Rejection sampling for complex fishery models facing limited data

Ernesto Carrella, Steven Saul, Richard M Bailey, Nicolas Payette, Katyana A. Vert-pre, Aarthi Ananthanarayanan, Michael Drexler, Chris Dorsett, Jens Koed Madsen

Abstract

It is difficult to estimate fishery models when data are limited or of poor quality. When information is a mix of quantitative and qualitative, biological and social, rejection sampling provides a model-agnostic process to incorporate data within simulation models.

Even so, information will often be so limited that parameter identification will remain hopeless.

Under these circumstances, rejection sampling should be used to build an ensemble of many "close enough" candidate fisheries that reproduce the few real-world observations we have. We can use this ensemble for policy making analysis even when all the "true" fishery parameters remain unknown.

We showcase this process with two models: a bio-economic delay difference equation model and an agent-based fishery model.

1 Introduction

Data poor fisheries often need assessment and management the most (Prince and Hordyk 2018). It is difficult to fit models when data are limited and of poor quality. For any given model there may be many alternative parameterizations that reproduce the few real world observations we have available. With limited evidence the problem is not how to replicate the data, rather it is to account for all the many ways in which replication is possible. We propose giving up on either looking for the "most likely" parameters or even the confidence intervals around them. Rather we propose producing a large ensemble of fisheries (i.e. combinations of parameters and histories) "close enough" to the data and then simulate policies on each fishery to draw conclusions that are robust to misspecification.

We can build this fishery ensemble via rejection sampling (see chapter 14 in Russell and Norvig 2009; see also "rejection filtering" in Hartig et al. 2011). At its core this simply means running the model many times with random parameters and rejecting all runs that do not match the data we observed. If we are willing to state priors on the parameter space, the resulting ensemble is the Bayesian posterior of all the "close enough" fisheries (this is the rejection approximate Bayesian computation of Beaumont 2009). If we are unwilling to state priors, the ensemble can still be used to look for robust policies (those that work unanimously for every member of the ensemble) or can be split into groups to gain a better understanding of which policies work when (scenario based uncertainty, as in Meier et al. 2016). This can be useful for strategic guidance under deep-uncertainty where multiple states of nature are possible.

While rejection sampling is computational expensive, it requires no likelihood function which allows us to match models with data that would otherwise be hard to use, in particular qualitative or socio-economic observations. It is also a useful technique when fitting simulation models whose likelihood exists but is intractable or unknown. More generally, rejection sampling substitutes deriving a likelihood function with running simulations, so that it can be used to fit any kind of model in a uniform way.

Rejection sampling (usually under the name of rejection filtering) is popular in ecology as an operationalization of pattern oriented modeling (Grimm et al. 2005). Rejection sampling is also at the basis of the "ensemble ecosystem" approach (Baker, Gordon, and Bode 2017) that generates a large set of qualitative Lotka-Volterra models and rejects all those that do not match a set of known responses to shocks. Rejection sampling is also present in fisheries as it powers the catch-MSY method (Martell and Froese 2013). We propose here to move past the catch-MSY approach of using rejection sampling to build confidence intervals around parameters (logistic r and K). We instead propose collecting the entire state history and parameter combinations so that we can simulate policy shocks across all accepted fisheries to look for patterns and vulnerabilities. Under-identification (the inability to retrieve parameters given data at hand) is a problem even in data-rich fishery models (Yin et al 2018); dealing with unknown parameters by studying an "irreducible" ensemble is an approach championed by the DLM toolkit (Carruthers et al. 2014; Hordyk and Carruthers 2018).

We explain the method in section 2. In section 3 we rank and optimize policies in a simple difference equation bio-economic model. In section 4 we apply the method to a POSEIDON agent-based model (Bailey et al. 2018) implementation of the Bungo Channel's hairtail fishery to show, inter alia, how a low SPR (spawning potential ratio, Hordyk et al. 2019) can be a predictor of imminent long-term economic improvement when contextualized by the model.

2 Methods

2.1 Rejection sampling

We start with a set of evidence measured in the real world $E = e_1, ..., e_k$ and a simulation model m that takes as input parameters $\Theta = \theta_1, ..., \theta_n$ to generate the synthetic evidence $\hat{E} = \hat{e}_1(\Theta), ..., \hat{e}_n(\Theta)$. We define a set of filters $f_1(E, \hat{E}), ..., f_k(E, \hat{E})$ as predicate functions returning true if, for some aspect of the evidence set, the synthetic evidence is "close enough" to the real one. Rejection sampling proceeds as follows:

1. Draw a set of random model parameters $\hat{\theta}_1, ..., \hat{\theta}_n$ from their prior distribution $\pi(\Theta)$
2. Run the model and generate set of synthetic evidence $\hat{e}_1(\hat{\Theta}), ..., \hat{e}_n(\hat{\Theta})$.

3. If at least one filter returns false (that is some synthetic evidence is not “close enough” to its real counterpart), reject (filter away) this simulation

We run through this loop until our computational budget is exhausted[1] and collect all the simulations that were not rejected. Each accepted simulation corresponds to a valid fishery.

We use the ensemble of all accepted simulations to perform two kinds of tasks. First, we can study the posterior distribution of the parameters or other state variables. Studying the evolution of the posterior as we add evidence (rejecting more runs) allows us to understand how each evidence shapes the posterior. Second, we can apply the same policies to each fishery in the ensemble to compare their effect and predict their overall success.

The definition of filters is deliberately vague as it depends on the kind of data available. If some evidence is statistical then the filters can incorporate it; for example if we have estimated SPR (spawning potential ratio) the filter could stipulate that the simulated SPR must be within the 95% confidence intervals of the real estimate (in effect turning this into an indirect inference problem, Gourieroux et al. 1993). If however some evidence is qualitative the filter will be more open; for example, if the expert opinion is that landings have recently decreased we should only filter out simulations where landings have increased without specifying any further restriction.

The advantage of defining filters this way is that it is relatively simple to test their sensitivity. We can run rejection sampling using wide filters that accept many fisheries and then progressively tighten the filters to study changes of the progressively smaller accepted set. Since filters are not weighted, they must unanimously pass, we can proceed by changing one filter at a time. As long as we store all runs before filtering, sensitivity analysis requires no new runs or parametrization.

As an example observe figure 2.1. In this fictitious example we are interested in x (a parameter or state variable in the model) and how it changes as evidence is used. In this example there are two pieces of evidence and the prior on x is normal. The posterior, that is the set of x for the runs that pass both evidence filters, is bimodal. However had we only used evidence one, the posterior would have been still normal. Using only evidence two however gives the posterior its bimodality. Here then it is evidence two that is important for the shape of the posterior of x.

[1] This follows the “reference table” approach in the Approximate Bayesian Computation literature (Pudlo et al, 2015); one can always choose the number of runs based on the quality of the accepted ensemble via, for example, variance stability as in Lee et al (2015)

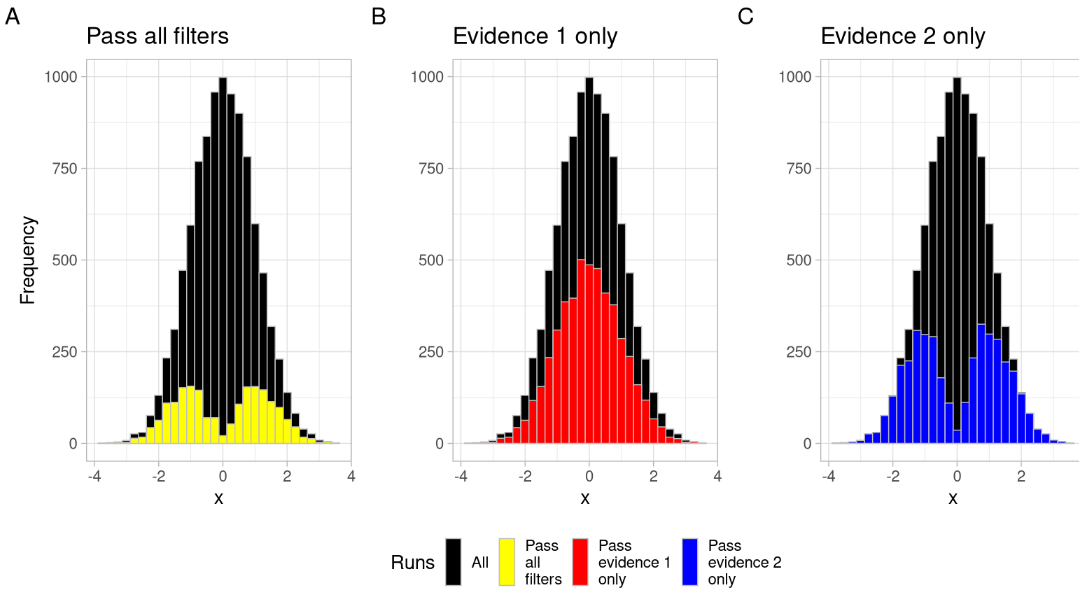

Figure 2.1: A simple example to understand how to look at the effect of evidence in rejection sampling. Panel A,B, and C show the distribution of x for all simulations in black. A shows the distribution for the simulations that passed both filters in yellow. B shows the distribution of x for runs that are accepted when we look only at evidence 1 in red. C shows the same using only evidence 2 in blue. It is clear that the shape of the posterior is due to evidence 2. Evidence 1 rejects about half of all the initial simulations in the prior but does not alter the overall posterior distribution of x the way evidence 2 does

The main advantage of rejection sampling compared to standard model estimation is that it is model agnostic as it does not require the user to know the likelihood function they are trying to parametrize. This makes it useful for complex simulation models where the likelihood is either unknown or too expensive to compute numerically. We also think that defining filters is easier than defining likelihood weights to harmonize multiple conflicting evidence.

3 Application 1: Bio-economic Model

3.1 Description

In this section we build a conceptual model showing how to assess the status of a stock and estimate policy effects via rejection sampling. We show that assessment and policy analysis are possible even though all the model parameters remain unidentified. The results in this section are unsurprising but they are useful to illustrate the method and account for the uncertainty still present as we move to policy optimization.

We use a bio-economic model (Smith 1969; Seijo, Defeo, and Salas 1998) where a single logistic fish stock is targeted by boats whose number changes yearly in proportion to profits. It can be summarised by two difference equations:

$$\Delta_t f = \phi[pqB_t - c]f_t;\ \Delta_t B = rB_t[1 - \frac{B_t}{K}] - qfB_t$$

B_t is the biomass at year t, f_t is the effort (expressed as number of boats), p is the revenue per ton caught, $c$is cost per effort, K is the carrying capacity, q is catchability and ϕ represent how quickly boats enter or exit the fishery as a proportion of the profits observed.

Table 3.1 lists the parameter priors. Notice in particular the high carrying capacity: we expect this to be potentially a large stock at virgin levels. How many years to simulate is also an unknown parameter (T) as we are pretending not to know the age of the simulated fishery.

We want to condition this model on three pieces of information. First, we know that "real" landings in the current year were between 10,000t and 15,000t. Second, we know that the fishery is currently not making profits. Precise profit data is hard to obtain even in data rich fisheries, so the filter here is very wide: today's profits must be negative. Third, we know that CPUE was higher ten years ago than it is today. This information cannot be sufficient to estimate even a simple bio-economic model. It is enough however to make meaningful policy analysis.

Table 3.1: Table containing the priors of the simple bio-economic model. All are assumed uniform. The numbers themselves are just wide neighborhoods around the example values from table 2.1 in Seijo, Defeo, and Salas (1998)

Parameter	Minimum	Maximum	Meaning
p	30$	90$	revenue per ton of fish
c	15,000$	60,000$	annual cost per boat
q	0.0002	0.0008	catchability
K	1,750,000t	7,000,000t	carrying capacity
ϕ	0.0000025	0.00005	profit to boat gain
r	0.1	0.6	logistic growth
T	40	65	age of the fishery

3.2 Rejection sampling results

We run the model 1,000,000 times. Each run uses a random set of parameters drawn from table 3.1. We discard all runs that do not match our evidence. That is, we discard runs where catches at year T are not between 10,000 and 15,000t; we discard runs where profits at time T are positive and we discard runs where $CPUE_T > CPUE_{T-10}$. We accept 3,256 runs.

Even though our acceptance rate is 0.3%, figure 3.1 shows how none of the parameters can be identified given the information we have. Even worse, figure 3.2 proves that MSY (maximum sustainable yield) cannot be identified from the joint distribution of the logistic parameters. This is to be expected given the data limitations: one should not be able to identify MSY with only one year of landings and two qualitative indicators.

Figure 3.1: The marginal posterior distribution for each of the parameters. While the distributions are not uniform as the prior was, the density is positive through the entire domain of each of the parameters. We failed to identify any parameter.

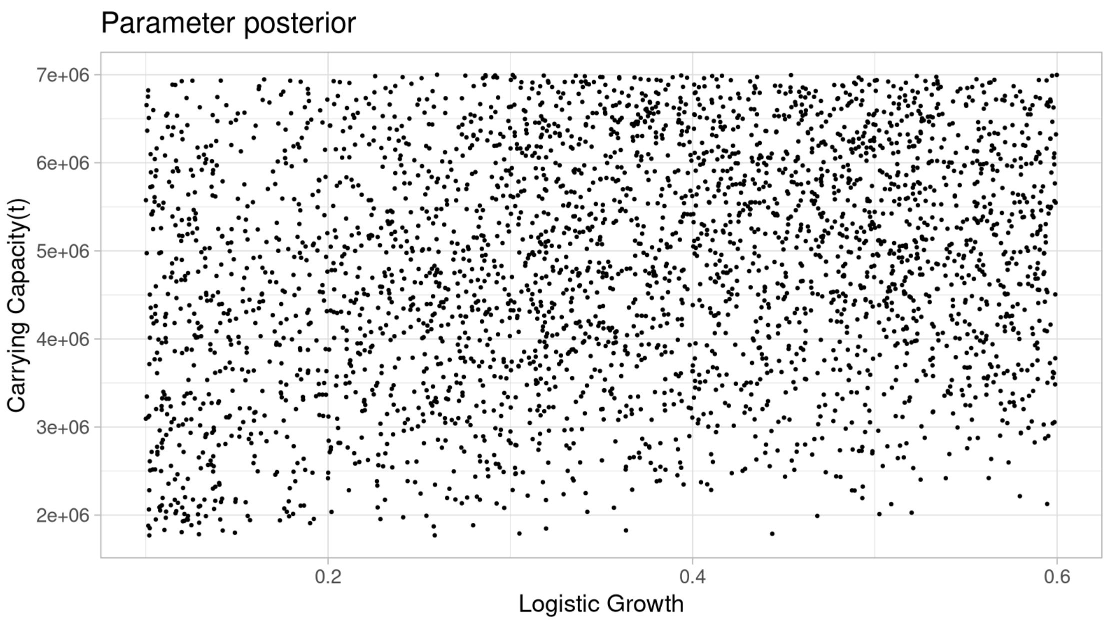

Figure 3.2: A scatter-plot where each point is an accepted run. Visually the density of points approximates the joint posterior distribution for logistic r and K. Again, no real identification is possible as fisheries have been accepted through the entire joint domain. It is therefore not possible to estimate MSY and other reference points given the evidence at hand. Compare this with the catch-MSY method which identifies a narrow triangle of values from which to compute MSY.

Even though no parameter can be identified, we still generate informative posteriors on some state variables, in particular current depletion $\frac{B_T}{K}$. In figure 3.3 we show the posterior distribution of depletion as we incrementally filter away simulations that do not match evidence. Total landings in tons alone do not help identify the status of the stock (as shown in panel B) . This is somewhat unexpected because landings are low compared to our high carrying capacity prior. However it is possible for landings to be low if we are barely scratching the surface of a large untapped stock (as shown in the bimodal $\frac{B_T}{K}$ posterior in panel B). Once we add the second piece of information (profits this year are negative) the depletion posterior in panel C shifts left: we know now that we are dealing with a depleted stock. Adding the third piece of information (CPUE today is lower than ten years ago) shifts the posterior further to the left, predicting a heavily depleted stock (as shown in panel D).

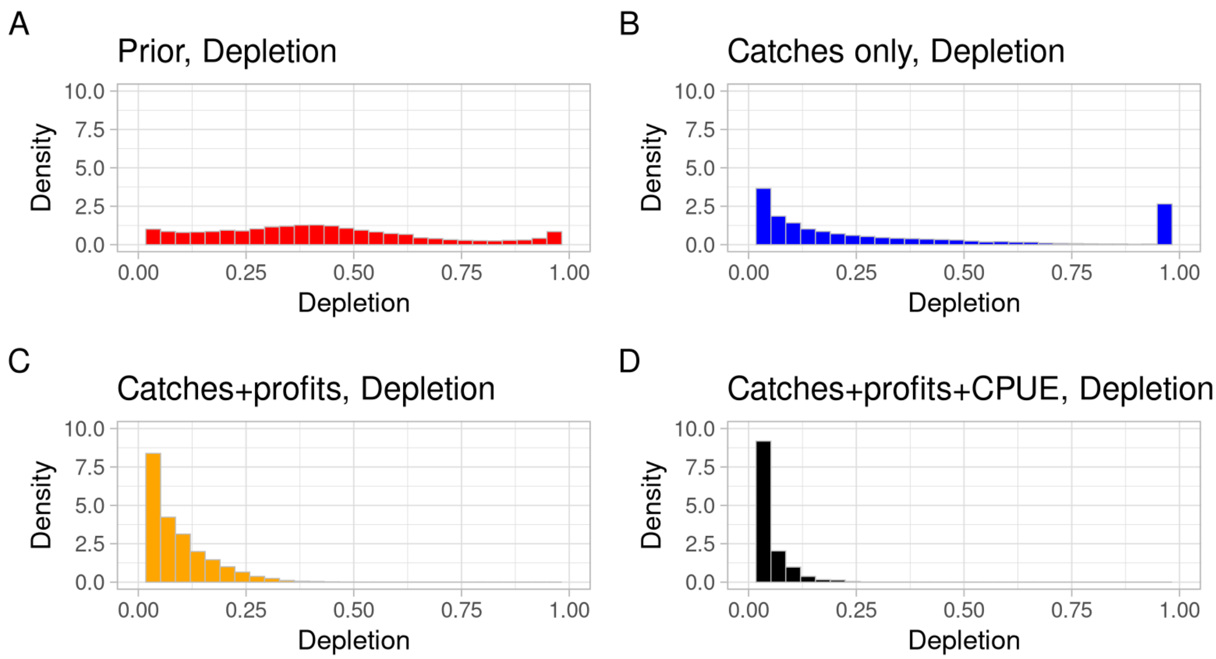

Figure 3.3: The distribution of this year depletion $\frac{B_T}{K}$ for all accepted simulations. A shows the prior: all simulations are accepted without looking at evidence; No depletion pattern is evident. B shows all the simulations whose last year's catch conforms to evidence; it is clear that this piece of information alone does not help with determining depletion. C shows the depletion of simulations after filtering for both catches and profits; we can now see that the fishery is overfished. D shows the depletion after filtering for catches, negative profits and CPUE being higher ten years ago; it is clear that the fishery is heavily depleted.

Negative profits and a decade-old fall in CPUE are critical pieces of information then. Because the simulated stock grows through a logistic curve, the drop in CPUE informs us that biomass must have been higher ten years ago. Negative profits tell us that the fishery is struggling with over-capacity, although it remains unclear whether the number of boats is unsustainable because of past overfishing or would be unsustainable even under optimal conditions. Combining these two indicators together with the high prior we had on K and the low posterior we observe in depletion $\frac{B_T}{K}$ implies that the current catches are small compared to what is achievable with a well managed fishery.

3.3 Policy analysis

We are interested in assessing the relative effectiveness of policies in achieving management objectives given the information we have on the fishery. Because the model is not identified, we cannot use parameters to identify a priori which policy will work. Instead we run each policy on each of 3,256 accepted fisheries in the ensemble.

We apply each policy to each fishery for 50 simulated years and average results across all years and fisheries. In table 3.2 we report average landings, profits and the percentage of years with depletion above $0.25K$and$0.5K$ for each simulated policy. Business as usual (no policy) results in landings of large quantities of fish, but inefficiently so as profits are overall negative. Most simulated fisheries remain trapped in a continuous boom-bust cycle. Setting a TAC as catches at year T (last year before policy) proves too restrictive: landings are minimal even after biomass recovers. This makes sense as we know the fishery is heavily depleted and current catches are likely far below sustainable values. Setting a maximum effort to 80% of boats at time T provides only moderately more landings than business as usual but larger profits as the policy prevents over-capacity during boom cycles.

We can compare these results to what they could have been had it been possible to identify MSY. Setting a TAC to the real MSY of each fishery provides the most landings but not the most profits as the fishery can still become crowded with many boats chasing an ever smaller share of the theoretically sound TAC.

Table 3.2: Average (across fisheries and time) results for each policy applied to every accepted fisheries. We list total landings (i.e. the average 50 years projected landings for all fisheries), total profits (i.e. the average 50 years sum of yearly profits for all accepted fisheries) and the percentage of simulated years any simulated fishery has B_t above 25% and 50% of carrying capacity

Average policy results

	Total Landings	Total Profits	Years $B_t >$ 25%K	Years $B_t >$ 50%K
Business as usual	6,497,560t	−$5,292,318	30.43%	13.59%
80% current boats	7,804,540t	$286,430,796	56.67%	36.90%
Last year catch TAC	432,365t	−$57,284,877	60.21%	51.12%
MSY TAC	9,561,784t	$68,281,386	57.93%	43.21%
Closed	0t	$0	69.78%	59.99%

We can then search for the policy that maximizes our objectives across the accepted fisheries(Bailey et al. 2018). Focusing on effort control, we look for the number of boats that would maximize profits over the next 50 years. Knowing the exact number of boats that would maximize profits requires a fully identified model. We instead set maximum effort to be a percentage of boats we observe this year (time T) and look for maximum profits "on average" across all accepted fisheries.

In table 3.3 we show the result of a grid search looking for the right % of current boats to set as maximum effort. Profits are maximized when setting maximum effort to between 50% and 60% of current effort.

Table 3.3: Average (across fisheries and time) results over 50 simulated years. We modify maximum effort allowed as a proportion of boats present at time T

Policy optimization

Looking for the maximum effort to impose as % of boats currently in the system

	Total Landings	Total Profits	Years B_t > 25%K	Years B_t > 50%K
10.0% current boats	2,359,611t	$122,542,205	**68.65%**	**58.42%**
20.0% current boats	4,155,692t	$208,558,019	67.43%	56.75%
30.0% current boats	5,468,769t	$263,642,642	66.17%	54.20%
40.0% current boats	6,390,060t	$294,053,626	64.72%	50.59%
50.0% current boats	7,009,285t	**$306,049,585**	62.89%	46.78%
60.0% current boats	7,409,192t	$305,568,079	60.84%	42.71%
70.0% current boats	7,656,195t	$297,696,871	58.70%	39.45%
80.0% current boats	7,804,540t	$286,430,796	56.67%	36.90%

90.0% current boats	7,914,020t	$275,320,579	54.74%	34.99%
100.0% current boats	**8,014,010t**	$266,004,973	53.09%	33.31%

4 Application 2: hairtail fishing in the Bungo channel

4.1 Description

In this example we model the Japanese coastal trolling fishery based in Usuki as described in Makino et al. (2017). As of 2011 it was composed of 45 boats, all smaller than 5 gross tons, targeting hairtail (*Trichiurus japonicus*). Sale price of hairtail depends on its size: letting it grow for 6 months can triple its value. The contribution of the Usuki trolling fishery to the overall fishing mortality is limited, however, as trolling boats share the stock with purse seiners who target other species and land hairtail as a byproduct. The fishery is not data-limited as Watari et al. (2017) presents a full stock assessment. The purpose of this example is one of data-degradation: we want to compare data-limited assessment against what is possible using the full data-set.

Makino et al. (2017) focuses on the economic side of the fishery: the number of fishers is in decline, most of them earn about 5-6M JPY which covers fixed and living costs but is not enough to save the 25M JPY needed to build a new boat. The focus on individual boats makes it ideal for an agent-based model analysis and here we create a simple POSEIDON (Bailey et al. 2018) application for it.

We sketch here the key ideas of the model but leave the full details in the appendix. The biological operating model is length-based, with daily growth between 5cm length-bins using the fixed boxcar method (Goudriaan 1986). Recruitment is noisy (with uniform distributed shocks) and happens twice a year (spring and autumn) assuming a Beverton-Holt stock-recruitment relationship. Geographically, fish biomass is spread out over a 5-by-5 grid, each cell a square 12km wide. We model each trolling boat separately. Boats can fish any day of the year and choose the fishing location by trial and error through the explore-exploit-imitate algorithm (Carrella, Bailey, and Madsen 2019).
We track the profits made by each boat. Those that manage to accumulate more cash than their target savings (after accounting for living and fixed costs) spawn a successor (that is, create a new boat that joins the fishery). Those that fail to cover their living costs will eventually leave.
The competing purse seiner fleet is not modeled explicitly and is instead a simple constant fishing mortality F applied to the stock at the end of each year.

We run the model for at most 45 years, starting the trolling fishery *ab novo* at year 1. We condition the model on six pieces of information. First, total landings (trolling and purse seiners combined) in the fishery have never been above 15,000t. This is evident from figure 2 in Watari et al. (2017). Second, current landings of the Usuki trolling fleet are between 250t and 1,850t. The lower bound is used as an historical target in Makino et al. (2017), the upper bound is the non-net fishing mortality computed for 2011 in Watari et al. (2017). Third, current SPR is below sustainable levels and we will accept fisheries whose current observed SPR is between .10 and .25. Watari et al. (2017) cites an SPR for the fishery of .21 in 2011. Fourth, the Usuki trolling fishery today has less than 60 active boats. Fifth, the Usuki trolling fishery represents at most 30% of the total landings for the stock. Sixth, the fishery is at least 30 years old.

Because our model is length rather than age based, we compute SPR via a length-based approximation, assuming median life-history parameters for the species (using Fishlife from Thorson et al. 2017; rFishBase from Boettiger, Lang, and Wainwright 2012). When computing SPR we do not assume we know the correct life-history parameters within the simulation. In other words, the simulated SPR we compute within the agent-based model is subject both to noise due to the length-to-age conversion as well as from the possibility of using the wrong life history parameters. As such we expect the SPR filter to be quite noisy.

This POSEIDON application is simple and lacks many details, *inter alia* it assumes knife-edge maturity (maturity is 0 for all fish below length at maturity, and 1 for all others), a simple geography, no seasonal variation in ex-vessel prices and simple stochasticity for the bi-annual spawning pulses. In spite of its simplicity it still contains many more parameters (see table 3.1) than observations available and we cannot hope for identification. We run the model 354,775 times[2] and accept 743 fisheries.

Table 4.1: A table with all unknown parameters from the POSEIDON simulation we need to randomize. Some (like cost data) have narrow priors but others (like catchability and steepness) are wide

Variable	Distribution	Meaning	Source
catchability	$U[10^{-5}, 10^{-9}]$	% of biomass caught per hour spent fishing	-
S_1	$U[6.2, 9.3]$	Logistic selectivity, first parameter	(size limit at 25cm)
S_2	$U[0.17, 0.26]$	Logistic selectivity, second parameter	(size limit at 25cm)
Cell with most biomass	$U[1, 5] \times U[1, 5]$	Cell in the 5x5 grid with	-

[2] 24 hours of computing time

		the most biomass	
Biomass smoothing	$U[0.1, 1]$	Parameter defining the geographical dispersion of biomass	-
Virgin recruits R_0	$U[37M, 62M]$	Number of annual recruits at virgin biomass	R_0 that would generate 10,000t to 200,000t of virgin biomass
Φ	$U[0.27, 0.31]$	Ratio R_0 to spawning stock biomass	Implied by knife-edge maturity from Fishlife median L_{mat}
Steepness	$U[0.2, 0.99]$	Proportion of current recruits to virgin when 20% of biomass left	-
Initial depletion	$U[0.7, 1]$	Proportion of biomass to virgin available as simulation starts	-
Life history parameters L_∞, K, L_{mat}, M	Drawn from Fishlife	Parameters governing the speed of growth for all fish	Drawn from Fishlife multi-normal distribution
Allometric parameters a, b	Bootstrap sampled from rFishbase	Parameters converting fish length to weight	Sampled with replacement from Fishbase
Hourly variable cost (JPY/hr)	$U[100, 140]$	Monetary cost per hour spent at sea (fuel,engine)	Makino et al. (2017)
Hourly effort cost (JPY/hr)	$U[700, 800]$	Monetary cost per hour	Makino et al. (2017)

		spent fishing (film, bait and boxing)	
Hold size (kg)	$U[1000, 3000]$	Maximum amount of fish transportable by a boat	5GT boats
Savings target (JPY)	$U[20M, 30M]$	Money to accumulate to spawn additional boat	Makino et al. (2017)
Yearly Expenses	$U[4.5M, 7M]$	Yearly expenditure (living costs, plus fixed costs)	Makino et al. (2017) and census
Daily probability of fishing	$U[.75, .9]$	Probability of going fishing each day of the year	Comparing boat numbers with total effort recorded
Yearly exogenous fishing mortality	$U[.3, .8]$	Mortality rate due to other fishers targeting the stock	-

4.2 Value of evidence

Conditioning on data creates colliders (Monty Hall effects in Burns and Wieth 2004): correlations between parameters that were originally independent. These correlations can overturn the *a priori* meaning of some evidence and parameters. Rejection sampling here is useful because it identifies these collider biases automatically and contextualizes the evidence.

Length-based SPR can be an ambiguous piece of evidence. We measure a low SPR when the fish caught is smaller than expected. This could be a sign of overfishing, but could also be a sign of our expectations being wrong and the fish being naturally smaller than assumed. We can observe low SPR without overfishing if the "real" $\frac{M}{K}$ ratio is higher than the one used to compute SPR (which in this example is assumed to be the median value from FishLife).

When $\frac{M}{K}$ is high, the fish grows slowly and is likely to die young. All else equals then a high $\frac{M}{K}$ makes the fishery less economically viable. Not everything else is equal among the accepted fisheries *(a posteriori)* however, because we know the fishery has survived for at least 30 years. Knowing both that $\frac{M}{K}$ is high and that the fishery survived implies that other model parameters must have changed to make it so.

Fisheries with high $\frac{M}{K}$ that survive are quantity-driven rather than value-driven. If fish matures slowly and is less likely to survive, boats need to catch it when it is shorter. Because smaller fish are less profitable, boats need to catch more of them per unit of effort. The end result is that high $\frac{M}{K}$ fisheries are associated with parameters that make each unit of effort more profitable. Because entry and exit are driven by savings and therefore profits per unit of effort, high $\frac{M}{K}$ fisheries that are accepted will sustain a larger fleet in the equilibrium.

More precisely, rejection sampling exploits the wide FishBase (Froese and Pauly 2012) priors on length-to-weight parameters a, b. Before rejection sampling, the length-to-weight parameters a, b are uncorrelated with $\frac{M}{K}$, as they are drawn independently. Among the accepted fisheries however the correlation coefficient is 0.88: the parameters must be all high or all low. In low $\frac{M}{K}$ low a, b fisheries, many fish will reach 50cm but each will weigh 200g; in high $\frac{M}{K}$ high a, b fisheries, fewer fish will reach 50cm but each will weigh 600g, tripling the revenue for the fisher. Other combinations wouldn't work: high $\frac{M}{K}$ low a, b would not be sustainable, low $\frac{M}{K}$ high a, $b$fisheries would sustain too large of a boat population above the set filters.

We can combine these two analyses together: a low SPR implies a high $\frac{M}{K}$ which implies a larger fleet in the long-run. Or in other words, the lower the SPR the more likely the fishery will have more boats in the longer run.

Figure 4.1 shows the evolution of the distribution of $\frac{M}{K}$ subject to different sets of filters. In panel A we show the prior, which we draw directly from Fishlife (Thorson et al. 2017), and the posterior after passing all filters. The posterior has shifted right (lower $\frac{M}{K}$) and has lower variance.

In panel B we compute how $\frac{M}{K}$ would have been distributed had the SPR filter been higher (above 0.3) or lower (below 0.1). A high SPR is associated with a low $\frac{M}{K}$ while a low SPR is associated with a bi-modal $\frac{M}{K}$: the right peak represents fisheries that are effectively struggling while the left peak represents fisheries for which SPR is a too pessimistic indicator because it assumes too low $\frac{M}{K}$.

Sensitivity of length-based SPR to mispecifications of $\frac{M}{K}$ is well understood (Hordyk et al. 2014; Froese et al. 2019b, 2019a;Hordyk et al. 2019) and here we are simply picking up the conflict between Fishlife's wide priors and using its modal values to compute SPR.

In panel C we revert back to the original filters but we look at the distribution of $\frac{M}{K}$ for fisheries where more than 60 boats survive. Again the distribution becomes bi-modal with most of the density given to fisheries with very high $\frac{M}{K}$ (above 2).

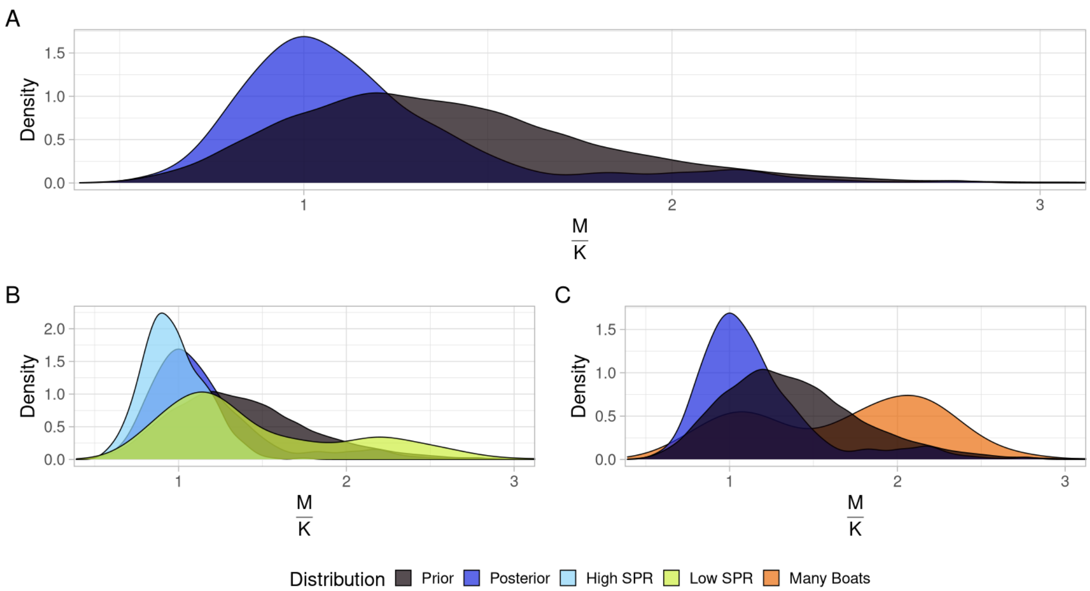

Figure 4.1: Ratio between natural mortality M and growth rate K; subfigure A shows prior and posterior distribution of the $\frac{M}{K}$, the prior is drawn from FishLife and the posterior represents the fisheries that passed all filters, the posterior has shifted mostly to the left ;subfigure B compares the $\frac{M}{K}$ posterior had we only SPR above 30 (cyan) or below 10 (green); subfigure C shows the posterior achieved by accepting only fisheries who have more than 60 boats active today (orange).

Combining these two collider effects explains why runs with lower SPR will "do better" (attract more boats) in the future, as shown in figure 4.2. As explained above, low SPR is associated with high $\frac{M}{K}$ and high $\frac{M}{K}$ is associated with quantity-driven fisheries with many boats. High participation fisheries are usually filtered away because of our 60 boat limit but it is possible that today's participation is temporarily low due to recent recruitment failures (5-10 years ago). In context then, a low SPR may indicate a high participation fishery that is only temporarily under-strength (that is, some boats have exited but will re-enter). This kind of fishery will eventually snap back to its long-term high participation equilibrium.

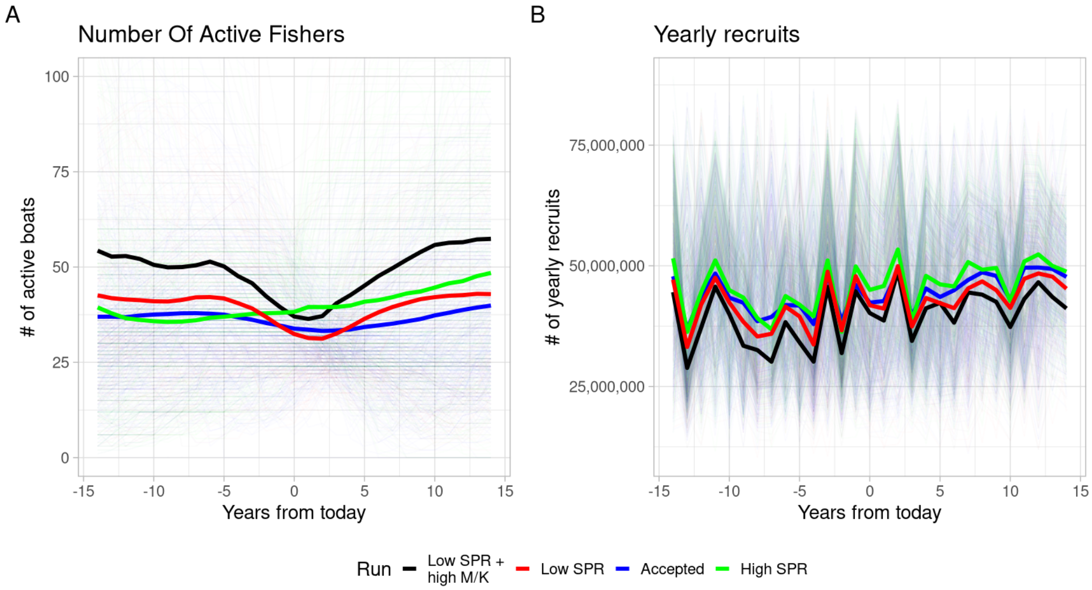

Figure 4.2: Projected and past realizations of participation and recruitment for accepted fisheries as well as fisheries that would have been accepted had SPR evidence been different. Each line is a separate simulation, bolded lines are median values. The subset of runs with low SPR and also M/K ratio above two have been grouped separately (in black). Low SPR today implies higher participation tomorrow if associated with higher M/K. This is better explained by looking at recruitment (subfigure B): these fisheries in equilibrium have higher participation but are over-correcting to low recruitment pulses 5 to 10 years ago and are likely to rebound.

4.3 Policy analysis

For the accepted fisheries, business as usual (which here implies only minimum catch size requirements) will not result in any long-term change in the number of fishers or biomass as shown in figure 4.3. It remains however a depleted fishery with biomass at approximately 15% of carrying capacity, which justifies additional policies.

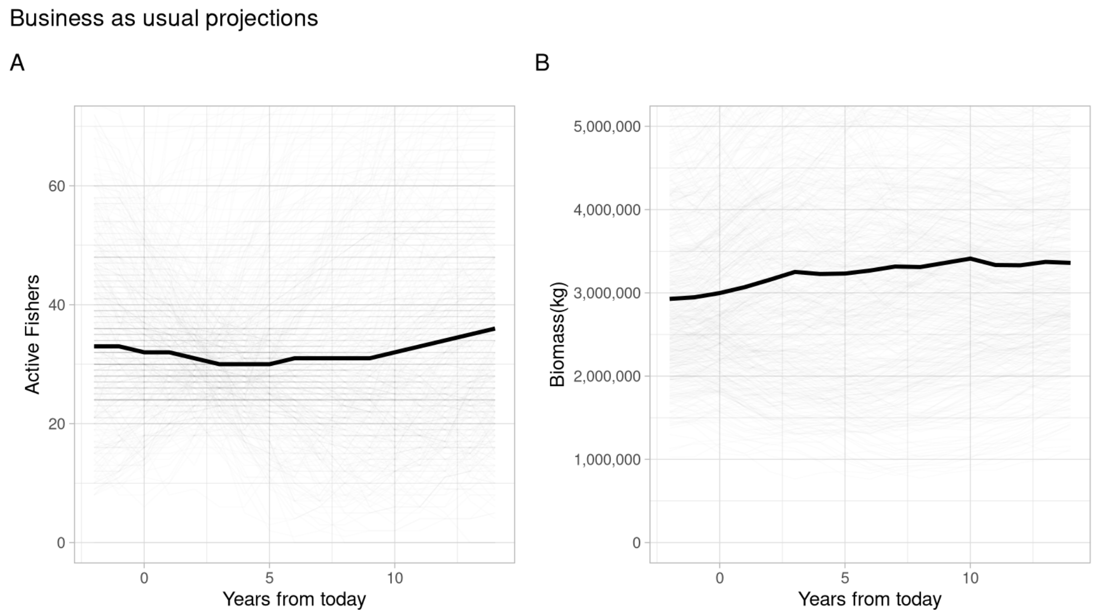

Figure 4.3: Projected participation and biomass for the accepted fisheries when no policy is implemented. Each line is a simulation,the bolded line is the median yearly value. While each individual run will exhibit noise and short term trends, the overall dynamic is one of long-term equilibrium around current values.

We assume here that we can only enforce policies on the Usuki trolling fleet and cannot affect fishing mortality outside of it. Because we constrained the Usuki fishery to represent at most 30% of the landings, regulations are unlikely to achieve much rebuilding of the fish stock.

However, it may be possible to improve the economic health of this fleet in two ways. First, management which results in marginal improvements in biomass may increase CPUE and therefore profitability. Second, policies that increase the average size of fish caught may increase earnings for each ton of fish landed.

We compare three policies. The first is to make the fishery closed access, i.e., disallowing new entrants to the fishery and re-entry. The second combines the closed access with a limit of 200 days of fishing a year. The third is to combine the closed entry with a mandated gear modification that selects for larger fish (specifically, we model this as a 10% increase in S_1 and a 10% decrease inS_2 of the logistic selectivity). Changing hook and bait types were suggested for this fishery in Hirose et al (2017).

The impacts of these policies across a range of outcomes are shown in figure 4.4 below. For the overall objective of increasing profits and biomass, closed entry alone performs just as well as more complex policies. We see that seasonal closures do succeed in enabling boats to achieve higher profits per trip and that the gear modification does increase revenue per unit caught; however, these benefits are negated over time by the unmanaged purse seine fishery which increases its relative importance in landings as landings from the Usuki fleet decline. Further, while seasonal closures generate higher profits per trip,these fail to cover fixed costs.

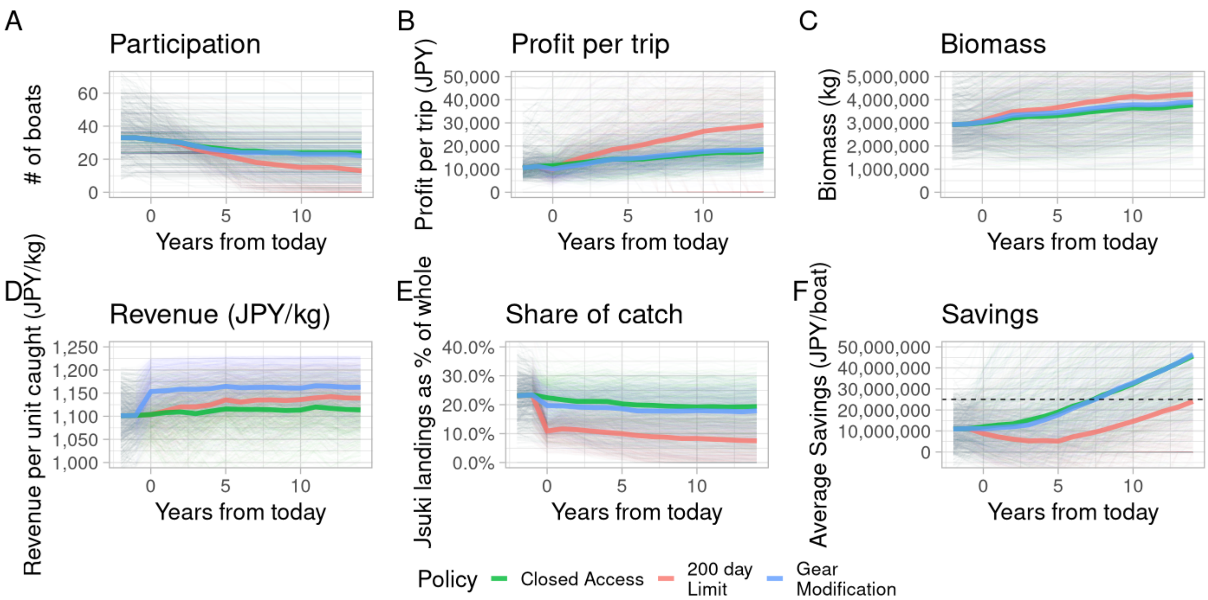

Figure 4.4: Key indicators projections for three policies (no entry or re-entry, 200 days fishing limit, higher selectivity); each line is a separate fishery, bolded lines are the median yearly value. Better selectivity increases revenue per kg caught (subfigure D) and fishing limit increases profits per trip (subfigure B). In general however all policies cause more landings to occur outside the Usuki trolling fishery (subfigure E) and selectivity improvements do not improve overall savings any better than just closing to new entrants (subfigure F)

4.4 Comparison to data-rich results

Biomass in the stock assessment produced in Watari et al. (2017) is estimated at 4,896t in 2011. Here the mean current biomass is 3,477t with standard deviation of 1,806t. This is a one standard deviation difference using only limited evidence.

The main difference between our work and the data-rich simulations in Makino et al. (2017) involves future projections. POSEIDON projections are more optimistic under the status quo. In the original data-rich work the unmanaged fishery constantly deteriorates, both biologically and economically. Our projections instead show a long term equilibrium around current values.

The discrepancy can be explained by the underlying differences in the economic model. In our model boats quit when earnings dry up and bank balances turn negative. This helps the fishery recover over time and avoid collapses. In Makino et al. (2017) however, fishing mortality is a free parameter and as such fishing pressure can continue unabated.

One contributing factor for discrepancies is our choice of life-history parameters. We pretended not to know the stock assessment parameters and instead used “off-the-shelf”

FishLife distributions as priors. These are wide and not necessarily centered near the real values from the stock assessment. For example, in its stock assessment Watari et al. (2017) assumes that the maximum age for hairtail is 5 (the maximum ever observed in the area). Using the Quinn and Deriso (1999) rule of thumb (Hordyk et al. 2014) this would imply a natural mortality of at least 0.9. The fishlife average is 0.39 with a standard deviation of 0.11. For this particular application then, better priors on life history parameters could have improved the performance and realism of rejection sampling.

5 Discussion and conclusion

We presented a method fit models to limited data for policy analysis accounting for identification failures and the resulting uncertainty. We presented two examples to show how the method is agnostic both to the kind of model used (difference equation or agent-based model) and the kind of evidence brought to bear.

5.1 Rejection sampling and the law of decreasing credibility

Fitting complex models to limited data generates identification issues(see Canova and Sala 2009): the data may not be enough to fix the value of all (or any) parameters. Some parameters may not affect the likelihood function (under-identification) or affect it only proportionally with others (partial identification). The curvature of the likelihood function around each parameter may be small enough to make likelihood maximization difficult in practice (weak identification).The common solution to lack of identification is to add more assumptions to achieve it with the data we have(see for example Manski 2007). These assumptions often involve data we couldn't collect and are subject to the "law of decreasing credibility"(Manski 2003): the more we want to reduce uncertainty the less realistic our assumptions become.

For certain questions however reducing the uncertainty on the parameter posterior is unnecessary. If we are interested in the average behaviour of the model given some observed data we may not require point estimation for the parameters. As long as we can sample from the set of fisheries that match the data, we can just average out their projections as the central limit theorem will bound their expectation.

It should be noted however that simulation alone does not avoid the "law of decreasing credibility." The model that generates candidate fisheries contains many assumptions of varying degrees of credibility. The advantage however is that the model assumptions tend to be about the dynamics of the system and theoretical causal mechanisms (the way things work) rather than about data and evidence we do not possess (the way things are). The cost of simulation is that we can only focus on average behaviour of the model which in some applications may not be enough to develop robust policies.

5.2 Theory as a substitute for data

Varian (1989) summarised the economics approach to under-identification with the phrase "theory as a substitute for data." When there are not enough observations for an extrapolation, theory can help connect the evidence we do have with the parameters we are interested in. This is useful both in the sense of telling us what parameter matters and what observations can help identify it.

Simulation and rejection sampling are a way to automatically connect theory and data in a model-agnostic way in exchange for a steep computational price. It is a computational substitute to drawing causal directed acyclic graphs connecting evidence and parameters through conditioning (Pearl 2017, 2014).

The cost of leaning on theoretical causal mechanisms is that filtering is only as good as the model we are using. We face a tradeoff: we would like the model to be simple so that it runs fast and produces a large ensemble of accepted fisheries, but also complicated enough that our causal mechanisms are realistic. We also need to add enough complexity in the model to produce the evidence we need to filter against. For example, if we observe catch-at-length then our model must contain age or length structure. As computing becomes cheaper, this tradeoff will tilt towards more complicated models.

5.3 Computational limits and solutions

Feeding a model random parameters in the hope that it will eventually pass all filters is computationally expensive. As we accumulate data and add filters, we will accept fewer runs. In theory this is not a problem: once we have enough data we can revert back to the usual likelihood-based stock assessment techniques and avoid this curse of dimensionality. In practice however we might find ourselves with enough data to make rejection sampling impractical while at the same time not enough to run a stock assessment.
Under these circumstances we need to improve the efficiency.

Computational efficiency could be improved in two ways. If we are willing to establish priors on the original parameters we can exploit advances in approximate Bayesian computation methods that replace simple rejection with Markov-chain Monte Carlo or sequential Monte Carlo(Beaumont et al. 2009). These techniques would become important with more evidence or slower models but are sometimes slower than plain rejection when applied to simple problems (Grazzini, Richiardi, and Tsionas 2017). An additional advantage is that the convergence of ABC methods is guaranteed under mild assumptions (Barber, Voss, and Webster 2015).

If we are not willing or able to state priors, Williams et al (2020) proposes using optimization via genetic algorithms (targeting both closeness to data and phenotype diversity) to sample efficiently a large number of valid scenarios. With this technique however no guarantee can be made about the representativeness of the sampled fisheries and we would need to analyze policy projections through deep-uncertainty methods (i.e. mini-max or looking only at unanimous results); this is the price we pay for lacking good priors.

Bibliography

Bailey, Richard M., Ernesto Carrella, Robert Axtell, Matthew G. Burgess, Reniel B. Cabral, Michael Drexler, Chris Dorsett, Jens Koed Madsen, Andreas Merkl, and Steven Saul. 2018. "A computational approach to managing coupled human–environmental systems: the POSEIDON model of ocean fisheries." *Sustainability Science*, June, 1–17. https://doi.org/10.1007/s11625-018-0579-9.

Baker, Christopher M., Ascelin Gordon, and Michael Bode. 2017. "Ensemble ecosystem modeling for predicting ecosystem response to predator reintroduction." *Conservation Biology* 31 (2): 376–84. https://doi.org/10.1111/cobi.12798.

Barber, Stuart, Jochen Voss, and Mark Webster. 2015. "The rate of convergence for approximate Bayesian computation." *Electronic Journal of Statistics* 9: 80–105. https://doi.org/10.1214/15-EJS988.

Beaumont, Mark A, Jean-Marie Cornuet, Jean-Michel Marin, and Christian P Robert. 2009. "Adaptive approximate Bayesian computation." *Biometrika* 96 (4): 983–90. https://doi.org/10.1093/biomet/asp052.

Beaumont, Mark A, Wenyang Zhang, and David J Balding. 2002. "Approximate Bayesian computation in population genetics." *Genetics* 162 (4): 2025–35. https://doi.org/Genetics December 1, 2002 vol. 162 no. 4 2025-2035.

Boettiger, C., D. T. Lang, and P. C. Wainwright. 2012. "Rfishbase: Exploring, manipulating and visualizing FishBase data from R." *Journal of Fish Biology* 81 (6): 2030–39. https://doi.org/10.1111/j.1095-8649.2012.03464.x.

Burns, Bruce D., and Mareike Wieth. 2004. "The collider principle in causal reasoning: Why the Monty Hall dilemma is so hard." *Journal of Experimental Psychology: General* 133 (3): 434–49. https://doi.org/10.1037/0096-3445.133.3.434.

Canova, Fabio, and Luca Sala. 2009. "Back to square one: Identification issues in DSGE models." *Journal of Monetary Economics* 56 (4): 431–49. https://doi.org/10.1016/j.jmoneco.2009.03.014.

Carrella, Ernesto, Richard M. Bailey, and Jens Koed Madsen. 2019. "Repeated discrete choices in geographical agent based models with an application to fisheries." *Environmental Modelling and Software* 111 (January): 204–30. https://doi.org/10.1016/j.envsoft.2018.08.023.

Carruthers, Thomas R., André E. Punt, Carl J. Walters, Alec MacCall, Murdoch K. McAllister, Edward J. Dick, and Jason Cope. 2014. "Evaluating methods for setting catch limits in data-limited fisheries." *Fisheries Research* 153 (May): 48–68. https://doi.org/10.1016/j.fishres.2013.12.014.

Cramer, Simon, and Torsten Trimborn. 2019. "Stylized Facts and Agent-Based Modeling." https://stooq.com.

Franke, Reiner, and Frank Westerhoff. 2016. "Why a simple herding model may generate the stylized facts of daily returns: explanation and estimation." *Journal of Economic Interaction and Coordination* 11 (1): 1–34. https://doi.org/10.1007/s11403-014-0140-6.

Froese, Rainer & Pauly, D. (2012). FishBase. Available at www.fishbase.org.

Froese, Rainer, Henning Winker, Gianpaolo Coro, Nazli Demirel, Athanassios C Tsikliras, Donna Dimarchopoulou, Giuseppe Scarcella, Wolfgang Nikolaus Probst, Manuel Dureuil, and Daniel Pauly. 2019a. "On the pile-up effect and priors for Linf and M/K: response to a comment by Hordyk et al. on 'A new approach for estimating stock status from length frequency data'." *ICES Journal of Marine Science* 76 (2): 461–65. https://doi.org/10.1093/icesjms/fsy199.

———. 2019b. "A new approach for estimating stock status from length frequency data." *ICES Journal of Marine Science* 76 (1): 350–51. https://doi.org/10.1093/icesjms/fsy139.

Gourieroux, C., Monfort, A., & Renault, E. (1993). Indirect Inference. *Journal of Applied Econometrics*, 8, 85–118. https://doi.org/10.2307/2285076

Goudriaan, J. 1986. "Boxcartrain Methods for Modelling of Ageing, Development, Delays and Dispersion." In *The Dynamics of Physiologically Structured Populations*, edited by J. A. J. Metz and O. Dierkmann, 1st ed., 453–73. Berlin: Springer.

Grazzini, Jakob, Matteo G. Richiardi, and Mike Tsionas. 2017. "Bayesian estimation of agent-based models." *Journal of Economic Dynamics and Control* 77 (April): 26–47. https://doi.org/10.1016/j.jedc.2017.01.014.

Grimm, Volker, Eloy Revilla, Uta Berger, Florian Jeltsch, Wolf M. Mooij, Steven F. Railsback, Hans Hermann Thulke, Jacob Weiner, Thorsten Wiegand, and Donald L. DeAngelis. 2005. "Pattern-oriented modeling of agent-based complex systems: Lessons from ecology." https://doi.org/10.1126/science.1116681.

Hartig, F., J. M. Calabrese, B. Reineking, T. Wiegand, and A. Huth. 2011. "Statistical inference for stochastic simulation models - theory and application." *Ecology Letters* 14 (8): 816–27. https://doi.org/10.1111/j.1461-0248.2011.01640.x.

Hirose, T., Sakurai, M., Watari, S., Ogawa, M., & Makino, M. (2017). Conservation of small hairtail Trichiurus japonicas by using hooks with large artificial bait: Effect on the trolling line fishery. *Fisheries Science*, 83(6), 879–885. https://doi.org/10.1007/s12562-017-1142-9

Hordyk, Adrian R., and Thomas R. Carruthers. 2018. “A quantitative evaluation of a qualitative risk assessment framework: Examining the assumptions and predictions of the Productivity Susceptibility Analysis (PSA).” *PLoS ONE* 13 (6). https://doi.org/10.1371/journal.pone.0198298.

Hordyk, Adrian R, Jeremy D Prince, Thomas R Carruthers, and Carl J Walters. 2019. “Comment on "a new approach for estimating stock status from length frequency data" by Froese et al. (2018).” *ICES Journal of Marine Science* 76 (2): 457–60. https://doi.org/10.1093/icesjms/fsy168.

Hordyk, Adrian R., Kotaro Ono, Keith Sainsbury, Neil Loneragan, and Jeremy Prince. 2014. “Some explorations of the life history ratios to describe length composition, spawning-per-recruit, and the spawning potential ratio.” In *ICES Journal of Marine Science*, 72:204–16. 1. https://doi.org/10.1093/icesjms/fst235.

Lee, J. S., Filatova, T., Ligmann-Zielinska, A., Hassani-Mahmooei, B., Stonedahl, F., Lorscheid, I., Voinov, A., Polhill, G., Sun, Z., & Parker, D. C. (2015). “The complexities of agent-based modeling output analysis”. *Journal of Artificial Societies and Social Simulation*, 18(4). https://doi.org/10.18564/jasss.2897

Makino, Mitsutaku, Shingo Watari, Taro Hirose, Kentaro Oda, Masahito Hirota, Atsushi Takei, Michio Ogawa, and Hiroshi Horikawa. 2017. “A transdisciplinary research of coastal fisheries co-management: the case of the hairtail Trichiurus japonicus trolling line fishery around the Bungo Channel, Japan.” *Fisheries Science* 83 (6): 853–64. https://doi.org/10.1007/s12562-017-1141-x.

Mangel, Marc, Alec D. MacCall, Jon Brodziak, E. J. Dick, Robyn E. Forrest, Roxanna Pourzand, and Stephen Ralston. 2013. “A perspective on steepness, reference points, and stock assessment.” Edited by Kenneth Rose. *Canadian Journal of Fisheries and Aquatic Sciences* 70 (6): 930–40. https://doi.org/10.1139/cjfas-2012-0372.

Manski, Charles F. 2003. *Partial identification of probability distributions*. New York, New York, USA: Springer.

———. 2007. *Identification for Prediction and Decision*. https://books.google.com.tw/books/about/Identification_for_Prediction_and_Decisi.html?id=KcW-820XgeQC&printsec=frontcover&source=kp_read_button&redir_esc=y#v=onepage&q&f=false.

Martell, Steven, and Rainer Froese. 2013. “A simple method for estimating MSY from catch and resilience.” *Fish and Fisheries* 14 (4): 504–14. https://doi.org/10.1111/j.1467-2979.2012.00485.x.

Mildenberger, Tobias Karl, Marc Hollis Taylor, and Matthias Wolff. 2017. “TropFishR : an R package for fisheries analysis with length-frequency data.” Edited by Samantha Price. *Methods in Ecology and Evolution* 8 (11): 1520–27. https://doi.org/10.1111/2041-210X.12791.

Pearl, Judea. 2014. "The Causal Foundations of Structural Equation Modeling." *Handbook of Structural Equation Modeling*, no. June: 68–91. http://ftp.cs.ucla.edu/pub/stat_ser/r370.pdf.

———. 2017. "A Linear 'Microscope' for Interventions and Counterfactuals DOI." *J. Causal Infer*. https://doi.org/10.1515/jci-2017-0003.

Prince, J. and Hordyk, A. 2018. What to do when you have almost nothing: a simple quantitative prescription for managing extremely data-poor fisheries. *Fish and Fisheries*, 20, 224-238. doi:10.1111/faf.12335.

Pudlo, P., Marin, J. M., Estoup, A., Cornuet, J. M., Gautier, M., & Robert, C. P. (2015). "Reliable ABC model choice via random forests". *Bioinformatics*, 32(6), 859–866. https://doi.org/10.1093/bioinformatics/btv684

Quinn, Terrance J., and Richard B. Deriso. 1999. *Quantitative Fish Dynamics*. Oxford University Press.

Russell, Stuart, and Peter Norvig. 2009. *Artificial Intelligence: a modern approach*. 3rd ed. Upper Saddle River, NJ, USA: Prentice Hall Press.

Seijo, Juan Carlos, Omar Defeo, and Silvia Salas. 1998. *Fisheries bioeconomics Theory, modelling and management*. FAO. http://www.fao.org/3/W6914E/W6914E02.htm.

Sisson, Scott A. 2018. *Handbook of Approximate Bayesian Computation*. https://doi.org/10.1201/9781315117195.

Smith, Vernon L. 1969. "On Models of Commercial Fishing." *Journal of Political Economy* 77 (2): 181–98. https://doi.org/10.1086/259507.

Sparre, Per, and Siebren C. Venema. 1998. *Introduction to tropical fish stock assessment Introduction to tropical fish stock assessment Part I : Manual*. Vol. 306. https://ci.nii.ac.jp/naid/10030499433/.

Thorson, James T., Stephan B. Munch, Jason M. Cope, and Jin Gao. 2017. "Predicting life history parameters for all fishes worldwide." 8. Vol. 27. https://doi.org/10.1002/eap.1606.

Varian, Hal. 1989. "What Use is Economic Theory?"

Watari, Shingo, Syunji Tokumitsu, Taro Hirose, Michio Ogawa, and Mitsutaku Makino. 2017. "Stock structure and resource management of hairtail Trichiurus japonicus based on seasonal broods around the Bungo Channel, Japan." *Fisheries Science* 83 (6): 865–78. https://doi.org/10.1007/s12562-017-1140-y.

Williams, T. G., Guikema, S. D., Brown, D. G., & Agrawal, A. (2020). Assessing model equifinality for robust policy analysis in complex socio-environmental systems. *Environmental Modelling and Software*, 134, 104831. https://doi.org/10.1016/j.envsoft.2020.104831

Yin, Y., Aeberhard, W. H., Smith, S. J., & Mills Flemming, J. (2019). Identifiable state-space models: A case study of the Bay of Fundy sea scallop fishery. Canadian Journal of Statistics, 47(1), 27-45.

Appendix A: length-based Bungo channel fishery model

Geography

The fishery is spread out over a 5 by 5 grid, each cell a square 12 km wide. The port is assumed to be the same for all boats and is placed at the top right corner of the grid. Each cell contains a separate population of fish that grows and is affected by natural mortality independently from the other cells. There is no fish movement between cells except for new recruits which are generated by pooling all mature populations across each cell and then distributing these recruits to each cell proportional to their carrying capacity.

The total carrying capacity, the proportion of recruits received each spawning period and the initial population are allocated to cells unequally. One cell at random in the grid is defined as the “peak” and other cells are allocated a diminishing proportion of initial population and carrying capacity depending on their Manhattan distance from the peak. More precisely the amount of carrying capacity allocated to a cell is proportional to d^{λ} where λ is a smoothing parameter selected by rejection sampling and $d$is the Manhattan distance.

Geography serves two purposes: distance from port represents increasing costs (in terms of time travelled as well as monetarily) and unequal distribution of fish represents some areas being more productive and better for fishing than others.

Biology

In each cell, the model tracks fish abundance in 5cm bins; growth and mortality occur daily and locally (i.e. independently for each cell) while recruitment occurs only twice a year and does so globally.

Fish growth is simulated using the fixed boxcar method (Goudriaan and van Roermund 1999) where daily growth for each bin is assumed to be the derivative of the Von Bertalanffy function measured at the mid point of the bin. For example given N_1 fish in bin 1 (the number of local fish of size between 0 and 5cm), the number of fish that graduates daily to the second bin is:

$$Graduates_{1\rightarrow 2} = N_1 \frac{K L_{\infty} - 2.5}{365*5}$$

After growth, daily natural mortality takes places where survivors are

$$N_{tomorrow} = N_{today} e^{-\frac{M}{365}}$$

where M is the instantaneous yearly mortality rate assumed constant across length bins. Number of fish per bin is allowed to be decimal and no rounding occurs after growth or mortality events.

We convert abundance in bin to weight (for the purpose of counting biomass and landings weights) we use the allometric length-to-weight formula of form:

$$W = aL^b$$

where a and b are allometric weight parameters drawn together from fishbase priors and subject to rejection sampling.

Recruitment occurs twice a year on 1st of May and 1st of October. The number of recruits follows the Beverton-Holt formula :

$$R = u \times \frac{4\,h\,R_0 SSB}{R_0 \phi(1-h)+(5h-1)SSB_0}$$

Where steepness h, virgin recruits $R_0$and ϕ are assumed unknown and found through rejection sampling; SSB is the current biomass of all mature fish. We assume knife-edge maturity given unknown length at maturity parameter L_{mat} (drawn from FishLife prior and then subject to rejection sampling). The multiplier u is uniformly distributed between 0.25 and 0.75, so that on average yearly Beverton-Holt recruits will spawn in equal numbers in spring and fall, but with the possibility of some years producing more or less recruits than expected by the yearly recruitment formula.

Notice that recruitment is modeled differently here than in the Hairtail stock assessment in Watari et al (2017): there recruitment is simulated twice a year as a fixed proportion of current spawning stock biomass with a regular 5 year cycle of strong recruitments, simulated by increasing the recruitment per spawning biomass multiplier for those years.

Recruits are computed for the whole fishery by pooling all spawning stock biomass, and are then re-allocated to each cell in their first bin (0 to 5cm) proportionally to the initial spatial distribution of biomass.

Fleet

Each boat in the hairtail fishery is a separate agent. The agent chooses one cell to fish for each trip they take through the use of the explore-exploit-imitate algorithm (Carrella et al. 2019).

Each boat faces the same variable and fixed costs: one hourly cost for travelling, an additional hourly cost to pay for each hour spent fishing and a yearly fixed cost to simulate cooperative memberships, repairs and depreciation. Boats also have a maximum hold size (described in weight rather than volume) which limits the amount of hairtail they can land in a single trip. Each trip lasts no more than 9 hours of active fishing and never more than 2 days total.

Boats who make profits throughout the year can save it and if they accumulate enough savings above a threshold they consume it to generate an additional boats.

Active boats will choose each day whether to go out fishing with a random draw (its probability defined by rejection sampling). Boats who made consistent losses and have negative "savings" can first reduce their total effort such that they only fish four months a year but if savings continue to be negative they quit the fishery entirely. The amount of years of negative savings agents are willing to tolerate is given by the inertia parameter in the rejection sampling table.

Boats fishing in a cell transform local biomass into catch through a logistic selectivity curve (logistic with respect to mid-length of each bin). Catch sells for different prices depending on its length, with fish below 25cm being unsellable (assuming therefore a size limit is in place).