PNAS
www.pnas.org

Supplementary Information for

Rebuilding global fisheries under uncertainty

Milad Memarzadeh, Gregory Britten, Boris Worm, Carl Boettiger

Corresponding Authors Milad Memarzadeh, Carl Boettiger
E-mail: memarzadeh.milad@gmail.com or cboettig@berkeley.edu

This PDF file includes:

Supplementary text
Figs. S1 to S4
Tables S1 to S3
References for SI reference citations

www.pnas.org/cgi/doi/10.1073/pnas.1902657116

Supporting Information Text

Data Source. Data for the analysis in this paper are obtained from the publicly available database of the RAM Legacy Stock Assessment Database (1). We defined well-sampled stocks as those that the length of data points (measurements of population biomass and recordings of the actual catch) were sufficient for model estimation. We have eliminated stocks that had less than 30 data points, as it is too small for estimation of the models to be accurate enough for analysis. It contains information from 109 well-sampled stock assessments worldwide (including Australia, New Zealand, Canada, the USA, Peru, South Africa, Argentina, Russia, Japan, Europe, and numerous multi-national stocks) and provides the time series of stock biomass and catch for all of them. The complete list of the RAM fisheries included in the analyses in this paper is presented in Table S1.

Some of the timeseries in the RAM Legacy data prior to 1970 include long sections that are perfectly flat or smooth trends which only start fluctuating more erratically after that time. This pattern is consistent with hindcast projections from stock assessment models, where the smooth regions represent pure model output without observations from that period. Such regions will lead us to underestimate the level of uncertainty in those stocks. Unfortunately, the RAM Legacy data provides no independent annotation to indicate what timeseries come only from such projections. While it would be possible to define a threshold to filter out any such timeseries that are too smooth, this would remove timeseries with little apparent uncertainty where existing approaches like MSY or MDP will perform best. To avoid drawing an arbitrary filter on the data that would favor the POMDP, we have simply included all of the time series.

Materials and Methods

In this section, we explain different parts of obtaining the optimal sustainable yield solution. It should be noted that the decision optimization framework proposed here is not limited to the specific functions and parameters used in this study and can be easily generalized. First, we outline the model framework generally, and then specify the particular formation.

Our formulation of fisheries management closely follows the literature (2, 3), which we summarize here. The fishery manager seeks to maximize the expected net present value of the fishery over the life-span of the ecosystem. In this study we have assumed of infinite time horizon for life-span of the ecosystem; extension to finite horizon problems is straight-forward and less computationally expensive. There are two sources of uncertainty in the process: uncertainty in the functional form and model parameters defining the population dynamics, and uncertainty in estimating the population biomass. The goal is to find the optimal policy that satisfies the above-mentioned objective:

$$\Pi_\tau^* = \text{argmax}_{q_t \in Q} \mathbb{E}_\Theta \left[\sum_{t=\tau}^{\infty} \gamma^t \sum_{b_t \in B} p(b_t) U(b_t, q_t) \right] \qquad [1]$$

where, $b \in B$ denotes the stock biomass, $q \in Q$ is the harvested biomass (the catch quota), with subscripts t denoting years, Π_τ^* denotes the optimal harvest policy at year τ, $\mathbb{E}$ is the expectation, Θ represent the uncertain population dynamics model (either the functional form, i.e. functional uncertainty, or parameter uncertainty), γ represents the discount factor, translating future economic returns to their net present value, U is the utility function, $p(b_t)$ is the probability distribution over the stock biomass (because of measurement error, the true stock biomass cannot be identified with certainty). While the set of possible biomass values B is the non-negative real numbers, for numerical optimization we replace this with values on a discrete, uniform grid of 100 points in $B \equiv [0, 2K]$, where K is the estimated carrying capacity of the stock. We consider all possible harvests, $Q \equiv B$.

Model formulation. The state of the system is b_t, which corresponds to the (true) fish stock biomass at time t which a manager only knows imperfectly through the distribution $p(b_t)$. The harvest policy is $h_t = \min(b_t, q_t)$, where q_t is the attempted catch at time t. For simplicity, we assume the utility (i.e. reward) is equal to the harvest, i.e. $U(b_t, q_t) = \min(b_t, q_t)$, with future rewards discounted by multiplicative discount factor γ. More complex utility functions are equally applicable.

The population dynamics is formulated according to foundational Gordon-Schaefer surplus production model (4).

$$b_{t+1} = \zeta_t \left(b_t + r b_t \left(1 - \frac{b_t}{K} \right) - q_t \right) \qquad [2]$$

Where, $b \in B$ and $q \in Q$ as defined above, with subscripts t denoting years, r the intrinsic productivity of the stock, K the carrying capacity and ζ_t is a random variable that captures the inherent stochasticity in population growth and follows a Gaussian distribution with mean unity and standard deviation σ_g (i.e. the process error, assumed to arise due to environmental variation). It should be noted that this specific model is being chosen for the results to be comparable to recent research in this field that have used the same model (2, 3), although the proposed method can be applied to any other more complex model (e.g. age-structure models (5)).

Due to measurement error, the observations of stock biomass are assumed to be noisy, $z_t = \epsilon^t b_t$, where the estimated biomass is defined as a random Gaussian noise $\epsilon^t \sim N(b_t, \sigma_m)$. Four alternative measurement errors have been evaluated in this article, $\sigma_m \in \{0, 0.05, 0.1, 0.2\}$. Both true and estimated biomass are normalized such that $B \in [0, 1]$, in which any probability density falling outside that region is placed on the corresponding boundary. This effectively truncates the noise distributions, though in practice these probabilities are small.

We choose Gaussian random noise to correspond to the stochastic differential equation for these population dynamics under environmental noise, see (6), but alternative noise formulations are possible for discrete-time dynamics, such as the log-normal.

The log-normal distribution is long-tailed, resulting in more frequent large deviations for the same variance that can exaggerate the role of uncertainty. (2) considered uniform noise on a bounded interval, which simplifies the numerics at the expense of realism. The qualitative performance of the different decision methods under different noise structure is consistent with the results presented here. Examination of alternative noise structures is supported by our software package, (7), and explored in (8, 9).

Model estimation. As mentioned above, the parameters of the population dynamics model are treated uncertain and need to be estimated from data. To do so, we use Markov chain Monte Carlo (MCMC) to sample from the posterior distribution of parameters given the observations (10) through NIMBLE (11) to estimate parameters r, K and σ_g. The Bayesian model is as follows:

$$p(b_{t+1} \mid b_t, r, K, q_t) \sim f(b_t, r, K, q_t) N(1, \sigma_g) \quad [3]$$

$$p(r) \sim \text{Uniform}(0, r_{max}) \quad [4]$$

$$p(K) \sim \text{Uniform}(0, K_{max}) \quad [5]$$

$$p(\sigma_b^t) \sim \text{Uniform}(0, \sigma_{max}) \quad [6]$$

Where f is the Gordon-Shaefer model from Eq. (2). Once the model parameters are estimated, the decision optimization phase proceeds in a two-step process, with the second step explained below. The estimated model parameters for all stocks is reported in Table S2. r_{max},K_{max},σ_{max} are equal to 2 (avoiding period-doubling/chaos), 1 (since biomass is normalized), and 0.5 (avoiding very large noise) respectively. Note that the model estimation does not attempt to independently estimate the observation error directly from the RAM stock assessments, as a reliable estimate would require significantly more data than are available. In place of such an estimate we evaluate all stocks using a range of four alternative measurement errors $\sigma_m \in \{0, 0.05, 0.1, 0.2\}$ each.

Decision optimization. We model the decision optimization step based on the framework of partially observable Markov decision processes (POMDP) (12, 13). POMDP is a generalization of Markov decision processes (MDP), which is a fundamental model for sequential decision making and optimal control (14, 15). MDP is known as stochastic dynamic programming (SDP) in the ecological literature (16, 17). Specifically, POMDP is a suitable framework for decision optimization in this context, as they allow the partial observability of the state space, which is the case here, when we have error in measuring the stock biomass.

Generally, POMDPs are defined as a tuple $(B, q, Z, T, O, U, \gamma, p(b_0))$, in which B represent the state of the system (corresponding to population biomass), q is available actions to the decision-maker (i.e. the agent, corresponding to catch quota), Z is the noisy measurements of the state, T represents the transition probability function, which defines dynamics of the system and in this manuscript is modeled by Gordon-Schaefer surplus production model (Eq. 2), O defines the accuracy of measurements (i.e. observation error, which in the manuscript is defined with a Gaussian distribution), U is the utility function as defined before, γ is the discount factor, and $p(b_0)$ is the initial belief state.

The belief state at each time t is the posterior probability of the state space, given history of observations formally described as $p(b_t) = p(b_t \mid \bar{q}_{t-1}, \bar{z}_t)$, where $\bar{z}_t = \{z_1, \ldots, z_t\}$ indicates the history of observations up to time t. Belief at time $t+1$ can be updated at each time given the new observation z_{t+1}, using Bayes' rule:

$$p(b_{t+1}) = \frac{O(b_{t+1}, q_t, z_{t+1}) \sum_{b \in B} T(b, q, b_{t+1}) p(b_t)}{\sum_{b'' \in B} O(b'', q_t, z_{t+1}) \sum_{b \in B} T(b, q_t, b'') p(b_t)} \quad [7]$$

In the context of POMDP, a policy $\pi : p(b) \rightarrow A$ is a mapping from belief state to actions. Basically, with annual measurement of the population biomass, the policy should characterize the harvest quota for that year. Optimal policy can be found by solving the recursive Bellman's equation (18):

$$V^*(p(b)) = \max_q \left[\sum_{b \in B} p(b) U(b, q) + \gamma \sum_{z \in Z} p\left(z \mid p(b), q\right) V^* \left(p(b')\right) \right] \quad [8]$$

Where, $V^*(p(b))$ is the optimal long-term utility of managing the fishery over its entire life-span with starting stock biomass measurement at $p(b)$, $p(b')$ is the updated belief according to Eq. (7), with $q = q_t$,$z = z_{t+1}$, $p(b_t)$, and conditional probability $p(z \mid p(b), q)$ is computed as:

$$p\left(z \mid p(b), q\right) = \sum_{b' \in B} O\left(b', q, z\right) \sum_{b \in B} T\left(b, q, b'\right) p(b) \quad [9]$$

Finding the optimal policy for management of the system (i.e. fishery) over its entire life-span are theoretically possible through dynamic programming, but in practice are computationally expensive with any more than a handful of states (due to the fact known as the curse of dimensionality) (14). We approximate the optimal solution using an advanced dynamic programming algorithm called point-based value iteration that is applicable to complex real-world scenarios (19, 20). We summarize this method in the next section, and we refer the interested reader to (21) for more details. We have developed a software for the decision optimization step in R that is publicly available at https://github.com/boettiger-lab/sarsop. In results presented in this article, the computational time of identifying the optimal management strategies for each stock (solving Eq. (1)) takes about [2.6 - 3.3] hours to reach a precision in the order of [0 - 0.14] for normalized stock biomass (where the domain of the population biomass is $B = [0, 1]$).

Point-based value iteration. As mentioned above, finding the optimal policy for management of the system in the POMDP setting using dynamic programming (i.e. classical value iteration algorithm) is computationally challenging (22). Point-based value iteration approaches speed up the optimization process by restricting the search in the belief space to only those that can be reached starting from the initial belief state, i.e. $p(b_0)$ (19). In particular, here we use one of the most effective point-based value iteration methods, successive approximations of the reachable space under optimal policies (SARSOP) (Kurniawati et al. 2008), which identifies the optimally reachable belief states and approximates the optimal value function using this set. SARSOP represents the state-of-the-art in solving POMDPs offline in terms of efficiency and accuracy. The main idea behind point-based value iteration algorithms are as follows: it is well known that (for any POMDP) the convex value function (as defined in Eq. (8)) can be approximated, at any time, by the envelope of a set of affine functions (12), on the belief's domain,

$$V_t^*(p(b)) \geq \max_{\alpha \in \Gamma_t} [\alpha^{\mathrm{T}} p(b)] \qquad [10]$$

where Γ_t is the set of so-called *alpha vectors* at time t, and T is vector transpose. Each alpha vector is of size $[B \times 1]$ and refers to a specific conditional policy (23). At each time step, t, conditional policy ϕ_t assigns current management action (i.e. catch) $q_t = q[\phi_t]$ and an action at each future time depending on the sequence of collected observations. Depending on the observation $z_{t+1} = z$ at the next time step, the conditional policy continues into a new one, $\phi_{t+1}(z)$. Hence, conditional policy ϕ_t can be described by the initial action $q[\phi_t]$ and the set of conditional policies $\phi_{t+1}(z)$, for each possible observation z. Consequently, from n_{t+1} number of conditional policies at time step $t+1$, $n_t = n_{t+1}^Z Q$ distinct possible conditional policies can be defined at time step t. However, most of these conditional policies can be pruned as they are dominated by other ones. If Γ_t contained all possible alpha vectors (and not only those reachable from the initial belief), then Eq. (10) can be written with an equal sign.

To build an alpha vector from the corresponding conditional policy, one can follow the following formula,

$$\alpha_{b,\phi_t} = U(b, q[\phi_t]) + \gamma \sum_{z \in Z} O(b, q[\phi_t], z) \sum_{b' \in B} T(b, q[\phi_t], b') \alpha_{b',\phi_{t+1}(z)} \qquad [11]$$

where $\alpha_{b',\phi_{t+1}(z)}$ is component b' of the alpha vector related to the conditional policy $\phi_{t+1}(z)$.

Alpha vectors for each time can be built by initializing a set of them at the end of the time horizon, and using Eq. (11) as a Bellman operator. However, as noted previously, the number of vectors grows exponentially and, after few steps backward, the complete set becomes intractable. A point-based value iteration method, such as those for solving POMDPs, can approximate the value function, in the set of beliefs reachable from the initial belief, with a limited number of relevant alpha vectors. Most of the modern POMDP solvers are different in the way they update alpha vectors (such as one provided in Eq. (11)) and heuristics they use for pruning dominated alpha vectors. For further details regarding different approaches to do these steps, refer to (24).

Discount rates. We choose a small but non-zero discount rate of 1% (typical estimates would be between 1% and 5%). Ignoring the discount rate could bias the results, particularly in regard to the economic claims of the paper, as it would suggest that profits infinitely far into the future are worth the same as profits made today. Therefore, some discount rate is appropriate. Differences in discount rate have little effect on the qualitative behavior shown here. Figure S4 compares POMDP solutions over a suite of 500 replicate simulations under two different discount rates, 1% and 4%. While very large discount rates can considerably alter the value of rebuilding stocks at all, the results are not sensitive within this typical band of of discounting, with stocks rebuilding comparably under POMDP with either discount rates. Simulations use $K = 0.75$, $r = 0.3$, $\sigma_g = 0.1$, $\sigma_m = 0.1$ with all replicates starting below B_{MSY}, at $b_0 = K/4$.

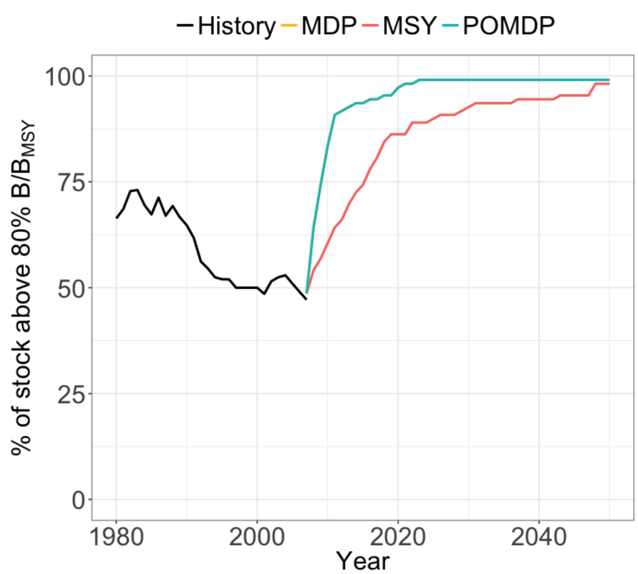

Fig. S1. Deterministic projections of recovery of the stocks under different management regimes, *black*: Historical observations, *green*: POMDP, *yellow*: MDP, *red*: MSY. It should be noted that in the deterministic case, POMDP and MDP will perform identical and recover the population faster than MSY as expected (MDP is theoretical optimum).

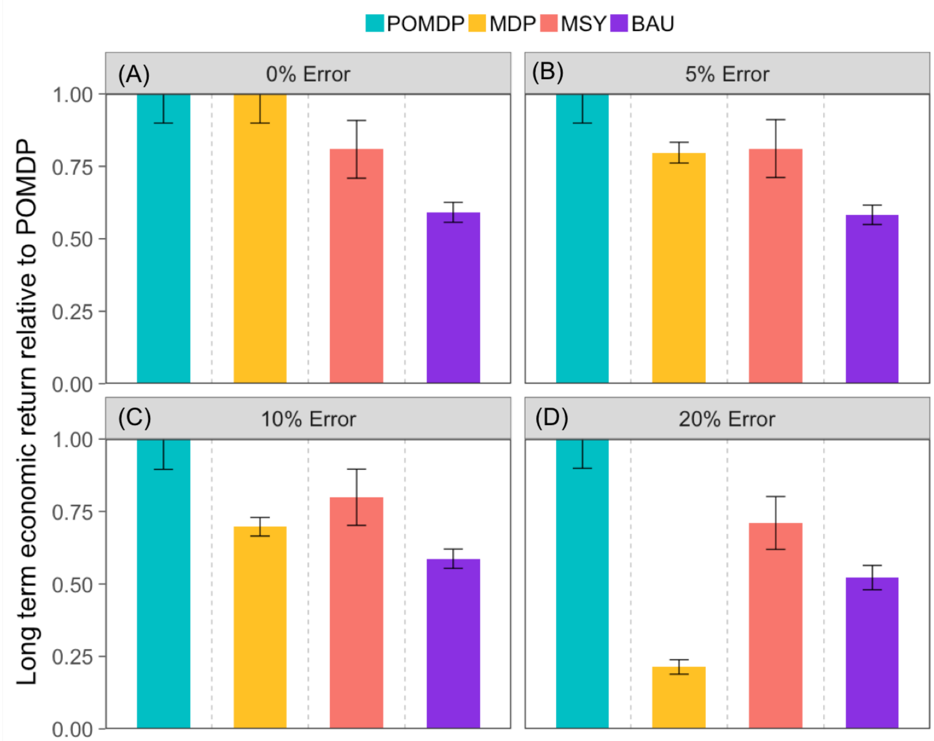

Fig. S2. Economic returns. This figure shows the long term economic return of different management regimes relative to POMDP, *green*: POMDP, *yellow*: MDP, *red*: MSY, and *purple*: BAU. This value is calculated as the expected sum of discounted immediate rewards of managing fisheries under each management and is shown in relative units to avoid additional uncertainty in predicting market values in real dollar terms. Different panels correspond to the intensity of the present measurement error in estimating the population biomass, i.e. (a) 0%, (b) 5%, (c) 10%, and (d) 20%. Bars show average trends over all stocks listed in Table S1 and hangers show +/- standard deviation.

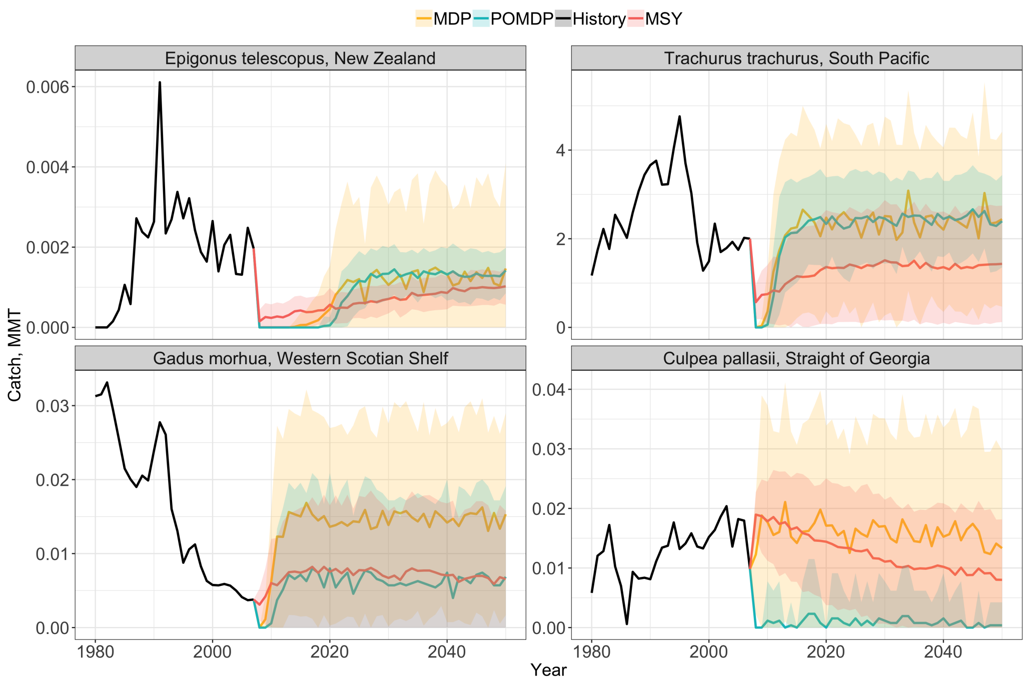

Fig. S3. Example projections of individual fish management actions. This figure shows four example management actions for (a) Black cardinalfish (*Epigonus telescopus*), (b) Horse mackerel (*Trachurus trachurus*), (c) Atlantic cod (*Gadus morhua*) and (d) Pacific herring (*Culpea pallasii*) based on different decision methods. Projections show averages over 500 replicate simulations at 10% measurement error. Lines show average trends over replicates and shades show +/- standard deviation.

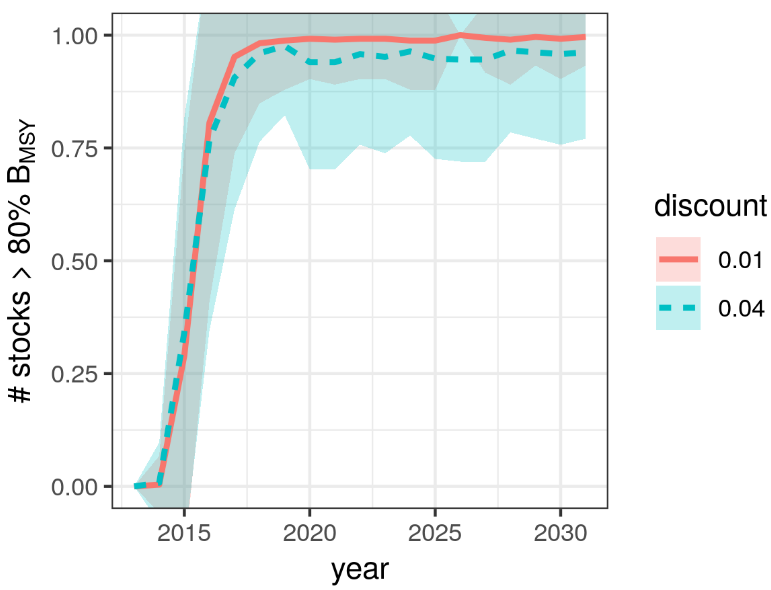

Fig. S4. Effect of discount rate on the POMDP solution. This figure shows stock rebuilding over 500 replicate simulations of a POMDP-managed fishery using either a 1% or 4% discount rate.

Table S1. Fisheries obtained from the RAM Legacy Stock Assessment Database for analyses in this paper.

file included separately

Table S2. Estimated parameters for all fisheries included in the study.

file included separately

Table S3. List of actual geographical areas of stocks used in this study and the area codes assigned to them.

file included separately

References

1. Ricard D, Minto C, Jensen O, Baum J (2011) Examining the knowledge base and status of commercially exploited marine species with the RAM Legacy Stock Assessment Database. *Fish and Fisheries* 13:380–398.
2. Sethi G, Costello C, Fisher A, Hanemann M, Karp L (2005) Fishery management under multiple uncertainty. *Journal of Environmental Economics and Management* 50(2):300–318.
3. Britten G, Dowd M, Kanary L, Worm B (2017) Extended fisheries recovery timelines in a changing environment. *Nature Communications* 8.
4. Schaefer MB (1957) Some Considerations of Population Dynamics and Economics in Relation to the Management of the Commercial Marine Fisheries. *Journal of the Fisheries Research Board of Canada* 14(5):669–681.
5. Holden MH, Conrad JM (2015) Mathematical Biosciences Optimal escapement in stage-structured fisheries with environmental stochasticity. *Mathematical Biosciences* 269:76–85.
6. Boettiger C (2018) From noise to knowledge: how randomness generates novel phenomena and reveals information. *Ecology Letters* 21(8):1255–1267.
7. Boettiger C, Ooms J, Memarzadeh M (year?) *sarsop: Approximate POMDP Planning Software.* R package version 0.4.0.
8. Memarzadeh M, Boettiger C (2019) Resolving the measurement uncertainty paradox in ecological management. *The American Naturalist* 193(5):645–660.
9. Memarzadeh M, Boettiger C (2018) Adaptive management of ecological systems under partial observability. *Biological Conservation* 224:9–15.
10. MacKay D (2003) *Information theory, inference, and learning algorithms.* (Cambridge University Press, Cambrdige, UK).
11. NIMBLE Development Team (2016) Nimble: An r package for programming with bugs models, version 0.5-1.
12. Smallwood RD, Sondik EJ (1973) The Optimal Control of Partially Observable Markov Decision Processes over a Finite Horizon.
13. Sondik EJ (1978) The Optimal Control of Partially Observable Markov Processes Over the Infinite Horizon : Discounted Costs. *Operations Research* 26(2):282–304.
14. Bertsekas D (1996) Dynamic Programming and Optimal Control. (Athena Scientific).
15. Sutton R, Barto A (1998) Reinforcement Learning: an Introduction. (The MIT Press, Cambridge, Massachusetts).
16. Marescot L, et al. (2013) Complex decisions made simple: a primer on stochastic dynamic programming. *Methods in Ecology and Evolution.*
17. Mangel M (1985) Decision and control in uncertain resource systems. p. 255.
18. Bellman R (1957) *Dynamic Programming.* (Princeton University Press, Princeton, NJ).
19. Pineau J, Gordon G, Thrun S (2003) Point-based value iteration: An anytime algorithm for POMDPs. *IJCAI International Joint Conference on Artificial Intelligence* pp. 1025–1030.
20. Kurniawati H, Hsu D, Lee WS (2008) SARSOP : Efficient Point-Based POMDP Planning by Approximating Optimally Reachable Belief Spaces. *Proceedings of Robotics: Science and Systems IV.*
21. Pozzi M, Memarzadeh M, Klima K (2017) Hidden-model processes for adaptive management under uncertain climate change. *Journal of Infrastructure Systems* 23(4):04017022.
22. Kaelbling LP, Littman ML, Cassandra AR (1998) Planning and Acting in Partially Observable Stochastic Domains. *Artificial Intelligence* 101(1-2):99–134.
23. Russell S, Norvig P (2010) *Artificial Intelligence: A Modern Approach.* (Prentice Hall).
24. Shani G, Pineau J, Roy N (2013) A survey of point-based POMDP solvers. *Auton Agents Multi-Agent Syst* 27(1):1–51.