Development of a new control rule for managing anthropogenic removals of protected, endangered or threatened species in marine ecosystems - Supplementary Information

Fanny Ouzoulias[1,*], Nicolas Bousquet[2], Mathieu Genu[3], Anita Gilles[4], Jérôme Spitz[3,5], and Matthieu Authier[3,6,*]

[1]Laboratoire de Biologie des Organismes et Ecosystèmes Aquatiques (BOREA), UMR 8067 - MNHN, CNRS, IRD, SU, UCN, UA, 75005, Paris, France
[2]Laboratoire Probabilités, Statistiques et Modélisation, UMR 8001 CNRS, Sorbonne Université, France
[3]Observatoire Pelagis, UAR 3462 CNRS - La Rochelle University, France
[4]Institute for Terrestrial and Aquatic Wildlife Research, University of Veterinary Medicine Hannover, Foundation, Büsum, Germany
[5]Centre d'Etudes Biologiques de Chizé, UMR 7372 CNRS-LRUniv, 79360 Villiers en Bois, France
[6]Corresponding author: mauthier@univ-lr.fr
[*]Equal authorship

November 23, 2023

Contents

List of Tables

List of Figures

1 Appendix - Likelihoods for the joint modeling of abundance and removals data

A sequence of observed removals $(R_0, \ldots, R_T)$ informs on θ via a likelihood function:

$$\ell(R_0, \ldots, R_T|\theta) \quad = \quad \ell(\{R_t\}|\theta) = \prod_{t=1}^{T} f(R_t|R_{t-1}, \theta) \tag{1}$$

where each conditional density function $f(R_t|R_{t-1}, \theta)$ is determined by a choice on the distribution of ε_t. Although the considered quantities are discrete and bounded in our setting, a log-normal assumption is a customary choice to model environmental stochasticity at it leads to more robust inferential results than discrete and / or bounded distributions (see references within **(author?)** 3):

$$\varepsilon_t \quad \sim \quad \log\mathcal{N}\left(-\frac{\sigma^2}{2}, \sigma\right),$$

which implies both $\mathbb{E}\left[\varepsilon_t|\mathcal{F}_{t-1}\right] = 1$ and $\mathbb{V}\left[\varepsilon_t|\mathcal{F}_{t-1}\right] = \sigma^2$. Accordingly, given Eq. 7 in main text, the conditional density of R_t in Eq. 1 becomes, for $t > 1$:

$$f(R_t|R_{t-1}, \theta) = \frac{1}{\sqrt{2\pi}\sigma R_t} \exp\left(-\frac{1}{\sigma^2} \exp\left\{\log R_t - \log g(R_{t-1}, \theta) + \frac{\sigma^2}{2}\right\}^2\right). \tag{2}$$

where $g(R_{t-1}, \theta)$ is given by Eq. 8 in main text.

One of the parameters in θ is K, the carrying capacity which needs data on absolute abundance to be estimated. Denote $\mathcal{I}$ the years for which an abundance estimate N_t^{obs}, observed with noise, is available. We denote $\mathcal{J}$ the years for which removals are observed, with $\mathcal{I} \subset \mathcal{J}$. The true abundance in year $t \in \mathcal{I}$ can be written as

$$N_t \quad = \quad \frac{R_t}{\rho}.$$

Assuming a log-normal distribution for observation errors ϵ_t' [2], $\forall t \in \mathcal{I}$:

$$N_t^{\text{obs}}|R_t, \theta, \tau_t \quad = \quad N_t \exp(\epsilon_t') \text{ with } \epsilon_t' \sim \log\mathcal{N}\left(-\frac{\tau_t^2}{2}, \tau_t\right),$$

where $\tau_t = \sqrt{\log\left(1 + \text{cv}_t^2\right)}$ and cv_t is the coefficient of variation associated with the estimated abundance N_t^{obs}. The observation model for each observed abundance datum is:

$$N_t^{\text{obs}}|R_t, \theta, \tau_t \quad \sim \quad \log\mathcal{N}\left(\log R_t - \log\rho - \frac{\tau_t^2}{2}, \tau_t\right) \tag{3}$$

The joint likelihood of the observed abundances and removals data is

$$\ell_\theta\left(\left\{N_t^{\text{obs}}\right\}_{t\in\mathcal{I}}, \{R_t\}_{t\in\mathcal{J}}\right) = \ell\left(\left\{N_t^{\text{obs}}\right\}_{t\in\mathcal{I}} | \{R_t\}_{t\in\mathcal{J}}, \theta\right) \times \prod_{t\in\mathcal{J}} \ell\left(\{R_t\}, \theta\right) \tag{4}$$

where $\ell\left(\{R_t\}, \theta\right)$ is given by Eq. 1. Under the assumption that abundances are observed independently from removals and given Eq. (3), one has:

$$\ell\left(\left\{N_t^{\text{obs}}\right\}_{t\in\mathcal{I}} | \{R_t\}_{t\in\mathcal{J}}, \theta\right) = \prod_{t\in\mathcal{I}} \frac{1}{\sqrt{2\pi}N_t^{\text{obs}}\tau_t} \exp\left(-\left(\log\left(\frac{\rho N_t^{\text{obs}}}{R_t}\right) + \frac{\tau_t^2}{2}\right)^2 \frac{1}{2\tau_t^2}\right). \tag{5}$$

The joint likelihood (Eq. 4) has been written in programming language `Stan` (4; see Appendix 2).

2 Appendix - Stan code for the Stochastic SPM

The model code in `Stan` syntax is stored as text data in a dataframe within the `RLA` package and can be accessed with:

```
library(rstan)
data(rlastan_models)
# use uniform priors
cat(rlastan_models$sspm_trend)
# compile model
rlastan <- rstan::stan_model(model_code = rlastan_models$sspm_trend,
                             model_name = "Anthropogenic Removals Threshold 2"
                             )
```

```
functions {
  real upper_bound_sigma(real phi, real r, real z) {
    // upper bound for sigma given phi, r and z
    // phi: extraction rate
    // r: growth rate
    // z: shape of the Pella-Tomlinson DD function
    real value;
    value = (1 + inv(z)) * log1p(z); // numerator
    value += -log(z) - log1m(phi + r * (z + 1) * inv(z)); // denominator

    return sqrt(expm1(value));
  }

  real removals_lpdf(vector y, int n, real phi, real r, real z,
 real sigma, real q, real D0, real B0, vector w
 ) {
    // log likelihood for removals
    // y: data
    // n: length of time series for y
    // phi: extraction rate
    // r: growth rate
    // z: shape of the Pella-Tomlinson DD function
    // sigma: environmental stochasticity
    // q: fraction
    // D0: initial depletion
    // N0: initial abundance
    // weights for likelihood
    real coef = exp(log(D0) - log(q) - log(phi) - log(N0));
    real value = 0.0;
    // currently does not include the first datum
    for(t in 2:n) {
      real mu = log(y[t-1] + ((z + 1) * r / z) * y[t-1] * (1 - pow(coef * y[t-1], z)) -
phi * y[t-1]) - 0.5 * square(sigma);
      value += w[t] * lognormal_lpdf(y[t] | mu, sigma);
    }

    return value;
  }

  real abundance_lpdf(vector y, real phi, real q, real w) {
    // log likelihood for abundance/biomass given removals
    // y: vector of length 3, with 1-removals, 2-estimated abundance/biomass, and
    //    3-associated coefficient of variation
    // phi: extraction rate
    // q: fraction
```

```
// scale parameter of log-normal distribution
    real tau = sqrt(log1p(y[3] * y[3]));
// location parameter of log-normal distribution
    real mu = log(y[1]) - log(q) - log(phi) - 0.5 * square(tau);

    return w * lognormal_lpdf(y[2] | mu, tau);
  }
}

data {
  int<lower = 2> n_year;
  int<lower = 1> n_survey;
  vector<lower = 0.0>[n_year] BYCATCH;     // removals time series
  vector<lower = 0.0>[n_year] Wr;          // weights for likelihood (removals time series)
  vector<lower = 0.0>[3] SURVEY[n_survey]; // abundance/biomass
  vector<lower = 0.0>[n_survey] Ws;        // weights for likelihood (survey time series)
  real<lower = 0.0> z;                     // shape of the Pella-Tomlinson DD function
  real<lower = 0.0> N0;                    // N0: initial abundance/biomass
  real<lower = 0.0, upper = 1.0> q;        // fraction
  real<lower = 0.0> upper_bound_r;         // upper bound for r
  real<lower = 0.0> upper_bound_phi;       // upper bound for phi
  real<lower = 0.0> lower_bound_D0;        // lower bound for D0, the initial depletion
}

transformed data {
  real nu = 1; // cauchy
  real prior_scale_slope = 0.05455187;     // see Cook et al. skeptical prior
  vector[n_survey] prop;
  vector[n_survey] survey_t;
  for(t in 1:n_survey) {
    prop[t] = log(SURVEY[t, 2]) - log(SURVEY[1, 2]);
    survey_t[t] = (t - 1) * 1.0 / (n_survey - 1);
  }
}

parameters {
  real<lower = 0.0, upper = 1.0> unscaled_r;    // growth rate
  real<lower = 0.0, upper = 1.0> unscaled_phi;  // extraction rate
  real<lower = 0.0, upper = 1.0> unscaled_sigma;// environmental stochasticity
  real<lower = lower_bound_D0, upper = 1.0> D0; // D0: initial depletion
  real aux_slope;                               // normal deviate
  real<lower = 0.0> scale_sq;                   // inverse gamma for scale mixture
  real<lower = 0.0, upper = 10> sigma_res;      // residual scale
}

transformed parameters {
  real r = unscaled_r * upper_bound_r;        // growth rate
  real phi = unscaled_phi * upper_bound_phi;  // extraction rate
  real sigma_max = upper_bound_sigma(phi, r, z);
  real sigma = unscaled_sigma * sigma_max;    // environmental stochasticity
  real slope = prior_scale_slope * aux_slope * sqrt(scale_sq); // cauchy
}

model {
  // priors
  aux_slope ~ normal(0.0, 1.0);               // normal auxiliary variate
  scale_sq ~ inv_gamma(0.5 * nu, 0.5 * nu);   // scale mixture (cauchy with nu = 1)
  unscaled_r ~ uniform(0.0, 1.0);
  unscaled_phi ~ uniform(0.0, 1.0);
  unscaled_sigma ~ uniform(0.0, 1.0);
  // weighted likelihoods
```

```
  // removals
  target += removals_lpdf(BYCATCH| n_year, phi, r, z, sigma, q, D0, N0, Wr);
  // surveys
  for(i in 1:n_survey) {
    target += abundance_lpdf(SURVEY[i]| phi, q, Ws[i]);
  }
  // trend in abundance
  target += normal_lpdf(prop| slope * survey_t, sigma_res);
}

generated quantities{
  real removal_limit = phi * fmin(1.0, exp(slope));
  real K = N0 * inv(D0);
}
```

3 Appendix - ASCOBANS ’unacceptable interactions’

The final depletion after T years of managing removals according to the fixed percentage rule of 1.7% of best available abundance estimate was recorded and averaged over the 100 simulations. Results are reported for two time horizons ($T \in 50, 100$) in Table 1. The ASCOBANS fixed percentage rule allowed to restore depleted populations to both 60% and 80% of K **on average** with unbiaised data, a result in line with **(author?)** [1]. However, this proportion dropped dramatically with biased data. Consistent use of the rule over 100 years resulted in more depleted populations compared to an horizon of 50 years suggesting also too high a removals rate.

Scenario	Time horizon T	$\Pr(D_T \geq 0.8)$	$\Pr(D_T \geq 0.6)$
Base case	50	0.52	0.72
	100	0.59	0.78
Overestimated abundance	50	0.16	0.40
	100	0.15	0.45
Underestimated removals	50	0.16	0.40
	100	0.15	0.45

Table 1: **Success rate of ASCOBANS fixed percentage rule**. Proportions are computed over 100 simulations, and used to approximate probabilities to reach a given depletion level after T years of management.

4 Appendix - Performance metrics

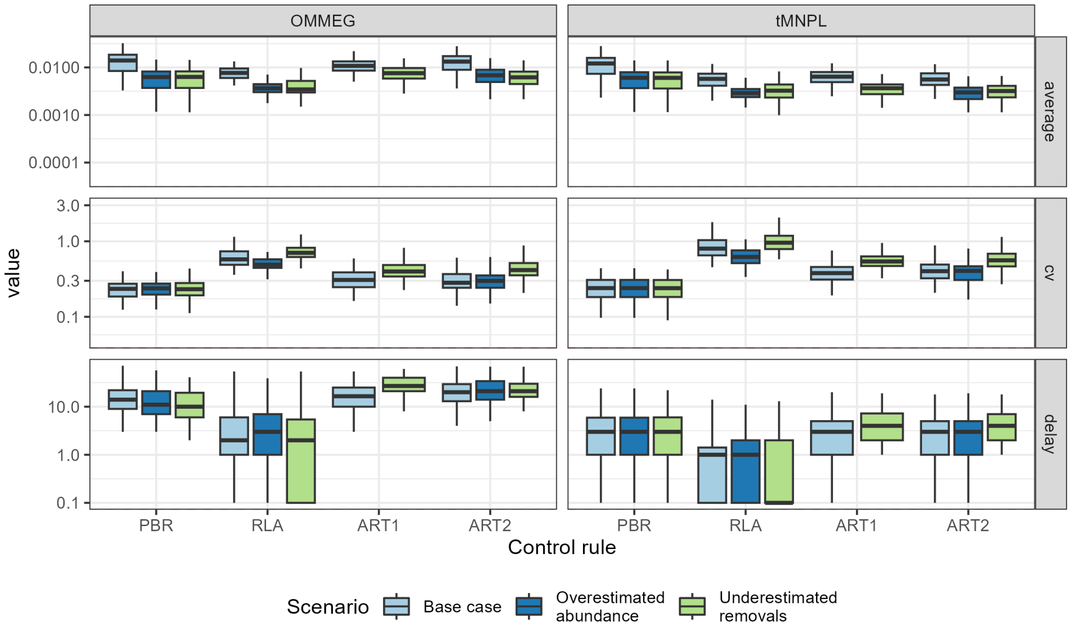

Figure 1: **Performance metrics**: Boxplots summarizing the distribution of variability (`cv`, unitless), mean (`average`, as a proportion of current abundance estimate; *e.g.* $0.01 = 1\%$), and `delay` in recovery (in years). The y-axis is on a logarithmic scale.

5 Appendix - Example of simulations

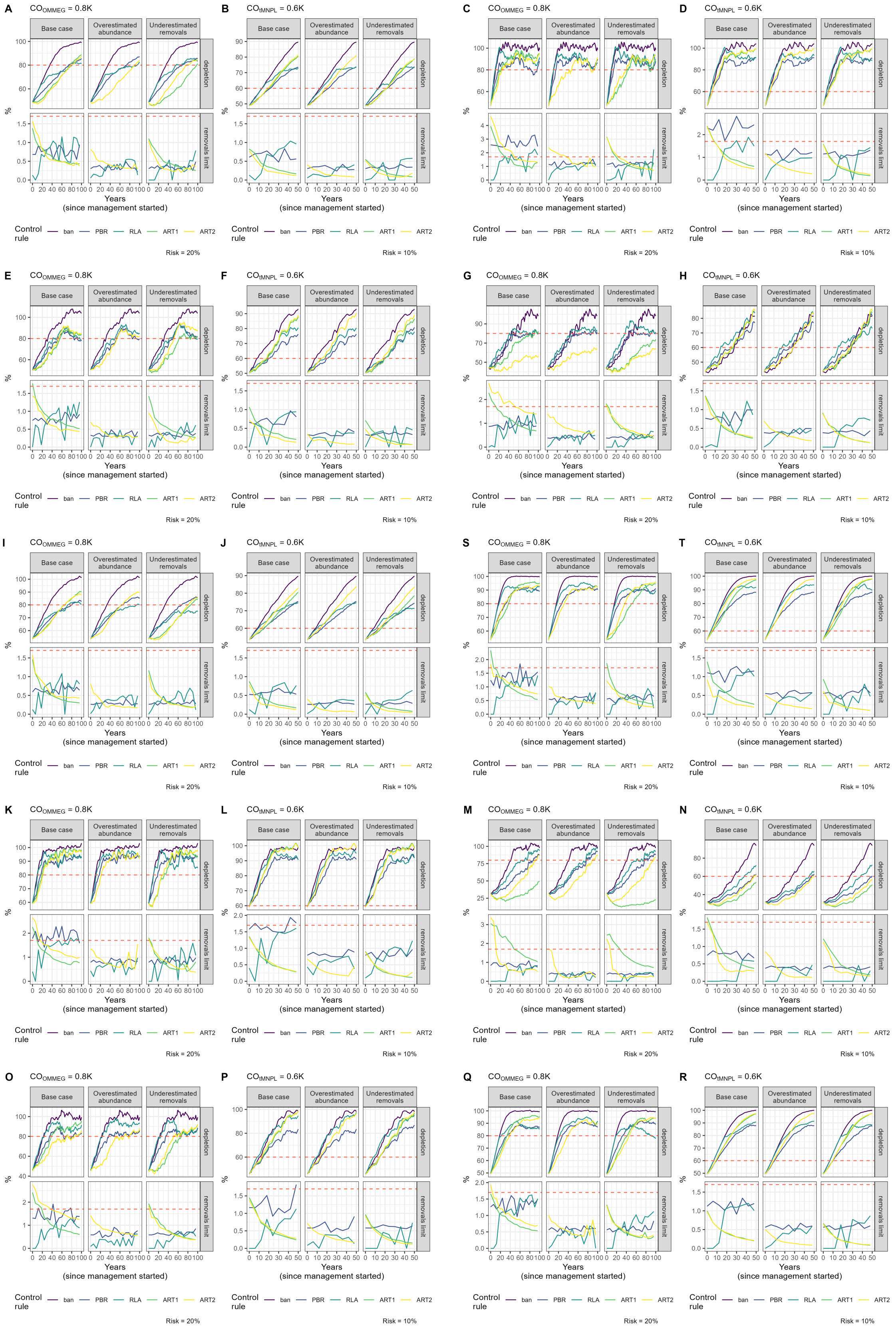

Table 2: **Example of simulations**: 10 simulations within MSE on assessing PETS long-term viability.

References

[1] ASCOBANS. Annex O - Report of the IWC-ASCOBANS Working Group on Harbour Porpoises. *Journal of Cetacean Research and Management*, 2(Supplement):297–305, 2000.

[2] J. Best and A. Punt. Parameterizations for Bayesian State-Space Surplus Production Models. *Fisheries Research*, 222:105411, 2020.

[3] N. Bousquet, T. Duchesne, and L. P. Rivest. Redefining the Maximum Sustainable Yield for the Schaefer Population Model Including Multiplicative Environmental Noise. *Journal of Theoretical Biology*, 254 1:65–75, 2008.

[4] B. Carpenter, A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, and A. Riddell. `Stan`: A Probabilistic Programming Language. *Journal of Statistical Software*, 76(1), 2017.