_{zero} - M_0)(1 - \frac{B_t}{B_0})$$

where, B_t is total biomass (males and females combined) in year t, M_t is the natural mortality at stock biomass B_t, M_0 is natural mortality at unfished biomass (i.e., carrying capacity), and M_{zero} is natural mortality at negligible stock size.

Both M_0 and M_zero were fixed at the same parameter values assumed by Logan et al. (2005). M_0 was set at 0.2 and M_zero was set at 0.18.

Need to add chages to reference point calculations .....

Table E.3.Statistical catch-age model using the Baranov catch.

Estimated parameters

$$\Theta = (R_0, \kappa, \bar{R}, \vartheta^2, \Gamma_{k,t}, \{\omega_t\}_{t=1-A}^{t=T}) \tag{E.15}$$

$$\sigma^2 = \rho/\vartheta^2, \quad \tau^2 = (1-\rho)/\vartheta^2 \tag{E.16}$$

Unobserved states

$$N_{t,a},\ B_t,\ Z_{t,a} \tag{E.17}$$

Initial states

$$N_{t,a,s} = \frac{1}{S}\bar{R}e^{\omega_{t-a}}\exp(-M)^{(a-1)}; \quad t=1; 2 \leq a \leq A \tag{E.18}$$

$$N_{t,a,s} = \frac{1}{S}\bar{R}e^{\omega_t}; \quad 1 \leq t \leq T; a=1 \tag{E.19}$$

$$v_{k,a,s} = \frac{1}{1+exp(-(l_{a,s}-\hat{a}_k)/\hat{\gamma}_k)} \tag{E.20}$$

$$F_{k,t} = \exp(\Gamma_{k,t}) \tag{E.21}$$

State dynamics (t $>$ 1)

$$B_{t,s} = \sum_a N_{t,a,s} f_{a,s} \tag{E.22}$$

$$Z_{t,a,s} = M + \sum_k F_{k,t} v_{k,t,a,s} \tag{E.23}$$

$$\hat{C}_{k,t} = \sum_s \sum_a \frac{N_{t,a,s} w_{a,s} F_{k,t} v_{k,t,a,s} \left(1-e^{-Z_{t,a,s}}\right)^{\eta_t}}{Z_{t,a,s}} \tag{E.24}$$

$$N_{t,a,s} = \begin{cases} N_{t-1,a-1,s}\exp(-Z_{t-1,a-1,s}) & a>1 \\ N_{t-1,a,s}\exp(-Z_{t-1,a,s}) & a=A \end{cases} \tag{E.25}$$

Recruitment model

$$R_t = s_o B_{t-k} e^{-\beta B_{t-k}+\delta_t-0.5\tau^2} \tag{E.26}$$

Table E.4. Definition of datasets denoted by gear index k in Tables E.1 - E.3 .

k	Dataset	Years with Catch Data	Years with CPUE Index	Selectivity Type	Selectivity Parameters($\hat{a}, \hat{\gamma}$)
1	Commercial fishery	1927-1989	1962-1989	Age-based	(4.45, 0.2)
2	Rec. fishery; limit $\leq$ 580mm	1962-1990	1982-2013	Age-based	(2.00, 0.2)
3	Rec. fishery; limit = 650mm	1991-2013	-	Length-based	(650, 15)

RESIDUALS, LIKELIHOODS, AND OBJECTIVE FUNCTION VALUE COMPONENTS

There are three major components to the overall objective function that are minimized while iSCAM is performing maximum likelihood estimation. These components consist of the likelihood of the data, prior distributions and penalty functions that are invoked to regularize the solution during intermediate phases of the non-linear parameter estimation. This section discusses each of these in turn, starting first with the residuals between observed and predicted states followed by the negative loglikelihood that is minimized.

Catch Data

It is assumed that the measurement errors in the catch observations are log-normally distributed, and the residuals given by:

$$\eta_{k,t} = \ln(C_{k,t} + o) - \ln(\hat{C}_{k,t} + o), \tag{E.27}$$

where o is a small constant (1.e-10) to ensure the residual is defined in the case of a 0 catch observation. The residuals are assumed to be normally distributed with a user-specified standard deviation σ_C. At present, it is assumed that observed catches for each gear k have the same standard deviation. The negative loglikelihood (ignoring the scaling constant) for the catch data is given by:

$$\ell_C = \sum_k \left[T_k \ln(\sigma_C) + \frac{\sum_t (\eta_{k,t})^2}{2\sigma_C^2} \right], \tag{E.28}$$

where T_k is the total number of catch observations for gear type k.

Commercial fishery catch data for Area 4B Lingcod are available in biomass units (tonnes), while catch from recreational fisheries are reported as numbers (pieces). iSCAM allows users to specify catch units as biomass or numbers; however, to accomodate the constraint that catches from all gears have the same standard deviation, we re-scale recreational catch to 10's of pieces so that it would be on the same scale as catch in tonnes from the commercial fishery. This re-scaling required us to specifythe parameters for the allometric weight-length relationship in (E.4) in units of mm to 100's of kg.

Relative Abundance Data

The relative abundance data are assumed to be proportional to biomass that is vulnerable to the sampling gear:

$$V_{k,t} = \sum_s \sum_a N_{t,a,s} e^{-\lambda_{k,t} Z_{t,a,s}} v_{k,a,s} w_{a,s}, \tag{E.29}$$

where $v_{k,a,s}$ is the sex- and age-specific selectivity of gear k, and $w_{a,s}$ is the mean-weight-at-age for sex s. A user-specified fraction of the total mortality $\lambda_{k,t}$ adjusts the numbers-at-age to correct for survey timing. We set $\lambda_{k,t}$ to 0.5 for all index observations in the commercial and recreational CPUE indices since fishery catch and the collection of CPUE data are the same process, and natural mortality occures throughout the fishing season.

The residuals between the observed and predicted relative abundance index is given by:

$$\epsilon_{k,t} = \ln(I_{k,t}) - \ln(q_k) + \ln(V_{k,t}), \tag{E.30}$$

where $I_{k,t}$ is the observed relative abundance index, q_k is the catchability coefficient for index k, and $V_{k,t}$ is the predicted vulnerable biomass at the time of sampling. The catchability coefficient q_k is evaluated at its conditional maximum likelihood estimate:

$$q_k = \frac{1}{N_k} \sum_{t \in I_{k,t}} \ln(I_{k,t}) - \ln(V_{k,t}),$$

where N_k is the number of relative abundance observations for index k (see Walters and Ludwig 1994 for more information). The negative loglikelihood for relative abundance data is given by:

$$\ell_I = \sum_k \sum_{t \in I_{k,t}} \ln(\sigma_{k,t}) + \frac{\epsilon^2_{k,t}}{2\sigma^2_{k,t}} \tag{E.31}$$

where

$$\sigma_{k,t} = \frac{\rho\varphi^2}{\omega_{k,t}},$$

where $\rho\varphi^2$ is the proportion of the total error that is associated with observation errors, and $\omega_{k,t}$ is a user specified relative weight for observation t from gear k. The $\omega_{k,t}$ terms allow each observation to be weighted relative to the total error $\rho\varphi^2$; for example, to omit a particular observation, set $\omega_{k,t} = 0$, or to give 2 times the weight, then set $\omega_{k,t} = 2.0$. For the current assessment, we assumed all observations have the same variance by setting $\omega_{k,t} = 1$. Note that if $\omega_{k,t} = 0$ then equation (E.31) is undefined; therefore, iSCAM adds a small constant to $\omega_{k,t}$ (1.e-10), which is equivalent to assuming an extremely large variance) to ensure the likelihood can be evaluated.

Option for Non-linear CPUE Relationship

The option to model relative abundance indices as a power function of vulnerable biomass was added to the version of iSCAM used for this assessment to allow replication of the approach taken in the last Strait of Gerogia Lingcod assessment (Logan et al. 2005). This approach was taken in 2005 because the qualified commercial CPUE index used as an abunance index was not expected to be linearly related to abundance. In this case, (E.29) is replaced by:

$$V_{k,t} = (\sum_s \sum_a N_{t,a,s} e^{-\lambda_{k,t} Z_{t,a,s}} v_{k,a,s} w_{a,s})^{\psi_k} \tag{E.32}$$

where ψ_k is an exponent determining the degree of linearity between the CPUE series from gear k and the vulnerable biomass available to gear k in year t. When ψ_k is less than 1, a given change in CPUE implies a greater relative change in exploitable abundance (hyperstability); when ψ_k = 1, CPUE is proportional to abundance; and when ψ_k is greater than 1, a given change in CPUE implies a lesser relative chance in abundance (hyperdepletion). The ψ_k

Table E.5. Fixed values used for exponent determining degree of linearity between vulnerable biomass and commercial CPUE (cCPUE) or recreational CPUE (rCPUE) series (ψ_k in (E.32)*) when non-linear CPUE option selected. Values are given for three different scenarios about the treatment of historic catch values in SubAreas 28 and 29 (i.e., the southeast quadrant).*

Scenario	ψ_{cCPUE}	ψ_{rCPUE}
SoG	0.878	1.759
noSE	0.646	0.611
halfSE	0.878	1.759

parameters for each gear k were estimated by Logan et al. (2005); however, we did not attempt to estimate these in our bridging analysis due to concerns of model overparameterization. Instead, ψ_k was fixed at the values estimated by Logan et al. (2005) in scenarios that used a non-linear CPUE relationship (Table E.5).

Stock-Recruitment

Annual recruitment and the initial age-composition are treated as latent variables in iSCAM, and residuals between estimated recruits and the deterministic stock-recruitment models are used to estimate unfished spawning stock biomass and recruitment compensation. The residuals between the estimated and predicted recruits is given by

$$\delta_t = \ln(\bar{R}e^{w_t}) - f(B_{t-y}) \tag{E.33}$$

where $f(B_{t-y})$ is given by (E.26), and y is the age at recruitment, which is set to 1 for this assessment. Note that a bias correction term for the lognormal process errors is included in (E.26).

The negative log likelihood for the recruitment deviations is given by the normal density (ignoring the scaling constant):

$$\ell_\delta = n\ln(\tau) + \frac{\sum_{t=1+k}^{T} \delta_t^2}{2\tau^2} \tag{E.34}$$

Equations (E.33) and (E.34) are key for estimating unfished spawning stock biomass and recruitment compensation via the recruitment models. The relationship between (s_o, β) and (B_o, κ) for the Ricker stock recruitment model is defined as:

$$s_o = \kappa/\phi_E \tag{E.35}$$

$$\beta = \frac{\ln(\kappa)}{B_o} \tag{E.36}$$

where s_o is the maximum juvenile survival rate, and β is the density effect on recruitment.

BAYESIAN ANALYSIS OF MODEL PARAMETERS & POLICY PARAMETERS

Bayesian estimation was done using the Markov Chain Monte Carlo (MCMC) method to approximate posterior distributions for estimated parameters. Marginal posterior distributions of each model parameter were constructed by using the metropolis algorithm built into ADMB to sample from the joint posterior distribution. This was accomplished by running iSCAM in -mcmc mode followed by the -mceval option. Prior distributions, estimation bounds, and initial values for

Table E.6. Details for estimation of parameters, including prior distributions with corresponding means and standard deviations, bounds between which parameters are constrained, and initial values.

Parameter	Prior distribution	Mean, standard deviation	Bounds	Initial value
SoG				
R_0	Uniform	–	[0.5, 3 300 000]	1 600 000
h	Normal	1.2, 10	[0.2, 208]	0.70
$\bar{R}$	Uniform	–	[0.5, 3 300 000]	89 000
ϑ	Gamma	20, 10	[0.1, 80]	20
noSE				
halfSE				

the MCMC procedure are shown in Table E.6. Marginal posterior densities were also produced for derived quantities such as MSY-based reference points using the steady-state age structured model described in Table E.2 and the associated text above.