

5th Meeting of the Scientific Committee

Shanghai, China, 23 - 28 September 2017

SC5-DW12

A simple delay-difference model for assessment of data-poor

Orange roughy stocks

Charles Edwards & Marie-Julie Roux (NIWA)



FORM 5

Final Research Report

Title of report: A simple delay-difference model for assessment of data-poor orange roughy stocks

Authors: Edwards, C. T. T. and Roux, M.-J.

Date: August 2017

Research Provider: National Institute of Water and Atmospheric Research

Project Code: SEA201622

Project Title: Analytical support for development of a new SPRFMO bottom fishing measure

Principal Investigator: Rowden, A.

Project Start Date: 1st February 2017

Project End Date: 1st August 2017

EXECUTIVE SUMMARY

Simple Bayesian biomass dynamics models can provide reliable estimates of stock depletion and maximum sustainable yield to inform management of data-limited orange roughy fisheries on the high seas. In orange roughy, recruitment to the spawning stock is lagged 30-40 years relative to current biomass. In this paper, a simple delay difference model that separates recruitment from other dynamics is developed and applied to assess whether the inclusion of a recruitment lag presently affects status evaluation for orange roughy stocks within the SPRFMO Convention Area. The results indicated consistency in stock status estimates between a standard surplus production model and a delay-difference model implemented with and without a 38-year lag on recruitment. This confirms the consistency of a cohort aggregated modelling approach and the reliability of biomass dynamic models for stock assessments of orange roughy fisheries within SPRFMO, where reliable and representative age data and fisheries-independent indices of abundance are not yet available.

OBJECTIVES

This is the Final Research Report for project SEA201622, and constitutes partial fulfilment of the following project deliverables as detailed in the proposal:

Overall Objective

1. To provide scientific services to support MPIs contribution to developing and implementing management measures for the South Pacific Regional Fisheries Management Organisation (SPRFMO)

Specific Objective

1. To finalise spatially disaggregated CPUE models and stock assessment models for all orange roughy management units in the western SPRFMO Area with sufficient data.

1. Introduction

The assessment of orange roughy *Hoplostethus atlanticus* stocks on the high seas is complicated by data availability and stock structure uncertainties. Within the South Pacific Regional Fisheries Management Organisation (SPRFMO) Convention Area, there are currently no reliable age data and fisheries-independent (i.e., trawl and/or acoustic surveys) indices of stock abundance for orange roughy. Commercial catch per unit effort (CPUE) data are the principal information source available to evaluate stock status. A data-limited approach was developed and implemented to evaluate stock status and fisheries sustainability (Roux et al. 2017). This approach combines the estimation of spatially-disaggregated CPUE indices of stock abundance and the fitting of a cohort-aggregated Bayesian biomass dynamics model (BDM) to estimate biomass trajectory and evaluate stock status. Until fisheries-independent abundance and age data are available, it has the potential to provide reliable estimates of stock depletion and maximum sustainable yield (MSY) reference points to inform management decisions for SPRFMO orange roughy stocks.

The applicability and validity of the BDM model to orange roughy was demonstrated by McAllister and Edwards (2016). In that study, a similar cohort-aggregated state-space biomass dynamics model (BSP) was fitted to an aggregate of stock trend data (including both fishery-dependent and fishery-independent indices) for a New Zealand domestic stock (East Chatham Rise orange roughy). The BSP model provided good fits, informative estimates of stock size, and estimates of depletion (relative to MSY or the carrying capacity) similar to those estimated by an age-structured production model such as CASAL.

Stock productivity in BDM is expressed as a single parameter r (the maximum intrinsic rate of population increase as biomass approaches zero), which is offered as an informative prior constructed using life history data (i.e., maturity, growth, recruitment and natural mortality information). In long-lived species such as orange roughy, recruitment to the spawning stock is lagged 30-40 years relative to the current biomass (Cordue 2014; Dunn 2007). Because of this, the absence of a recruitment lag in the BDM model structure may lead to bias in fisheries that have been occurring for longer than the age of recruitment. In this paper, we developed and applied a delay-difference model implemented in BDM to 1) test the validity of the equilibrium-level recruitment assumption; and 2) assess whether the inclusion of a recruitment lag presently affects status estimation for orange roughy fisheries within SPRFMO. This is done using a case study example focusing on the Central Louisville Ridge orange roughy stock.

2. Model description

Here we present a model of the form:

$$B_{t+1} = g(B_t) + r(B_t)$$

where $g()$ is a growth function accounting for somatic growth and mortality, and $r()$ is a recruitment function. Separating recruitment from other dynamic processes distinguishes this model class from other biomass dynamic models (such as the logistic surplus production model) and gives a more intuitive representation of this important determinant of the populations response to harvesting.

The growth function is most easily represented as a simple linear predictor:

$$g(B) = \psi \cdot B \cdot (1 - H)$$

where H is the harvest rate. The recruitment function can take a number of forms, including a constant recruitment level. For this study we assumed a Bertone-Holt function, which has more realistic reference point values:

$$r(B) = \frac{\alpha \cdot B}{1 + \beta \cdot B}$$

Since both $g(B)$ and $r(B)$ are functions of B , we assuming that B represents both the recruited and mature sections of the population biomass (i.e. that the maturation and selectivity ogives are the same), which is a common assumption for cohort aggregated models of this type. The model is best formulated as a proportion of K , with notation $x_t \equiv B_t/K$, which avoids high correlations between states and the carrying capacity, thereby increasing the rate of convergence of MCMC sampling:

$$x_{t+1} = \psi \cdot x_t \cdot (1 - H_t) + \frac{\alpha \cdot x_t}{1 + \beta \cdot x_t} \quad (1)$$

where $H = C/(x \cdot K)$ is the catch relative to available biomass and with initial condition x_0 . The recruitment parameters are calculated using steepness h :

$$\alpha = \frac{4h}{\rho(1-h)} \quad (2a)$$

$$\beta = \frac{5h-1}{1-h} \quad (2b)$$

where ρ is a scalar such that $r(x=1) \cdot \rho = 1$; i.e. the recruitment at unexploited (equilibrium) carrying capacity multiplied by ρ equals one. To ensure an equilibrium at K it is necessary for $r(x=1) = 1 - \psi$ and therefore $\rho = (1 - \psi)^{-1}$.

An advantage of this class of model is that it can easily incorporate a recruitment lag by simply re-writing the dynamic equation as:

$$x_{t+1} = \psi \cdot x_t \cdot (1 - H_t) + \frac{\alpha \cdot x_{t-a+1}}{1 + \beta \cdot x_{t-a+1}} \quad (3)$$

where $a \geq 1$ is the age at recruitment.

2.1. Reference points

Here we derive the depletion relative to K at maximum sustainable yield (x^{MSY}). To find the equilibrium harvest rate we set $x_{t+1} = x_t$ in equation 1 and solve for H to obtain:

$$H_0(x) = 1 - \frac{1}{\psi} + \frac{r(x)}{\psi \cdot x} \quad (4)$$

Then to find x^{MSY} we solve the equation

$$\frac{d}{dx} (H_0 \cdot x) = 0$$

assuming the second derivative is less than one at this point (i.e. it is a maximum). For the model described here:

$$\frac{d}{dx} (H_0 \cdot x) = 1 - \frac{1}{\psi} + \frac{\alpha}{\psi \cdot (1 + \beta \cdot x)^2}$$

which implies that:

$$x^{MSY} = \frac{1}{\beta} \cdot \left(\sqrt{\frac{\alpha}{1-\psi}} - 1 \right) \quad (5)$$

We note that since this is an equilibrium condition, it is valid even if there is assumed to be a recruitment lag (since we are assuming that $x_t = x_{t-a}$). The harvest rate at MSY is obtained from substituting equation 5 into equation 4, and MSY is then:

$$MSY = H_0(x^{MSY}) \cdot x^{MSY} \cdot K \quad (6)$$

$$= K \cdot \left(\left(1 - \frac{1}{\psi} \right) \cdot x^{MSY} + \frac{r(x^{MSY})}{\psi} \right) \quad (7)$$

The productivity function, along with x^{MSY} is shown in figure 1.

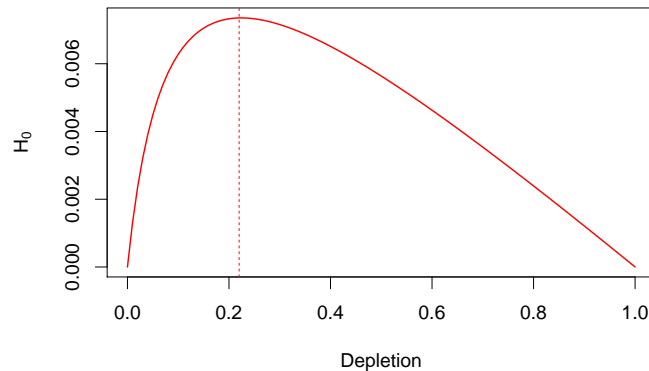


Figure 1: Sustainable harvest rate as a function of depletion x , with maximum shown at $x^{MSY} = 0.22$.

2.2. Prior specifications and estimation framework

Estimation was carried out within a Bayesian state-space framework using MCMC, for the parameter set $\theta = \{\psi, K, x_0, q\}$, where q is the catchability for the single available abundance index I . The model can be written as:

$$\begin{aligned}\bar{x}_{t+1} &= g(x_t) + r(x_t) \\ x_t &\sim LN(\ln(\bar{x}_t) - \sigma_{[p]}^2/2, \sigma_{[p]}) \\ I_t &\sim LN(\ln(q \cdot x_t) - \sigma_{[o]}^2/2, \sigma_{[o]})\end{aligned}$$

where \bar{x} is the deterministic expectation of the depletion, and $LN(\cdot)$ refers to the log-normal distribution. These distributions are formulated so as to have expectations: $E[x] = \bar{x}$, and $E[I] = q \cdot x$. The initial depletion x_0 also follows a log-normal distribution with variance $\sigma_{[p]}^2$ and expectation $E[x_0] = 1$. Following the approach taken in recent applications of this type of model, no attempt was made to estimate the observation or process error variances. Instead, the standard errors were fixed on input at $\sigma_{[p]} = 0.05$ and $\sigma_{[o]} = 0.2$. The catchability scalar was calculated as a “nuisance” parameter. These assumptions follow those made by Roux et al. (2017).

Assumed life-history parameters are listed in table 1. Catch and abundance data are listed in table 2 and plotted in figure 2.

It was necessary to specify priors for ψ and K . A uniform prior (with wide bounds) was assumed for $\ln(K)$, which is proportional to $1/K$ and therefore gives lower weight to higher K values.

$$\ln(K) \sim U(.,.)$$

with the upper and lower bounds for $\ln(K)$ chosen subjectively so as to not impinge of the parameter space explored during estimation.

A prior on ψ can be derived from life-history data in the following manner. Assuming recruitment and maturity occur according to the same ogive, we first calculate the average mass increment of recruited fish:

$$w^+ = \frac{1}{n} \sum_a \frac{w_a}{w_{a-1}} p_a m_a$$

where w , p and m are the mass, number per recruit and maturity at age respectively. The expected ψ represents combined effects of somatic growth and natural mortality:

$$E[\psi] = \exp(-M)(1 + w^+)$$

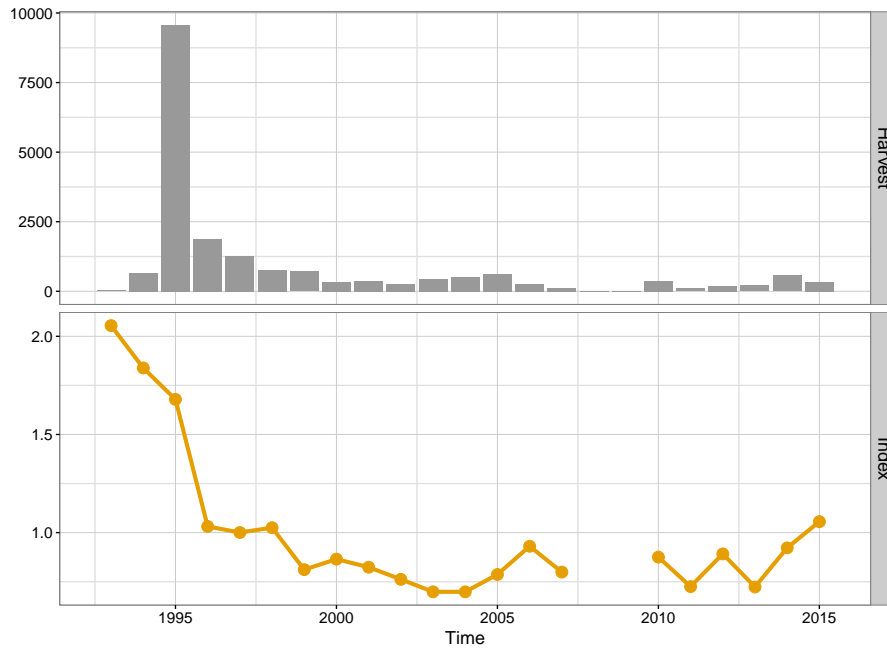


Figure 2: Input catch and abundance data.

which also gives an implicit prior for ρ from our assumption of equilibrium at K . Since recruitment at equilibrium is positive, we can also infer that $\psi < 1$. From the available life-history data (table 1), we calculated $E[\psi] = 0.986$. It is difficult to specify an informative prior for ψ that will allow mixing so close to the upper bound. We therefore used the logit transform, with an informative normal prior:

$$\text{logit}(\psi) \sim N(\mu, 1)$$

Although the probability density function for ψ can be defined, there is no closed form analytical solution for the expected value in terms of μ (Johnson 1949), and we therefore cannot back-calculate a value for μ from $E[\psi]$. However it is possible to approximate the inverse-logit function using a normal cumulative density $\Phi(\lambda \cdot \psi)$ with $\lambda = \pi/5.35$. To find μ , we can then solve:

$$E[\psi] \approx \int \Phi(\lambda \cdot z) \cdot N(z; \mu, 1) dz = \Phi\left(\frac{\mu}{\sqrt{\lambda^{-2} + 1}}\right)$$

which gives $\mu = 4.318$. Using this paramaterisation, the prior for ψ is shown in figure 3. Values for ρ and the α recruitment parameter were sampled from the posterior assuming a fixed steepness (table 1):

$$\rho_i = \frac{1}{1 - \psi_i}$$

$$\alpha_i = \frac{4h}{\rho_i(1 - h)}$$

where i represents an independent posterior sample from the MCMC chain.

3. Model fits and outputs

The model was written using the Stan modelling language (<https://github.com/stan-dev/rstan>), and executed within the **bdm** R-package (<https://github.com/cttedwards/bdm>). Three model runs were performed using the data listed in tables 1 and 2:

1. Delay-difference model (**Base model**; equation 1)

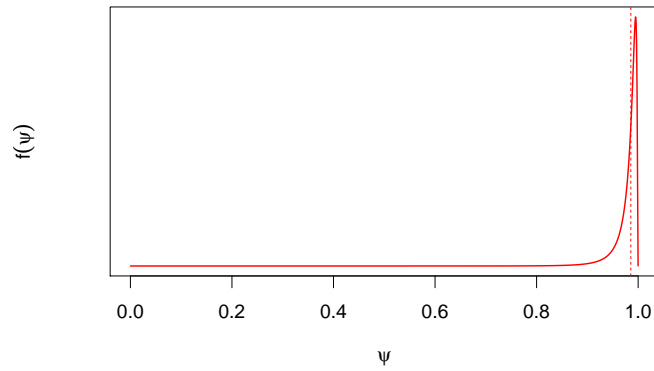


Figure 3: Prior distribution of ψ obtained from life-history data with $E[\psi] = 0.986$.

Table 1: Life-history data for orange roughy

Parameter	Notation	Value
Asymptotic age (y)	a_{∞}	130
Natural Mortality	M	0.045
Age at maturity (y)	a_{mat}	38
Growth (von Bertalanffy)	l_{∞}	37.63
(cm)	k	0.065
	t_0	-0.5
Length-weight	a	0.0921
	b	2.71
Recruitment steepness	h	0.75

Table 2: Catch and abundance data from Louisville Central for model fitting

Year	Catch (tonnes)	Abundance Index
1993	25.000	1.000
1994	657.000	0.895
1995	9566.000	0.817
1996	1889.000	0.502
1997	1277.000	0.487
1998	760.000	0.499
1999	712.500	0.395
2000	331.663	0.421
2001	371.472	0.401
2002	251.419	0.371
2003	443.464	0.340
2004	508.693	0.340
2005	629.621	0.383
2006	271.892	0.453
2007	117.008	0.389
2008	0	–
2009	0	–
2010	371.000	0.426
2011	101.000	0.353
2012	185.000	0.434
2013	215.000	0.352
2014	571.000	0.449
2015	341.000	0.514

2. Delay-difference model with a lag of $a = a_{mat}$ (equation 3)

3. Surplus production model, similar in approach to that applied by Roux et al. (2017)

All runs assumed the same depletion at MSY (i.e. $x^{MSY} = 0.224$).

For each model fit, four MCMC chains were run, each of 3000 iterations, with the first half of each chain discarded. Convergence diagnostic plots for the base model are shown in figure 4, with the fit in figure 5. Both indicate good convergence. The resultant estimated depletion dynamics are shown in figure 6a. Comparative plots of the depletion dynamics are also shown for the alternate model runs in figures 6b and 6c.

4. Discussion and Conclusions

The results demonstrate consistency in stock status estimates between a biomass dynamics model (BDM) used in SPRFMO orange roughy stock assessments (Roux et al. 2017) and a delay-difference model implemented with and without a 38-year lag on recruitment.

The delay-difference model structure allowed to distinguish recruitment from other dynamics processes to provide a more realistic representation of the likely response of orange roughy to harvesting. The absence of recruitment lag effects was expected in this case, because the fishery for orange roughy on the central Louisville Ridge only started in the mid-1990s. Since the ages at recruitment and maturity for orange roughy are greater than 30 years, the fishery has yet to impact the new recruits (i.e., all recruits to the fishery so far have arisen from spawning biomasses yet unimpacted by harvesting). The same the same conditions apply to all SPRFMO orange roughy fisheries. Only exception within SPRFMO is the Northwest Challenger fishery, which is nearing 35 years. A

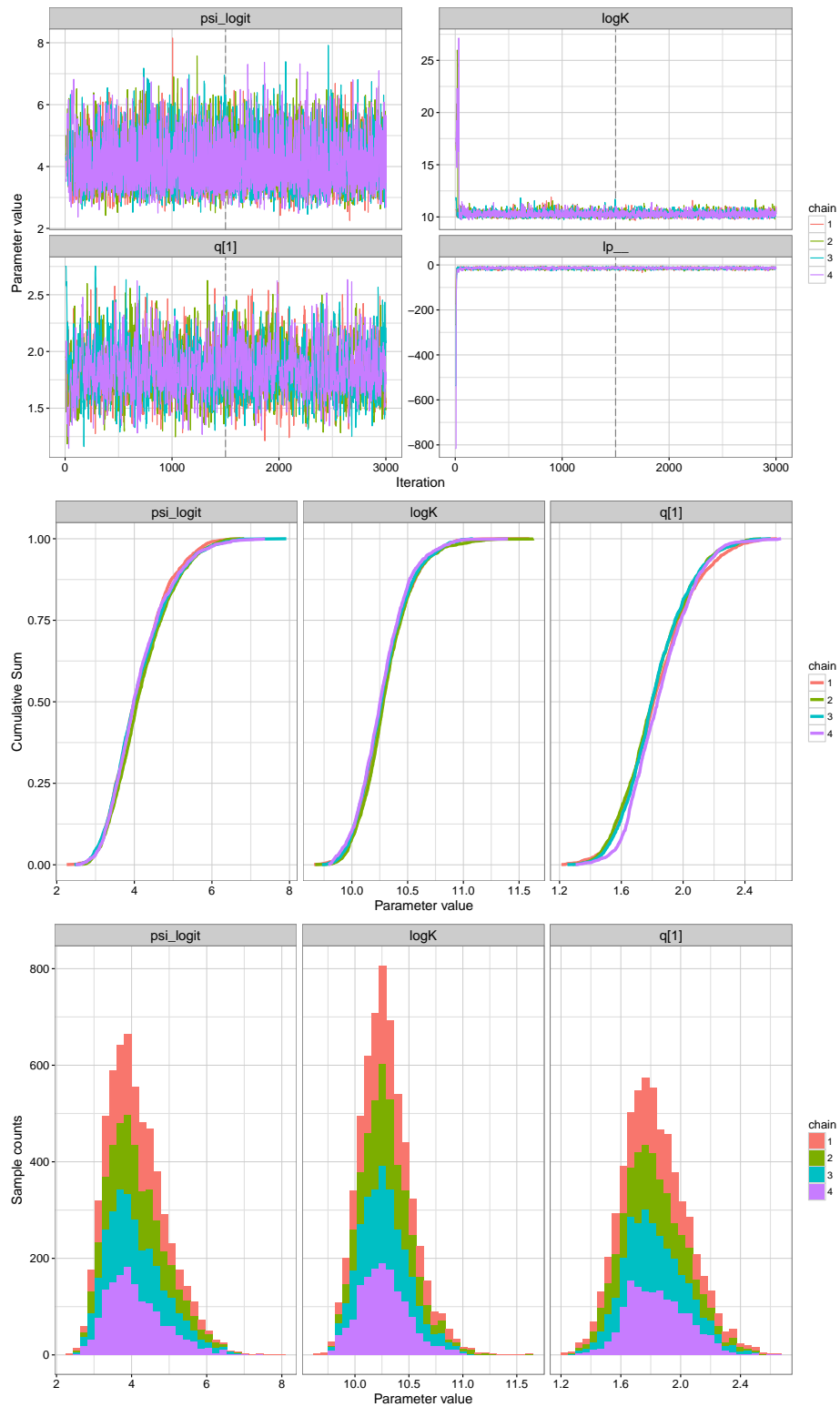


Figure 4: Trace, cumulative sum and (stacked) histogram plots of estimated parameters for fit of the base model to orange roughy data.

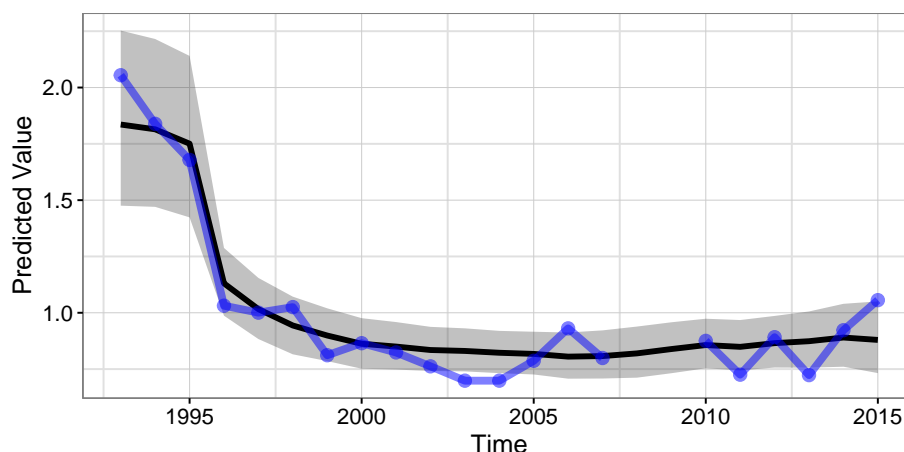


Figure 5: Base model fit to catch rate index. The predicted index is shown as a median value with 95% credibility intervals.

precautionary approach for future projections for Northwest Challenger orange roughy therefore, would be to implement a delay-difference model with recruitment lag alongside BDM, when exploring alternative harvest rate options.

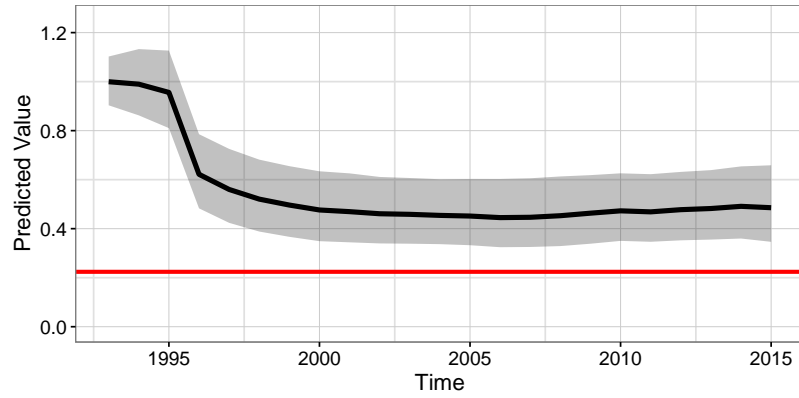
This paper confirms that the assumption of equilibrium recruitment is presently valid for orange rough stocks within the SPRFMO Convention Area. Thus, until reliable and representative age data and fisheries-independent abundance indices are available for these fisheries, stock assessments performed using simple biomass dynamics models can provide reliable estimates of stock status and reference points on which to base science advice for management.

There are several advantages in implementing simple Bayesian state-space production models for stock assessments of data-limited orange roughy on the high seas. These include limited complexity, ease of application, flexibility and high predictive performance. Limited complexity ensures that the results can be easily interpreted and communicated for enhanced transparency. Ease of application and low data requirements make this class of models practical in that the assessments can be easily updated. The models implemented herein have demonstrated flexibility to fit and predict annual changes in the abundance data (McAllister and Edwards 2016), allowing the testing of simple and straightforward assumptions with high predictive performance.

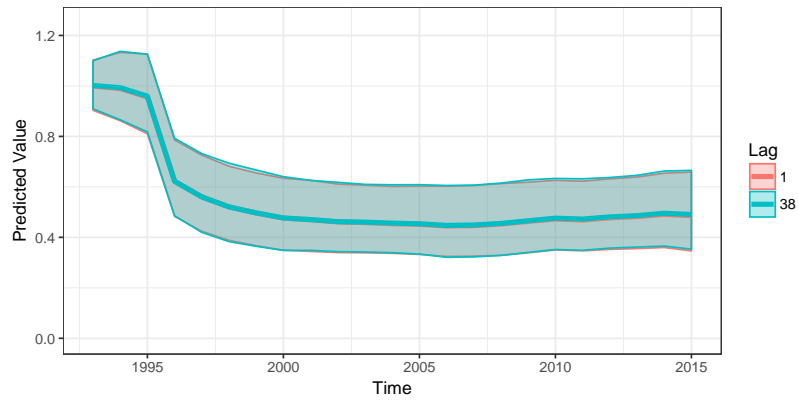
5. Acknowledgements

This work was funded by the Ministry for Primary Industries, New Zealand, under project code SEA201622. Dr. Ian Doonan kindly reviewed the manuscript.

(a) Base model run showing depletion relative to x^{MSY} .



(b) Comparison of base model (Lag = 1) with the model with lagged recruitment (Lag = 38).



(c) Comparison of base model (DD) with the surplus production model applied by Roux et al. (2017).

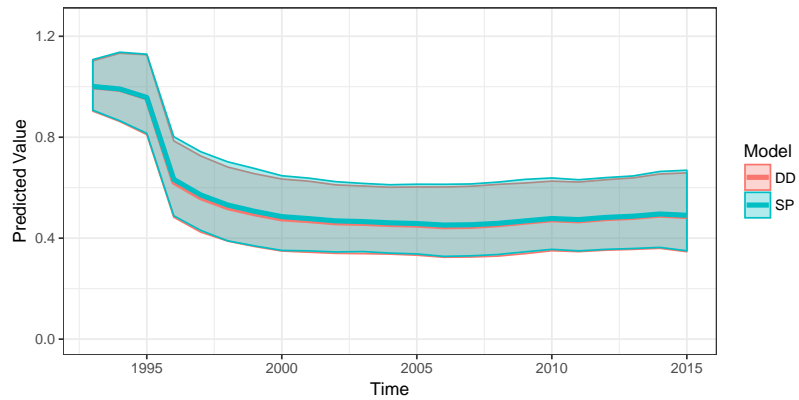


Figure 6: Estimated depletion from alternate model runs. The mean and 95% credibility intervals are shown.

6. References

- Cordue, P. A., 2014. The 2014 orange roughy stock assessments. New Zealand Fisheries Research Document 2014 (50).
- Dunn, M. R., 2007. CPUE analysis and assessment of the East Chatham Rise orange roughy stock (part of ORH 3B) to the end of the 200405 fishing year. New Zealand Fisheries Research Document 2007 (08).
- Johnson, N., 1949. Systems of frequency curves generated by methods of translation. *Biometrika* 36, 149–176.
- McAllister, M. K., Edwards, C. T. T., 2016. Applications of a Bayesian surplus production model to New Zealand fish stocks. New Zealand Fisheries Research Document 2016 (52).
- Roux, M.-J., Doonan, I., Edwards, C. T. T., Clark, M. R., 2017. Low information stock assessment of orange roughy *Hoplostethus atlanticus* in the South Pacific Fisheries Management Organisation convention area. New Zealand Fisheries Research Document 2017 (01).

Appendix: Stan model code

```

data {
  int T;
  int I;
  int lag;
  real index[T,I];
  real harvest[T];
  real h;
  real sigmao[T,I];
  real sigmap;
}
parameters {
  real<lower=3,upper=30> logK;
  real psi_logit;
  real<lower=0> x[T];
}
transformed parameters {

  real q[I];

  // variance terms
  real sigmao2[T,I];
  real sigmap2;

  real rho;
  real alpha;
  real beta;
  real psi;

  psi = inv_logit(psi_logit);
  rho = 1/(1 - psi);
  alpha = 4*h / (rho * (1 - h));
  beta = (5*h - 1) / (1 - h);

  for(t in 1:T)
  for(i in 1:I)
  sigmao2[t,i] = square(sigmao[t,i]);
  sigmap2 = square(sigmap);

  // compute mpd catchability assuming
  // variable sigmao over time assuming
  // uniform prior on q
  {
    real sum1;
    real sum2;
    real p;
    for(i in 1:I){
      sum1 = 0.0;
      sum2 = 0.0;
      p = 0.0;
      for(t in 1:T){
        if(index[t,i]>0.0 && x[t]>0.0) {
          sum1 = sum1 + log(index[t,i]/x[t])/sigmao2[t,i];
          sum2 = sum2 + 1/sigmao2[t,i];
          p = p + 1.0;
        }
      }
      if(p>2.0) { q[i] = exp((0.5 * p + sum1) / sum2);
      } else q[i] = 0.0;
    }
  }
}
}
model {

  // prior densities for
  // estimated parameters

```

```

// *****
logK ~ uniform(3.0, 30.0);
psi_logit ~ normal(4.318,1);
#psi ~ beta(1.123611, 0.01226281);

// state equation
// *****
{
  real H;
  real r;
  real mu;
  real x_lag;

  x[1] ~ lognormal(log(1.0) - sigmap2/2, sigmap);

  for(t in 2:T) {

    if(t <= lag) { x_lag = x[1];
    } else x_lag = x[t-lag];

    H = fmin(exp(log(harvest[t-1]) - logK - log(x[t-1])), 1.0);
    r = alpha * x_lag / (1 + beta * x_lag);

    mu = psi * x[t-1] * (1 - H) + r;

    if(mu > 0.0) mu = log(mu) - sigmap2/2;
    else mu = log(0.01) - sigmap2/2;

    x[t] ~ lognormal(mu, sigmap);
  }

  // observation equation
  // *****
  {
    real mu;
    for(i in 1:I){
      for(t in 1:T){
        if(index[t,i]>0.0 && x[t]>0.0 && q[i]>0.0) {
          mu = log(q[i]*x[t]) - sigmao2[t,i]/2;
          index[t,i] ~ lognormal(mu, sigmao[t,i]);
        }
      }
    }
  }

  // apply penalty for H>0.95
  // *****
  for(t in 1:T){
    real H_; H_ = harvest[t]/exp(log(x[t]) + logK);
    if(H_>0.95) {
      target += -log(H_/0.95) * (1/sigmap2);
    }
  }

  generated quantities {

    real biomass[T];
    real depletion[T];
    real harvest_rate[T];

    real current_biomass;
    real current_depletion;
    real current_harvest_rate;

    real msy;
    real depletion_at_msy;
    real biomass_at_msy;

```

```
real harvest_rate_at_msy;

real observed_index[T,I];
real predicted_index[T,I];

for(t in 1:T) {
biomass[t] = x[t] * exp(logK);
depletion[t] = x[t];
harvest_rate[t] = harvest[t]/exp(log(x[t]) + logK);
}

current_biomass = biomass[T];
current_depletion = x[T];
current_harvest_rate = harvest_rate[T];

depletion_at_msy = (1 / beta) * (sqrt(alpha / (1 - psi)) - 1);
biomass_at_msy = depletion_at_msy * exp(logK);
harvest_rate_at_msy = 1 - 1 / psi + (alpha * depletion_at_msy / (1 + beta * depletion_at_msy)) / (psi * depletion_at_msy);
msy = harvest_rate_at_msy * biomass_at_msy;

for(i in 1:I){
for(t in 1:T){
observed_index[t,i] = index[t,i];
predicted_index[t,i] = q[i]*x[t];
}
}
}
```