Appendix A: Code and supplementary figures for simulation and policy analyis

2018-01-30

Contents

library(parallel)
library(gridExtra)

options(mc.cores=parallel::detectCores())

Definitions: Population Model and Utility function		
Calculating MSY, TAC, and CE	2	
Solving the POMDP problem	4	
Introduce a discrete grid	4	
Define the POMDP matrices	5	
Compute the POMDP solution	5	
Simulations	6	
Simulating the static policies under uncertainty	6	
Simulating the POMDP policies under uncertainty	7	
Figure S2: Full simulation plots	7	
Figure S3: Economic value	8	
POMDP simulations overestimating measurement uncertainty	9	
Policy plots	11	
Static policies	11	
	11	
Comparing POMDP Policies	13	
Evolution of the POMDP prior belief	17	
References	19	
The POMDP calculations presented here are performed using the sarsop R package, which we have written to provide a general purpose implementation of the SARSOP algorithm of Kurniawati, H and Lee (2008). The version used in this analysis can be installed from GitHub using the commandevtools::install_github("boettiger-lab/sarsop@0.5.0").	lsu,	
<pre>#devtools::install_github("boettiger-lab/sarsop@0.5.0") library(sarsop) # Our main POMDP package library(tidyverse) # for munging and plotting</pre>		

Definitions: Population Model and Utility function

```
r <- 0.75
K <- 1

## Unlike classic Graham-Schaefer, this assumes harvest</pre>
```

log_dir <- "../data/observe_harvest_recruit" # Store the computed solution files here</pre>

```
## occurs right after assessment, before recruitment
f <- function(x, h){
    s <- pmax(x-h, 0)
    s + s * r * (1 - s / K)
}</pre>
```

A simply utility (reward) function: a fixed price per fish with no cost to fishing. Alternate models can easily be considered, we focus on this simple utility throughout as this is the choice typically focused on in the optimal control literature, implicitly assumed by the basic MSY model, and does not change the general results. Note that setting a smaller discount will take longer to converge to smooth POMDP solution, but result in S_star closer to the simple B_MSY.

```
reward_fn <- function(x,h) pmin(x,h)
discount <- 0.95</pre>
```

Calculating MSY, TAC, and CE

The approach below uses a generic optimization routine to find the stock size at which the maximum growth rate is achieved. (Note that this also depends on the discount rate since future profits are worth proportionally less than current profits).

Since we observe -> harvest -> recruit, we would observe the stock at its pre-harvest size, $X_t \sim B_{MSY} + H_{MSY}$.

```
\#B\_MSY \leftarrow S\_star # recruit first, as in classic Graham-Schaefer

B\_MSY \leftarrow f(S\_star,0) # harvest first, we observe the population at B\_MSY + h

\#MSY \leftarrow f(B\_MSY,0) - B\_MSY # recruit first

MSY \leftarrow B\_MSY - S\_star # harvest first

F\_MSY \leftarrow MSY / B\_MSY

F\_TAC = 0.8 * F\_MSY
```

As a basic reference point, simulate these three policies in a purely deterministic world. Unlike later simulations, here we consider all states an actions exactly (that is, within floating point precision). Later, states and actions are limited to a discrete set, so solutions can depend on resolution and extent of that discretization.

```
MSY_policy <- function(x) F_MSY * x
TAC_policy <- function(x) F_TAC * x

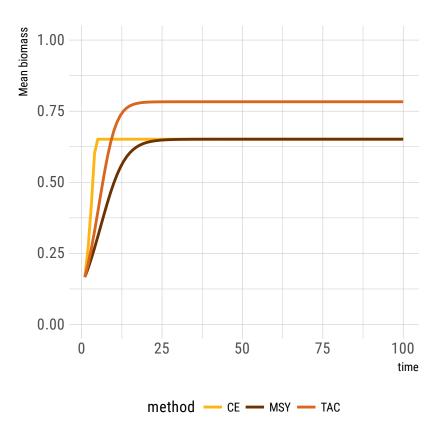
## ASSUMES harvest takes place *before* recruitment, f(x-h), not after.
escapement_policy <- function(x) pmax(x - S_star,0)

x0 <- K/6
Tmax = 100
do_det_sim <- function(policy, f, x0, Tmax){
    action <- state <- obs <- as.numeric(rep(NA,Tmax))
    state[1] <- x0

for(t in 1:(Tmax-1)){</pre>
```

With no stochasticity, MSY leads to the same long-term stock size as under the constant escapement rule, but takes longer to get there. This level is essentially B_{MSY} , though because the model considered here implements events in the order: observe, harvest, recruit; rather than observe, recruit, harvest, we see the stock at the pre-harvest size of $B_{MSY} + H_{MSY}$ (that is, K/2 + rK/4). More conservative rules, such as a harvest set to 80% of MSY result in faster recovery of the stock than under MSY, but slower than under constant escapement. Due to the reduced maximum harvest rate, such rules lead to stock returning to a value higher than B_{MSY} .

```
det_sims %>%
  ggplot(aes(time, state, col=method)) +
  geom_line(lwd=1) +
  coord_cartesian(ylim = c(0, 1)) +
  theme(legend.position = "bottom") +
  ylab("Mean biomass")
```



Solving the POMDP problem

Introduce a discrete grid

```
## Discretize space
states <- seq(0,2, length=100)
actions <- states
observations <- states</pre>
```

We compute the above policies on this grid for later comparison.

```
index <- function(x, grid) map_int(x, ~ which.min(abs(.x - grid)))

policies <- data.frame(
# CE = index(pmax(f(states,0) - S_star,0), actions), # obs,recruit,harv, f(x_t) - h_t
    CE = index(escapement_policy(states), actions), # obs,harv,recruit, f(x_t - h_t)
    MSY = index(MSY_policy(states), actions),
    TAC = index(TAC_policy(states), actions))</pre>
```

Define the POMDP matrices

We compute POMDP matrices for a range of sigma_g and sigma_m values. (We consider sigma_g = 0.02 rather than precisely zero here to avoid a degeneracy in the discretized transition matrix. Calculations for the precise deterministic solution can be solved without discretization as shown above, and match the results of this nearly-zero noise level in the discrete simulation, as expected.)

sigma_g	$sigma_m$	scenario
0.02	0.00	1
0.10	0.00	2
0.15	0.00	3
0.02	0.10	4
0.10	0.10	5
0.15	0.10	6
0.02	0.15	7
0.10	0.15	8
0.15	0.15	9

Compute the POMDP solution

The POMDP solution is represented by a collection of alpha-vectors and values, returned in a *.policyx file. Each scenario (parameter combination of sigma_g, sigma_m, and so forth) results in a separate solution file

Because this solution is computationally somewhat intensive, be sure to have ~ 4 GB RAM per core if running the 9 models in parallel. Alternately, readers can skip the evaluation of this code chunk and read the cached solution from the policyx file using the *_from_log functions that follow:

```
dir.create(log_dir)
system.time(
  alphas <-
    parallel::mclapply(1:length(models),
    function(i){</pre>
```

```
log_data <- data.frame(model = "gs",</pre>
                          r = r,
                          K = K
                          sigma_g = meta[i, "sigma_g"][[1]],
                          sigma_m = meta[i,"sigma_m"][[1]],
                          noise = "normal",
                          scenario = meta[i, "scenario"][[1]])
  sarsop(models[[i]]$transition,
         models[[i]] $observation,
         models[[i]]$reward,
         discount = discount,
         precision = 0.00000002,
         timeout = 15000,
         log_dir = log_dir,
         log_data = log_data)
})
```

We can read the stored solution from the log:

```
meta <- meta_from_log(data.frame(model="gs", discount=discount), log_dir) %>%
    mutate(scenario = as.character(scenario)) %>%
    left_join(meta) %>%
    arrange(scenario)
alphas <- alphas_from_log(meta, log_dir)</pre>
```

Simulations

Having calculated the POMDP solution in terms of these alpha vectors, we can easily compute the POMDP policy for any given prior and update the prior given further observations according to Bayes rule. We assume a uniform prior belief at the start of the POMDP simulation. These tasks are performed by the sim_pomdp function. Simulating the static strategies (CE, MSY, TAC) is easier, since the policy is uniquely determined by the observed state; we simply need to simulate the stochastic growth and process and measurement with error at each time step an then compute the corresponding policy. The sim_pomdp function performs such static simulations when given a pre-specified policy instead of a set of alpha vectors.

Simulating the static policies under uncertainty

```
map_dfr(policies, do_sim, .id = "method")
}, .id = "scenario")
```

Simulating the POMDP policies under uncertainty

We then combine the resulting data frames, transition (mutate) the units from grid indices (state, action $\in 1...N$) to continuous values (states, actions $\in [0,2]$), and select necessary columns before writing the data out to a file.

```
sims <- bind_rows(static_sims, pomdp_sims) %>%
   left_join(meta) %>% ## include scenario information (sigmas; etc)
   mutate(state = states[state], action = actions[action]) %>%
   select(time, state, rep, method, sigma_m, sigma_g, value)

write_csv(sims, file.path(log_dir, "sims.csv"))
```

Figure S2: Full simulation plots

We the results varying over different noise intensities, sigma_g, and sigma_m. Note that Figure 1 of the main text shows only the case of sigma_g = 0.15, sigma_m = 0.1 and omits MSY to simplify the presentation.

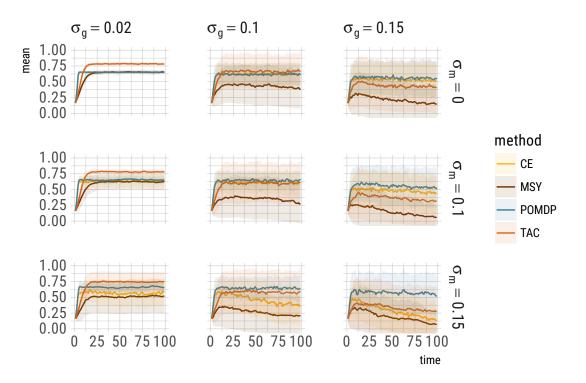
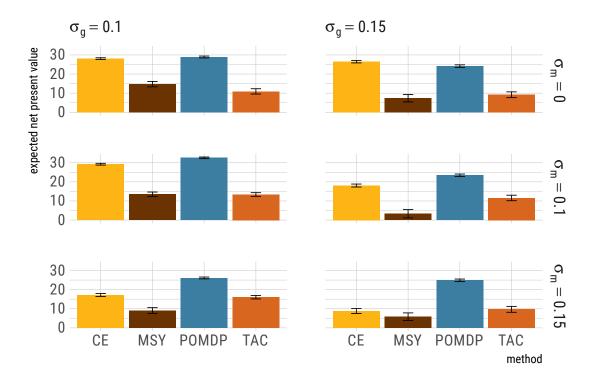


Figure S3: Economic value

Economic value is shown under varying noise levels under each strategy.

```
sims %>%
  select(time, value, rep, method, sigma_m, sigma_g) %>%
  filter(sigma_g %in% c("0.1", "0.15")) %>%
  group_by(rep, method, sigma_m, sigma_g) %>%
  summarise(npv = sum(value)) %>%
  group_by(method, sigma_m, sigma_g) %>%
  summarise(net_value = mean(npv), se = sd(npv) / mean(npv)) %>%
  ungroup() %>%
  ggplot(aes(method, net_value, ymin=net_value-se, ymax=net_value+se, fill=method)) +
  geom_bar(position=position_dodge(), stat="identity") +
  geom_errorbar(size=.3,width=.2,
               position=position_dodge(.9)) +
  facet_grid(sigma_m~sigma_g,
             labeller = label_bquote(sigma[m] == .(sigma_m),
                                     sigma[g] == .(sigma_g))) +
  theme(legend.position = "none") +
  ylab("expected net present value")
```



POMDP simulations overestimating measurement uncertainty

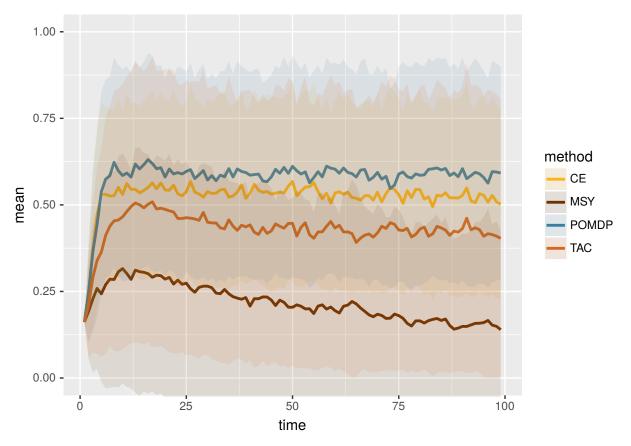
While the POMDP approach requires an estimate of the measurement error, the precise distribution of measurement errors will itself be unknown in most cases. However, the POMDP approach is quite robust to overestimation of the measurement error. As an extreme example of this, we consider the case where the POMDP solution assumes the largest measurement error level considered here, $\sigma_m = 0.15$, while performing simulations in which measurements occur without error:

```
set.seed(12345)
true <- 3 \# sigma_g = 0.15, sigma_m = 0
source(system.file("examples/pomdp_overestimates.R", package="sarsop")) # portable example code
pomdp_overest_sims <-</pre>
  map2_dfr(models, alphas, function(.x, .y){
    pomdp_overestimates(transition = .x$transition,
            model_observation = .x$observation,
            reward = .x$reward,
            discount = discount,
            true_observation = models[[true]]$observation,
            x0 = x0,
            Tmax = Tmax,
            alpha = .y,
            reps = reps)$df %>%
              mutate(method = "POMDP") # include a column labeling method
           },
           .id = "scenario"
```

)

Combine the POMDP simulations results from over-estimating measurement error with the previous results from the static policies for comparison:

```
overest <-
bind_rows(static_sims, pomdp_overest_sims) %>%
  left join(meta, by = "scenario") %% ## include scenario information (sigmas; etc)
  mutate(state = states[state], action = actions[action]) %>%
  select(time, state, rep, method, sigma_m, sigma_g, value) %>%
  group_by(time, method, sigma_m, sigma_g) %>%
  summarise(mean = mean(state), sd = sd(state)) %>%
  filter(sigma_g == "0.15") %>%
  ungroup()
overest %>% filter(method != "POMDP", sigma_m == "0") %>%
  bind rows(
  overest %>% filter(method == "POMDP", sigma_m == "0.15")) %>%
  select(-sigma_m, -sigma_g) %>%
  write_csv(file.path(log_dir, "overest_sims.csv"))
read_csv(file.path(log_dir, "overest_sims.csv")) %>%
  ggplot(aes(time, mean, col=method, fill=method)) +
  geom_line(lwd = 1) +
  geom_ribbon(aes(ymax = mean + sd, ymin = mean-sd), col = NA, alpha = 0.1) +
  coord_cartesian(ylim = c(0, 1))
```



Policy plots

Static policies

First we gather the static policies (CE, TAC, MSY) into a table and convert into terms of harvest and escapement.

```
policy_table <- tibble(state = 1:length(states)) %>%
  bind_cols(policies) %>%
  gather(policy, harvest, -state) %>%
  mutate(harvest = actions[harvest], state = states[state]) %>%
  mutate(escapement = state - harvest)
```

TAC harvests are always smaller than MSY harvests, but unlike the deterministic optimal solution, TAC and MSY solutions never go to zero. The discrete grid makes these appear slightly stepped.

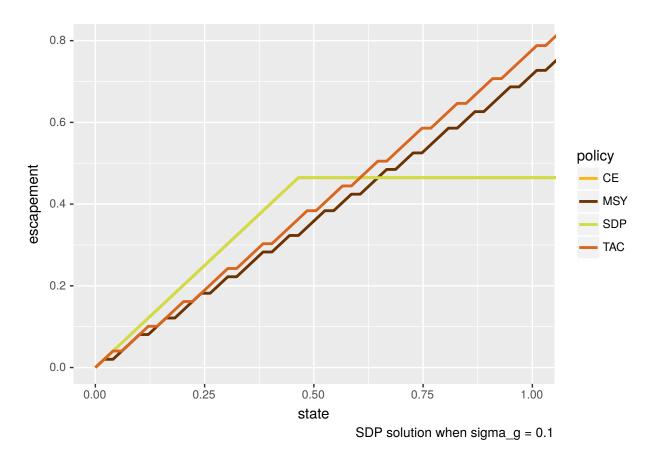
Policies under uncertainty: S = D

Because Reed (1979) has proven that the optimal escapement in the stochastic case is equivalent to the deterministic case, S=D, we have simply relied on the deterministic calculation for constant escapement. We can instead calculate the optimally policy for such a stochastic but fully observed Markov Decision Process (MDP) directly given the discritized model transition matrix using Stochastic Dynamic Programming (SDP). Reed (1979) essentially tells us that for small growth noise (satisfying or approximately satisfying Reed's self-sustaining condition) that the stochastic optimal policy is equal to the deterministic optimal policy. We can confirm this numerically as follows.

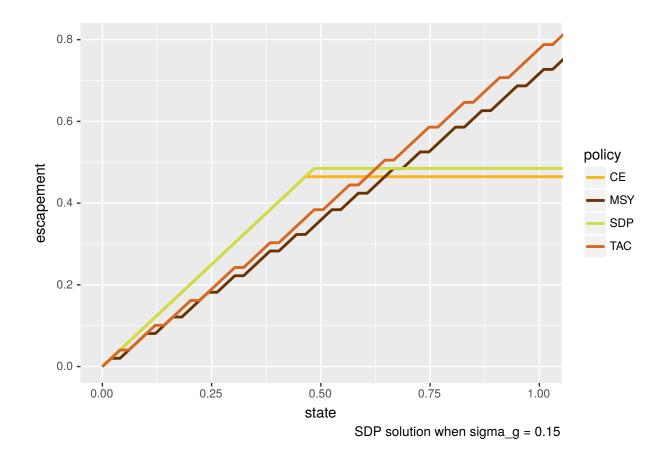
First we grab the transition matrix we have already defined when sigma_g = 0.1, and then solve the SDP using the MDPtoolbox package:

```
i <- meta %>% filter(sigma_g == 0.1, sigma_m ==0) %>%
  pull(scenario) %>% as.integer()
m <- models[[i]]
mdp <- MDPtoolbox::mdp_policy_iteration(m$transition, m$reward, discount)</pre>
```

We plot the resulting policy in comparison with the other static policies, which shows that even for such a large noise we get a stochastic constant escapement equal to that of of the deterministic CE calculation:



Repeating this larger stochastic growth $sigma_g = 0.15$, shows only slight deviation from CE (S > D), as Reed (1979) proves for sufficiently large noise to violate the self-sustaining criterion. Technically since our noise model in normally distributed none of our populations are self-sustaining for infinite time, but that deviation is quite small, as seen here. Note also the solution is still a constant-escapement type.)



Comparing POMDP Policies

The comparison of POMDP policy is yet more complicated, but the POMDP policy cannot be expressed merely in terms of a target harvest (or escapement) level given an estimation of the stock size (state). The optimal solution for the partially observed system must also reflect all prior observations of the system, not merely the most recent observation, as the system is not Markovian in the observed state variable. We summarize this history as a prior "belief" about the state, which is updated according to Bayes rule after each observation. Note that Sethi et al. (2005) plots solutions with measurement uncertainty without any reference to the prior. Bayes Law tells us that without a statement of prior belief over the true state, P(x) we cannot go from the measurement model that defines the probability of observing y given a true state x; P(y|x), into the probability of the true state given the measurement, P(x|y). Further, this prior P(x) should be updated with respect to each subsequent observation, we cannot simply assume it is a fixed value as Sethi et al. (2005). This explains their counter-intuitive finding that increased uncertainty should result in increased harvest rates with increased uncertainty.

Let us look at the POMDP solutions under various priors focusing on the case of moderate uncertainty, $\sigma_g = 0.15$ and $\sigma_m = 0.1$. (Recall we have already solved the POMDP solution for this model in the simulations above, as defined by the alpha vectors, so we can quickly load that solution now.)

```
i <- meta %>%
  filter(sigma_g == 0.15, sigma_m == 0.1) %>%
  pull(scenario) %>% as.integer()
m <- models[[i]]
alpha <- alphas[[i]] # we need the corresponding alpha vectors</pre>
```

We will consider what the POMDP solution looks like under a few different prior beliefs. A uniform prior

sounds like a conservative assumption, but it is not: it puts significantly more weight on improbably large stock values than other priors. (Loading the α vectors from our POMDP solution computed earlier, we can then compute a POMDP given these α , the matrices for transition, observation, and reward, and the prior we are using)

```
unif_prior = rep(1, length(states)) / length(states) # initial belief
unif <- compute_policy(alpha, m$transition, m$observation, m$reward, unif_prior)</pre>
```

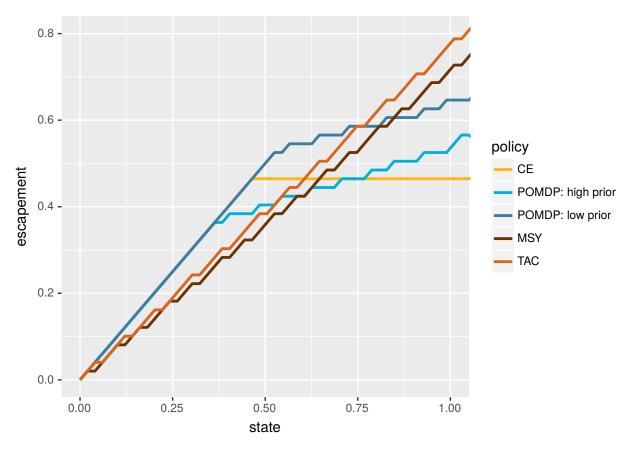
For a representative set of priors, we will consider priors centered at the target B_{MSY} size (or S^* in the language of Reed), at half B_{MSY} (i.e. K/4), and at $\frac{3}{4}K$, each with a standard deviation of $\sigma_m = 0.1$ (i.e. the uncertainty around a single observation of a stock at that size.)

```
prior_star <- m$observation[, which.min(abs(states - S_star)),1]
prior_low <- m$observation[,which.min(abs(states - 0.25 * K)),1]
prior_high <- m$observation[,which.min(abs(states - 0.75 * K)),1]

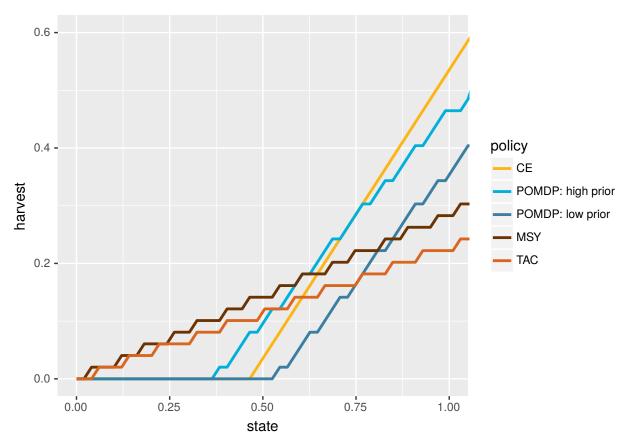
star <- compute_policy(alpha, m$transition, m$observation, m$reward, prior_star)
low <- compute_policy(alpha, m$transition, m$observation, m$reward, prior_low)
high <- compute_policy(alpha, m$transition, m$observation, m$reward, prior_high)</pre>
```

We gather these solutions into a single data frame and convert from grid indices to continuous values

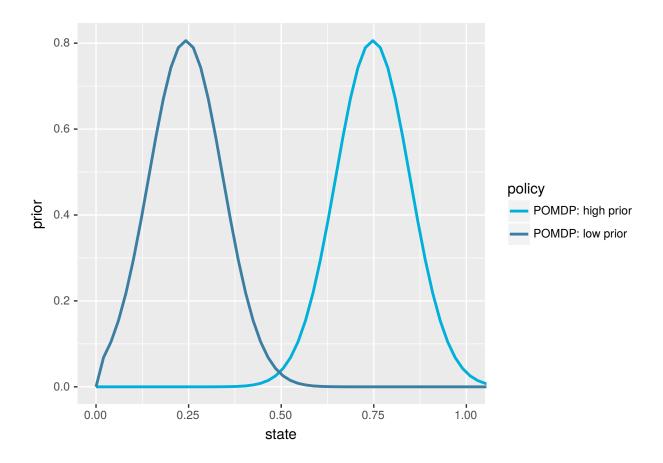
Plot these policies in terms of both Harvest and Escapement:



```
p2 <- recode_policies %>%
    ggplot(aes(state, harvest, col = policy)) +
    geom_line(lwd=1) +
    coord_cartesian(xlim = c(0,1), ylim=c(0,.6))
p2
```

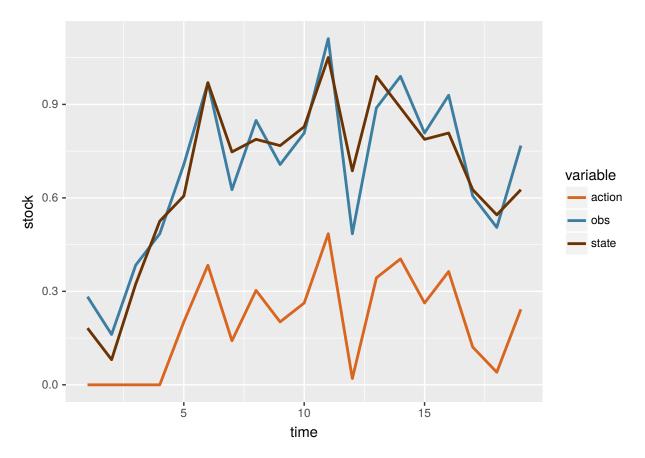


```
p3 <- read_csv(file.path(log_dir, "priors.csv")) %>%
ggplot(aes(state, prior, col = policy)) +
   geom_line(lwd=1) +
   coord_cartesian(xlim = c(0,1))
p3
```



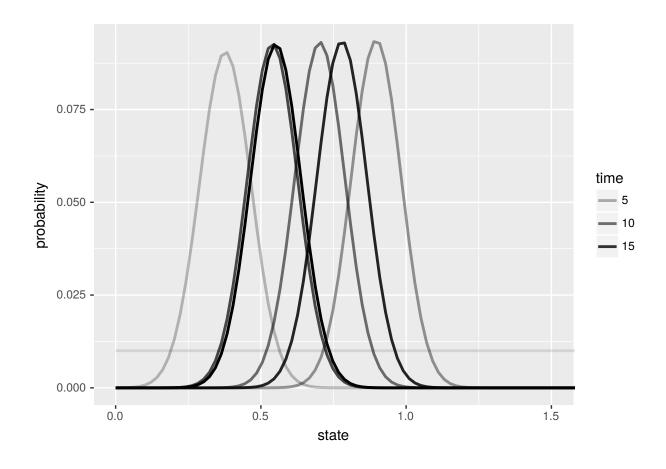
Evolution of the POMDP prior belief

To get a sense of how the prior belief (and consequently the POMDP policy) is continually adapting over the course of a simulation (unlike the static policies), it is instructive to examine a single simulation rather than the ensemble of replicates considered above.



Here we plot the evolution of the prior belief in state over the course of the above simulation in increasingly darker curves. Note that the beginning belief is uniform, but quickly tightens into a distribution centered around the region of the most recent observation. Note that this belief distribution never converges: as the true state constantly changes in due to stochasticity in population recruitment, so too does the belief distribution constantly change in response to new observations of that true state.

```
sim$state_posterior %>%
  data.frame(time = 1:Tmax) %>%
  filter(time %in% seq(1,Tmax, by = 3)) %>%
  gather(state, probability, -time, factor_key =TRUE) %>%
  mutate(state = states[as.numeric(state)]) %>%
  ggplot(aes(state, probability, group = time, alpha = time)) +
  geom_line(lwd = 1) +
  coord_cartesian(xlim = c(0,1.5))
```



References

Kurniawati, Hanna, David Hsu, and Wee Sun Lee. 2008. "SARSOP: Efficient Point-Based POMDP Planning by Approximating Optimally Reachable Belief Spaces." Proceedings of Robotics: Science and Systems IV, w/o page numbers. https://www1.comp.nus.edu.sg/ $\{\sim\}$ leews/publications/rss08.pdf $\$ \backslash $\$ hhttp://www.roboticsp:

Reed, William J. 1979. "Optimal escapement levels in stochastic and deterministic harvesting models." *Journal of Environmental Economics and Management* 6 (4). Elsevier: 350–63. doi:10.1016/0095-0696(79)90014-7.

Sethi, Gautam, Christopher Costello, Anthony Fisher, Michael Hanemann, and Larry Karp. 2005. "Fishery management under multiple uncertainty." *Journal of Environmental Economics and Management* 50 (2): 300–318. doi:10.1016/j.jeem.2004.11.005.