Model equations for the SEFRA seabird model

Charles Edwards, Tom Peatman, Dave Goad, Darcy Webber

27 September 2021

Equations needed to represent the Spatially Explicit Fisheries Risk Assessment framework (MPI, 2017), are described. A list of model terms is given in Table 1. This is `sra` version 0.0.2.

Numbers available to fishing

The number of adults per species (s) is defined using the number of breeding pairs:

$$N_s^{\text{adults}} = 2 \cdot \frac{N_s^{\text{BP}}}{P_s^{\text{B}}}$$

The number of adults available to fishing within the New Zealand EEZ, per species (s) and month (m) is:

$$N_{s,m}^{\text{avail. adults}} = N_s^{\text{adults}} \cdot P_{s,m}^{\text{in eez}} \cdot (1 - P_s^{\text{B}} \cdot P_{s,m}^{\text{on nest}})$$

Outside of the breeding season $P_{s,m}^{\text{on nest}} = 0$ and all adults within the EEZ are available to fishing. The number of breeding pairs (N_s^{BP}) and probability of breeding (P_s^{B}) are estimated with informed priors during the model fit (biological input priors are listed in the R-package: sraInputsBio). The probability in zone ($P_{s,m}^{\text{in eez}}$) and probability of being on nest ($P_{s,m}^{\text{on nest}}$) are currently fixed on input.

Spatial distribution and overlap

The spatial distribution of the species is described using a density term $d_{s,m,x}$, which is the number of individuals of species s, per kilometre squared, within raster grid x, during month m. Calculation of this density from the raw input data is described in sraInputsBio). The density is assumed constant across years and normalised to sum to one across grids. The value $d_{s,m,x}$ can be thought of as the binomial sampling probability of an individual being in grid x. The expected number of individuals at that location is therefore:

$$N_{s,m,x}^{\text{avail. adults}} = d_{s,m,x} \cdot N_{s,m}^{\text{avail. adults}}$$

If each fishing event i is allocated to a unique grid x, and assuming a uniform distribution of individuals within the grid and that individuals do not move, then $p_{i,s,m}$ is the value of $d_{s,m,x}$ at the location of fishing event i:

$$p_{i,s,m} = d_{s,m,x} \qquad \text{for} \quad i \in x$$

The overlap for fishing event i is defined as:

$$\text{overlap}_{i,s,m} = \text{effort}_{i,m} \cdot p_{i,s,m}$$

and the density overlap:

$$\underbrace{\text{density overlap}_{i,s,m}}_{\mathbb{O}_{i,s,m}} = \underbrace{\text{effort}_{i,m}}_{a_{i,m}} \cdot p_{i,s,m} \cdot \underbrace{N^{\text{avail. adults}}_{s,m}}_{\mathbb{N}_{s,m}}$$

for which we introduce the notation $\mathbb{O}_{i,s,m}$, $a_{i,m}$ and $\mathbb{N}_{s,m}$ (MPI, 2017). The density overlap is the expectation from a binomial distribution, with sampling probability $p_{i,s,m}$ and number of samples equal to $a_{i,m} \cdot \mathbb{N}_{s,m}$. It is proportional to the number of encounters between birds of species s and fishing event i. But because effort units are arbitrary the density overlap cannot be compared between fishing methods that use different units (meaning that vulnerabilities can also not be compared).

Effort is aggregated by fishery group j, and fishing method k, so that fishing event i is a subset of fishing group j, and fishing group j is a subset of fishing method k (i.e. $i \in j \in k$), adopting the convention that only the subscript of the largest group is reported. The density overlap for fishing group j is the summation across events:

$$\mathbb{O}_{j,s,m} = \sum_{i \in j} a_{i,m} \cdot p_{i,s,m} \cdot \mathbb{N}_{s,m}$$

which is specific to the species s and month m.

Expected captures and deaths

The capture probability per unit of density overlap is described by the vulnerability $\upsilon_{j,s}$, which is defined at the level of the fishing group j and species s. The total captures for fishing group j, per species and month is expected to be:

$$\underbrace{\text{captures}_{j,s,m}}_{C_{j,s,m}} = \upsilon_{j,s} \cdot \mathbb{O}_{j,s,m}$$

noting calculation of the density overlap using all known fishing effort per fishery group, $a_{j,m}$. The observable captures per month are calculated using $p^{obs}_{k,z}$, which is the probability of observing a capture given that an observer is present and the capture has occurred (equal to the inverse of the cryptic mortality multiplier $\kappa^{-1}_{k,z}$). The product of $\upsilon_{j,s}$ and $p^{obs}_{k,z}$ gives the catchability $q_{j,s}$. The number of observable captures is therefore expected to be:

$$\begin{aligned} \text{observable captures}_{j,s,m} &= \upsilon_{j,s} \cdot \mathbb{O}_{j,s,m} \cdot p^{obs}_{k,z} \\ &= q_{j,s} \cdot \mathbb{O}_{j,s,m} \end{aligned}$$

Finally, we use the observed fishing intensity $a'_{j,m}$ and associated observed density overlap $\mathbb{O}'_{j,s,m}$ to calculate the expected number of observed captures:

$$\underbrace{\text{observed captures}_{j,s,m}}_{C'_{j,s,m}} = q_{j,s} \cdot \mathbb{O}'_{j,s,m}$$

The $p^{obs}_{k,z}$ is currently partitioned according to the fishing method and species group, although this is under review. The probability of surviving capture is defined using the parameter $\Psi_{j,z}$, which represents the probability of a live capture at the level of the fishing group j and species group z (with s a subset of z), and $\omega_{j,z}$, which is the probability of a live capture surviving post-release. The probability of death is $\Psi_{j,z} \cdot (1 - \omega_{j,z}) + (1 - \Psi_{j,z})$, which can be simplified to predict the number of deaths and observed deaths:

$$\underbrace{\text{deaths}_{j,s,m}}_{D_{j,s,m}} = \underbrace{\text{captures}_{j,s,m}}_{C_{j,s,m}} \cdot (1 - \Psi_{j,z} \cdot \omega_{j,z})$$

$$\underbrace{\text{observed deaths}_{j,s,m}}_{D'_{j,s,m}} = \underbrace{\text{observed captures}_{j,s,m}}_{C'_{j,s,m}} \cdot (1 - \Psi_{j,z} \cdot \omega_{j,z})$$

Capture type

It is possible to separate the catchability according to the capture type. If we use the notation $C_{t,j,s,m}$ to refer to captures of a specific type t by fishery group j, with total captures by that group equal to the sum across types: $C_{j,s,m} = \sum_t C_{t,j,s,m}$; then we can write the catch equation as:

$$C_{j,s,m} = q_{j,s} \cdot \mathbb{O}_{j,s,m}$$
$$C_{1,j,s,m} + C_{2,j,s,m} + \ldots = q_{j,s} \cdot (\pi_1 + \pi_2 + \ldots) \cdot \mathbb{O}_{j,s,m}$$

with $\sum_t \pi_t = 1$. The probability of observing a capture type t by fishery group j then becomes the product $\pi_t \cdot q_{j,s}$.

Being able to represent the capture type within each fishery group is important to account for changing fishing practices. For example, changes in the trawl fishery have changed the proportion of captures that occur on the warp or on the net. It also allows the cryptic mortality multiplier to be properly defined. Captures on the warp in particular are less likely to be observed and should therefore have a higher cryptic mortality. In the longline fisheries, captures during hauling will have a lower cryptic mortality than captures during setting.

Regression equations

The model is fitted to the observed number of captures and deaths. If $C'_{j,s,m}$ is the observed number of captures for fishery group j, species s and month m, then the expectation is:

$$\lambda_{j,s,m} = q_{j,s} \cdot \mathbb{O}'_{j,s,m}$$

and the likelihood is:

$$C'_{j,s,m} \sim Poisson(\lambda_{j,s,m})$$

For fishery groups with more than one capture type:

$$C'_{t,j,s,m} \sim Poisson(\pi_t \cdot \lambda_{j,s,m})$$

The probability of death is included as a separate likelihood, using the number of observed live captures $L'_{j,s,m} \leq C'_{j,s,m}$:

$$L'_{j,s,m} \sim Binomial(C'_{j,s,m}, \Psi_{j,z})$$

The probabilities of post-release survival $\omega_{j,z}$ do not form part of the likelihood equations, since there are no data which could inform their estimation. Similarly for $p^{obs}_{k,s}$, which also do not form part of the likelihood. Instead $\omega_{j,z}$ and $p^{obs}_{k,s}$ are included as distributions, which allow the total number of deaths to be predicted with appropriate uncertainty.

The catchability is a function of method (k), fishery group (j) and species group (z) covariates:

$$\log(q_{j,s}) = \mu_k + \beta_j + \beta_z + \varepsilon_{s,j}$$

where $\varepsilon_{s,j}$ is a species and fishery group specific random effect centred on zero. Coefficients β_j and β_z are constrained to sum to zero. The probability of live captures is:

$$\text{logit}(\Psi_{j,z}) = \gamma_j + \gamma_z$$

Coefficients γ_j and γ_z are constrained to sum to zero.

Estimation

All estimation was performed within a Bayesian framework using rstan (Stan Development Team, 2020).

Table 1: Summary of model terms

	Description
Subscripts	
i	Event
j	Fishing group
k	Method
t	Capture type
s	Species
z	Species group
m	Month
x	Raster grid
Estimated parameters	
N_s^{BP}	Number of breeding pairs
P_s^{B}	Annual probability of breeding
μ_k, β_j, β_z, ε_{sj}	Catchability coefficients
γ_j, γ_z	Survivorship coefficients
π_t	Probability of capture type
Derived parameters	
N_s^{adults}	Total number of adults
$\mathbb{N}_{s,m}$	Number of adults available to fishing
$q_{j,s}$	Catchabilty
$\upsilon_{j,s}$	Vulnerability
$\Psi_{j,z}$	Probability alive given capture
Input covariates	
$P_{s,m}^{\text{in eez}}$	Probability of an adult being in the EEZ
$P_{s,m}^{\text{on nest}}$	Probability of a breeding adult being on the nest
$d(s,m,x)$	Number of adults per km^2
a_i	Fishing effort intensity
Derived covariates	
$p_{i,s,m}$	Proprtionate number of individuals available to fishing
$\mathbb{O}_{i,s,m}$	Density overlap
Observational data	
$C'_{i,s,m}$	Number of captures
$L'_{i,s,m}$	Number of live captures

References

MPI (2017). Spatially Explicit Fisheries Risk Assessment (SEFRA), Chapter 3, Aquatic Environment and Biodiversity Annual Review. pages pages 20 – 57.

Stan Development Team (2020). RStan: the R interface to Stan. R package version 2.21.2.