Bayesian Inference and Decision Theory – A Coherent Framework for Decision Making in Natural Resource Management

Robert M. Dorazio and Fred A. Johnson

U.S. Geological Survey and U.S. Fish and Wildlife Service

Florida Caribbean Science Center

Gainesville, Florida

http://www.fcsc.usgs.gov/Staff/dorazio.pdf

A Common Goal in Problems of Natural Resource Management

To develop a procedure for specifying a sequence of management actions that is optimal with respect to a particular set of management objectives.

- Sometimes called "management strategy"
- Optimal actions often depend on time or state of an ecological system

EXAMPLES

- **Harvest Management**
  - fisheries management
  - wildlife management
  - forestry
- **Plant or Animal Control** (removal of undesirable species)
  - pest control
  - non-indigenous species
- **Stocking and Translocation**
  - re-establishment of metapopulations
- **Habitat Management**
  - fire management
  - water-level manipulation

Formal Solutions of Such Problems Require:

- Finite set of alternative management actions
- Objective function – values the consequences of alternative management actions
  - unambiguously defined (usually requires elicitation!!)
  - may include multiple objectives
  - may include competing objectives
- Predictive model(s) – formulate dynamics of ecological system in terms of quantities relevant to management objectives
- Monitoring program – follows evolution of the system and responses to management

Dynamic Programming

A backward-induction algorithm for computing an optimal sequence of decisions provided:

- objective is a temporal sum of decision values
- system dynamics are modeled as a Markov process
- discrete state space

Discrete state space induces a 2nd "curse of dimensionality":

- system description is limited in complexity
- posterior uncertainty in model parameters is specified using a set of discrete parameter values and information states (= model weights)
  - adequacy?
  - additional theory needed to update model weights

Decision-making is not just sequential – it's adaptive!

monitoring (collection of data relevant to management objective)

$\downarrow$

assessment (analysis of data, prediction of consequences of proposed management actions, selection of actions most likely to achieve management objectives)

$\downarrow$

implementation (actions or manipulations intended to achieve management objectives)

$\downarrow$

monitoring

$\vdots$

- There is an iterative updating of beliefs that includes learning from data and making decisions in the presence of uncertainty to achieve an overall management objective.

Bayesian Inference – a probabilistic approach for

- sequentially updating beliefs (specified in terms of model parameters) as new information is acquired through monitoring
- predicting the consequences of future management actions, while properly accounting for *posterior uncertainty* in the updated beliefs

Decision Theory – a rigorous framework in which

- management objectives are specified as a function of model predictions (and/or parameters)
- the expected consequences of any particular sequence of management actions are calculated by integrating over the *posterior uncertainty* in both model parameters and predictions

The potential applicability of the Bayesian paradigm has been recognized previously (*Ellison 1996, Bergerud and Reed 1998, Wade 2000*); however, recent advances in statistical theory and computation now allow fairly complex, and hopefully more realistic, models to be fitted and used in decision making.

Inference and Decision-Making in a Problem of Habitat Management (A Hypothetical Example)

- motivated by an actual problem
- greatly simplified; yet includes several features that are common in problems of natural resource management
- illustrates the general utility of Bayesian inference and decision theory provided by modern methods of Bayesian computation

Year 1 Observations

burn	chop	graze
graze	burn	chop
chop	burn	graze

$\longrightarrow$

y_{11}	y_{21}	y_{31}
y_{41}	y_{51}	y_{61}
y_{71}	y_{81}	y_{91}

Management Actions

$\boldsymbol{X}_1 = (\boldsymbol{x}_{11}, \ldots, \boldsymbol{x}_{91})^T$

Vegetation

Responses, $\boldsymbol{y}_1$

Year 2 Predictions

?	?	?
?	?	?
?	?	?

$\longrightarrow$

Proposed Management Actions

$\tilde{\boldsymbol{X}}_2 = (\tilde{\boldsymbol{x}}_{12}, \ldots, \tilde{\boldsymbol{x}}_{92})^T$

Possible Management Objectives

- Target levels of observables
  - 50% vegetation cover and 50% open-water habitat
  - xx% cover of particular species of plants (desirable vs. exotic)
- Minimize total cost of management actions
- Maximize learning about effects of management actions

Solution of our management problem requires:

- annual monitoring of plot-specific vegetation responses to management
- model(s) of sequence of vegetation responses in each plot
- method for updating model parameters and predictions as new data are acquired
- unambiguous statement of management objectives (defined in terms of model parameters, model predictions of observables, or costs of management actions)
- procedure for selecting an optimal sequence of future management actions based on past data

“Centered” Parameterization of Management Actions:

Plot	ManagementAction	VegetationCover
1	$\boldsymbol{x}_{1t} = (1, 0)^T \Rightarrow$ burning	y_{1t}
2	$\boldsymbol{x}_{2t} = (1, 0)^T \Rightarrow$ burning	y_{2t}
3	$\boldsymbol{x}_{3t} = (0, 1)^T \Rightarrow$ cutting	y_{3t}
$\vdots$	$\vdots$	$\vdots$
n	$\boldsymbol{x}_{nt} = (0, 1)^T \Rightarrow$ cutting	y_{nt}

Only $q = 2$ management actions are illustrated for ease of presentation.

Then, consider a first-order autoregressive model for the sequence of vegetation responses in each plot:

$$\begin{pmatrix} Y_{i1} \\ Y_{i2} \\ Y_{i3} \\ \vdots \\ Y_{i\tau} \end{pmatrix} \sim N\left(\begin{pmatrix} \boldsymbol{x}_{i1}^T\boldsymbol{\beta} \\ \boldsymbol{x}_{i2}^T\boldsymbol{\beta} \\ \boldsymbol{x}_{i3}^T\boldsymbol{\beta} \\ \vdots \\ \boldsymbol{x}_{i\tau}^T\boldsymbol{\beta} \end{pmatrix}, \quad \frac{\sigma^2}{(1-\rho^2)} \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{\tau-1} \\ \rho & 1 & \rho & & \\ \rho^2 & \rho & 1 & & \\ \vdots & & & \ddots & \\ \rho^{\tau-1} & & & & 1 \end{pmatrix} \right)$$

Each element of $\boldsymbol{\beta}$ $(q \times 1)$ corresponds to the mean vegetation cover associated with a distinct management action.

Equivalently,

$$(Y_{it} \mid \boldsymbol{x}_{it}, \boldsymbol{\beta}, \sigma^2, \rho, y_{i,t-1}, \boldsymbol{x}_{i,t-1}) \sim \begin{cases} N(\boldsymbol{x}_{it}^T\boldsymbol{\beta}, \ \sigma^2/(1-\rho^2)) & \text{if } t = 1 \\ N\left(\boldsymbol{x}_{it}^T\boldsymbol{\beta} + \rho\,(y_{i,t-1} - \boldsymbol{x}_{i,t-1}^T\boldsymbol{\beta}), \ \sigma^2\right) & \text{if } t > 1 \end{cases} \tag{1}$$

Assuming conditional independence among among plot-specific responses (i.e., no spatial dependence) yields the joint density

$$f(\boldsymbol{y}_t \mid \boldsymbol{X}_t, \boldsymbol{\theta},\, \boldsymbol{y}_{t-1}, \boldsymbol{X}_{t-1}) = \prod_{i=1}^{n} f(y_{it} \mid \boldsymbol{x}_{it}, \boldsymbol{\theta},\, y_{i,t-1}, \boldsymbol{x}_{i,t-1}),$$

where $f(y_{it} \mid \boldsymbol{x}_{it}, \boldsymbol{\theta},\, y_{i,t-1}, \boldsymbol{x}_{i,t-1})$ specifies the conditional distribution in (1) and $\boldsymbol{\theta} = (\boldsymbol{\beta}, \rho, \sigma^2)^T$.

Bayes Theorem

Provides a coherent, probabilistic framework for updating our beliefs and for quantifying our uncertainty about model parameters and model predictions as new observations are made

Inference

At $t = 1$ the posterior distribution is

$$p(\boldsymbol{\beta}, \sigma^2, \rho \mid \boldsymbol{y}_1, \boldsymbol{X}_1) = \frac{f(\boldsymbol{y}_1 \mid \boldsymbol{X}_1, \boldsymbol{\beta}, \sigma^2, \rho)\ \pi(\boldsymbol{\beta}, \sigma^2, \rho)}{\int f(\boldsymbol{y}_1 \mid \boldsymbol{X}_1, \boldsymbol{\theta})\ \pi(\boldsymbol{\theta})\, d\boldsymbol{\theta}}.$$

At $t = 2$ the posterior distribution is

$$p(\boldsymbol{\beta}, \sigma^2, \rho \mid \boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{X}_1, \boldsymbol{X}_2) = \frac{f(\boldsymbol{y}_2 \mid \boldsymbol{X}_2, \boldsymbol{\beta}, \sigma^2, \rho,\ \boldsymbol{y}_1, \boldsymbol{X}_1)\ p(\boldsymbol{\beta}, \sigma^2, \rho \mid \boldsymbol{y}_1, \boldsymbol{X}_1)}{\int f(\boldsymbol{y}_2 \mid \boldsymbol{X}_2, \boldsymbol{\theta},\ \boldsymbol{y}_1, \boldsymbol{X}_1)\ p(\boldsymbol{\theta} \mid \boldsymbol{y}_1, \boldsymbol{X}_1)\, d\boldsymbol{\theta}}.$$

In the tth year, the posterior distribution is

$$p(\boldsymbol{\theta} \mid \boldsymbol{y}_1, \ldots, \boldsymbol{y}_t, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_t) = \frac{f(\boldsymbol{y}_t \mid \boldsymbol{X}_t, \boldsymbol{\theta},\ \boldsymbol{y}_{t-1}, \boldsymbol{X}_{t-1})\ p(\boldsymbol{\theta} \mid \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{t-1}, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_{t-1})}{\int f(\boldsymbol{y}_t \mid \boldsymbol{X}_t, \boldsymbol{\psi},\ \boldsymbol{y}_{t-1}, \boldsymbol{X}_{t-1})\ p(\boldsymbol{\psi} \mid \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{t-1}, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_{t-1})\, d\boldsymbol{\psi}}.$$

Prediction

$$p(\tilde{\boldsymbol{y}}_t \mid \tilde{\boldsymbol{X}}_t,\, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{t-1}, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_{t-1}) = \\ \int f(\tilde{\boldsymbol{y}}_t \mid \tilde{\boldsymbol{X}}_t, \boldsymbol{\theta},\, \boldsymbol{y}_{t-1}, \boldsymbol{X}_{t-1})\; p(\boldsymbol{\theta} \mid \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{t-1}, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_{t-1})\, d\boldsymbol{\theta}$$

- depends on past (observed) data and proposed management actions $\tilde{\boldsymbol{X}}_t$
- integrates over the posterior uncertainty of model parameters
- can be extended for a sequence of proposed management actions , e.g., $p(\tilde{\boldsymbol{y}}_t, \ldots, \tilde{\boldsymbol{y}}_{t+\tau} \mid \tilde{\boldsymbol{X}}_t, \ldots, \tilde{\boldsymbol{X}}_{t+\tau},\, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{t-1}, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_{t-1})$

Selection of Optimal Management Actions

- management objectives are specified in terms of utility (or loss) functions
- "optimal" = management actions with highest *expected* utility (or lowest *expected* loss) averaging over the posterior uncertainty of model parameters and predictions

Example 1 (Target level of vegetation)
Consider the loss function $l(\tilde{\boldsymbol{y}}_2, c) = \sum_{i=1}^{n} |\tilde{y}_{i2} - c|$, where c is a target level of vegetation cover. The expected loss is

$$\begin{aligned} \bar{l}(\tilde{\boldsymbol{X}}_2 \mid \boldsymbol{y}_1, \boldsymbol{X}_1) &= \mathrm{E}_{(\tilde{\boldsymbol{y}}_2 | \tilde{\boldsymbol{X}}_2, \boldsymbol{y}_1, \boldsymbol{X}_1)} \left[l(\tilde{\boldsymbol{y}}_2, c) \right] \\ &= \int l(\tilde{\boldsymbol{y}}_2, c) \; p(\tilde{\boldsymbol{y}}_2 \mid \tilde{\boldsymbol{X}}_2, \boldsymbol{y}_1, \boldsymbol{X}_1) \, d\tilde{\boldsymbol{y}}_2. \end{aligned}$$

If an optimal set of future management actions $\tilde{\boldsymbol{X}}_2^*$ exists, it equals

$$\tilde{\boldsymbol{X}}_2^* = \arg \min_{\tilde{\boldsymbol{X}}_2} \left[\bar{l}(\tilde{\boldsymbol{X}}_2 \mid \boldsymbol{y}_1, \boldsymbol{X}_1) \right].$$

Equivocal Responses in Vegetation Cover

At the end of year 1, suppose we observe the following responses to 2 types of management actions (denoted by $X_1 = 1$ and $X_1 = 2$):

Plot	X_1	y_1
1	1	0.15
2	2	0.55
3	2	0.85
4	1	0.45

Sample mean response to action 1 = 0.30
Sample mean response to action 2 = 0.70

Given these data and a loss function $l(\tilde{\boldsymbol{y}}_2, 0.50) = \sum_{i=1}^{n} |\tilde{y}_{i2} - 0.50|$, we need to select $\tilde{\boldsymbol{X}}_2$ that minimizes the total expected loss.

There are 16 $(= 2^4)$ possible values of $\tilde{\boldsymbol{X}}_2$ to be compared:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1	2	2	1	1	1	1	1	1	2	2	2	2	2	2
1	2	1	2	1	1	1	2	2	2	1	1	1	2	2	2
1	1	1	1	1	2	2	1	2	2	1	2	2	1	2	2
1	1	1	1	2	1	2	2	1	2	2	1	2	2	1	2

Assuming $\rho = 0$ and mutually independent priors:

$\pi(\beta_1) = \pi(\beta_2) \sim \mathrm{U}(0, 1)$

$\pi(\sigma^{-2}) \sim \mathrm{Gamma}(.1, .1)$

yields posterior mean responses $\hat{\beta}_1 = 0.36$ and $\hat{\beta}_2 = 0.64$ and approximately *equal* expected losses (1.59, Monte Carlo SE $= 0.01$) for all 16 values of $\tilde{\boldsymbol{X}}_2$.

Favored Responses in Vegetation Cover

At the end of year 1, suppose we observe the following responses to 2 types of management actions (denoted by $X_1 = 1$ and $X_1 = 2$):

Plot	X_1	y_1
1	1	0.15
2	2	0.55
3	2	0.65
4	1	0.25

Sample mean response to action 1 = 0.20
Sample mean response to action 2 = 0.60

Assuming $\rho = 0$ and mutually independent priors:
$\pi(\beta_1) = \pi(\beta_2) \sim \mathrm{U}(0, 1)$
$\pi(\sigma^{-2}) \sim \mathrm{Gamma}(.1, .1)$

yields posterior mean responses $\hat{\beta}_1 = 0.25$ and $\hat{\beta}_2 = 0.58$.

The expected loss for each of the 16 values of $\tilde{\boldsymbol{X}}_2$ is

1	2	3	4	5	6	7	8
1.55	1.49	1.49	1.46	1.50	1.51	1.47	1.45
9	**10**	**11**	**12**	**13**	**14**	**15**	**16**
1.46	1.41	1.46	1.46	1.41	1.41	1.41	1.36

(Monte Carlo SE = 0.01)

Thus, $\tilde{\boldsymbol{X}}_2 = (2, 2, 2, 2)^T$ is the optimal set of management actions.

Equivocal, but Correlated, Responses in Vegetation Cover

At the end of year 2, suppose we observe the following responses to 2 types of management actions:

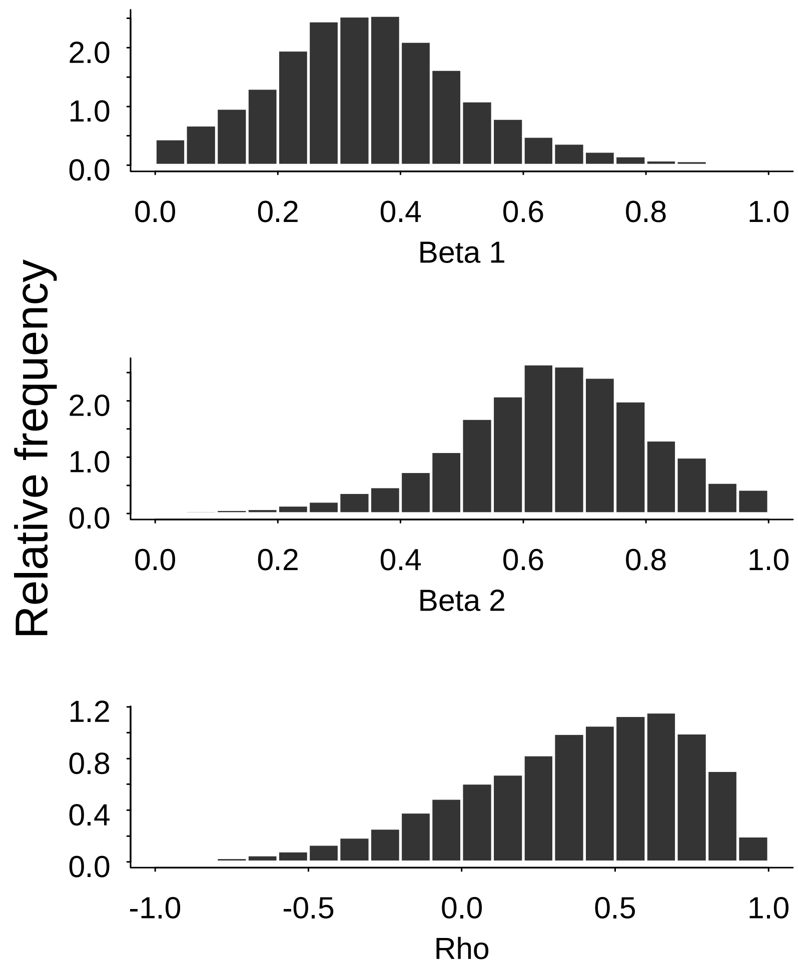
Plot	X_1	y_1	X_2	y_2
1	1	0.15	1	0.25
2	2	0.55	2	0.50
3	2	0.85	2	0.75
4	1	0.45	1	0.50

Sample mean response to action 1 = 0.3375
Sample mean response to action 2 = 0.6625

Assuming mutually independent priors:
$\pi(\beta_1) = \pi(\beta_2) \sim \mathrm{U}(0, 1)$
$\pi(\sigma^{-2}) \sim \mathrm{Gamma}(.1, .1)$
$\pi(\rho) \sim \mathrm{U}(-1, 1)$

yields posterior mean responses $\hat{\beta}_1 = 0.35$ and $\hat{\beta}_2 = 0.65$. The posterior of ρ is highly skewed: mean = 0.37, median = 0.43.

Histogram of the posterior distributions of β_1, β_2, and ρ estimated from the 2 years of vegetation responses



The expected loss for each of the 16 values of $\tilde{\boldsymbol{X}}_3$ is

1	2	3	4	5	6	7	8
1.077	1.008	1.069	1.005	1.143	1.077	1.151	1.077
9	**10**	**11**	**12**	**13**	**14**	**15**	**16**
1.007	1.077	1.148	1.072	1.147	1.071	1.000	1.074

(Monte Carlo SE = 0.005)

Thus, design #15 ($\tilde{\boldsymbol{X}}_3 = (2, 2, 2, 1)^T$) provides the optimal set of management actions; however, designs #2, #4, and #9 provide almost the same expected loss given the Monte Carlo error in the estimates.

Modeling Other Sources of Uncertainty

- Environmental variability
- Errors in sampling, measurement, or application of management actions
- Alternative forms of temporal dependence
- Spatial dependence and variability

Example 2 (Averaging posterior predictions over posterior model uncertainty)

Let $\{H_1, H_2, \ldots, H_K\}$ denote the set of K candidate models, and let $\{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_K\}$ denote their parameters.

At $t = 1$ the posterior distribution of $\boldsymbol{\theta}_k$ for the kth model is

$$p(\boldsymbol{\theta}_k \mid \boldsymbol{y}_1, \boldsymbol{X}_1, H_k) = \frac{f(\boldsymbol{y}_1 \mid \boldsymbol{X}_1, \boldsymbol{\theta}_k, H_k)\; \pi(\boldsymbol{\theta}_k \mid H_k)}{m(\boldsymbol{y}_1 \mid \boldsymbol{X}_1, H_k)},$$

where the marginal likelihood of the data given model H_k $(= m(\boldsymbol{y}_1 \mid \boldsymbol{X}_1, H_k) = \int f(\boldsymbol{y}_1 \mid \boldsymbol{X}_1, \boldsymbol{\theta}_k, H_k)\; \pi(\boldsymbol{\theta}_k \mid H_k)\, d\boldsymbol{\theta}_k)$ is finite provided $\pi(\boldsymbol{\theta}_k \mid H_k)$ is proper.

If K is finite (and not too large), the posterior uncertainty in a particular model also may be computed using Bayes theorem:

$$p(H_k \mid \boldsymbol{y}_1, \boldsymbol{X}_1) = \frac{m(\boldsymbol{y}_1 \mid \boldsymbol{X}_1, H_j)\; \pi(H_k)}{\sum_{j=1}^{K} m(\boldsymbol{y}_1 \mid \boldsymbol{X}_1, H_j)\; \pi(H_j)}$$

assuming $\sum_{j=1}^{K} \pi(H_j) = 1$.

The values of $p(H_k \mid \boldsymbol{y}_1, \boldsymbol{X}_1)$ are used to weigh the model-specific predictions and thereby average over model uncertainty:

$$p(\tilde{\boldsymbol{y}}_2 \mid \tilde{\boldsymbol{X}}_2, \boldsymbol{y}_1, \boldsymbol{X}_1) = \sum_{k=1}^{K} p(\tilde{\boldsymbol{y}}_2 \mid \tilde{\boldsymbol{X}}_2, \boldsymbol{y}_1, \boldsymbol{X}_1, H_k)\; p(H_k \mid \boldsymbol{y}_1, \boldsymbol{X}_1)$$

In practice, each draw from $[\tilde{\boldsymbol{y}}_2 \mid \tilde{\boldsymbol{X}}_2, \boldsymbol{y}_1, \boldsymbol{X}_1, H_k]$ *is selected with probability* $p(H_k \mid \boldsymbol{y}_1, \boldsymbol{X}_1)$ *to obtain a draw from* $[\tilde{\boldsymbol{y}}_2 \mid \tilde{\boldsymbol{X}}_2, \boldsymbol{y}_1, \boldsymbol{X}_1]$.

Example 3 (A sequence of management actions)

We cannot seriously expect our model assumptions to remain valid indefinitely long; however, suppose we have observed $\boldsymbol{X}_1$ and $\boldsymbol{y}_1$ and want to predict an optimal sequence of future management actions $(\tilde{\boldsymbol{X}}_2^*, \tilde{\boldsymbol{X}}_3^*, \ldots, \tilde{\boldsymbol{X}}_\tau^*)$ to be implemented in the next $\tau - 1$ years.

Then minimize the expected loss

$$\begin{aligned}
\bar{l}(\tilde{\boldsymbol{X}}_2, \ldots, \tilde{\boldsymbol{X}}_\tau \mid \boldsymbol{y}_1, \boldsymbol{X}_1) &= \mathrm{E}_{(\tilde{\boldsymbol{y}}_2, \ldots, \tilde{\boldsymbol{y}}_\tau \mid \tilde{\boldsymbol{X}}_2, \ldots, \tilde{\boldsymbol{X}}_\tau, \boldsymbol{y}_1, \boldsymbol{X}_1)} \left[l(\tilde{\boldsymbol{y}}_2, \ldots, \tilde{\boldsymbol{y}}_\tau, c) \right] \\
&= \int l(\tilde{\boldsymbol{y}}_2, \ldots, \tilde{\boldsymbol{y}}_\tau, c) \; p(\tilde{\boldsymbol{y}}_2, \ldots, \tilde{\boldsymbol{y}}_\tau \mid \tilde{\boldsymbol{X}}_2, \ldots, \tilde{\boldsymbol{X}}_\tau, \, \boldsymbol{y}_1, \boldsymbol{X}_1) \, d\tilde{\boldsymbol{y}} \\
&\doteq \frac{1}{R} \sum_{r=1}^{R} l(\tilde{\boldsymbol{y}}_2^{(r)}, \ldots, \tilde{\boldsymbol{y}}_\tau^{(r)}, c)
\end{aligned}$$

where $\tilde{\boldsymbol{y}} = (\tilde{\boldsymbol{y}}_2, \ldots, \tilde{\boldsymbol{y}}_\tau)^T$.

Based on our model, random draws from $[\tilde{\boldsymbol{y}}_2, \ldots, \tilde{\boldsymbol{y}}_\tau \mid \tilde{\boldsymbol{X}}_2, \ldots, \tilde{\boldsymbol{X}}_\tau,\, \boldsymbol{y}_1, \boldsymbol{X}_1]$ may be computed by taking random draws from an appropriately ordered sequence of conditional posterior predictive distributions since

$$\begin{aligned} p(\tilde{\boldsymbol{y}}_2, \ldots, \tilde{\boldsymbol{y}}_\tau \mid \tilde{\boldsymbol{X}}_2, \ldots, \tilde{\boldsymbol{X}}_\tau,\, \boldsymbol{y}_1, \boldsymbol{X}_1) &= p(\tilde{\boldsymbol{y}}_2 \mid \tilde{\boldsymbol{X}}_2, \boldsymbol{y}_1, \boldsymbol{X}_1)\, p(\tilde{\boldsymbol{y}}_3 \mid \tilde{\boldsymbol{y}}_2, \tilde{\boldsymbol{X}}_2, \tilde{\boldsymbol{X}}_3, \boldsymbol{y}_1, \boldsymbol{X}_1) \\ &\quad \cdots\; p(\tilde{\boldsymbol{y}}_\tau \mid \tilde{\boldsymbol{y}}_2, \ldots, \tilde{\boldsymbol{y}}_{\tau-1}, \tilde{\boldsymbol{X}}_2, \ldots, \tilde{\boldsymbol{X}}_\tau, \boldsymbol{y}_1, \boldsymbol{X}_1). \end{aligned}$$

Example 4 (Dual-control problem)

Express competing management objectives in terms of:

$u_1(\tilde{\boldsymbol{y}}_2, c)$ = utility of achieving target level of vegetation cover

$u_2(\boldsymbol{\theta})$ = utility of learning about model parameters identified with treatment effects (e.g., Kullback-Leibler distance between posteriors $[\boldsymbol{\theta} \mid \tilde{\boldsymbol{y}}_2, \tilde{\boldsymbol{X}}_2, \boldsymbol{y}_1, \boldsymbol{X}_1]$ and $[\boldsymbol{\theta} \mid \boldsymbol{y}_1, \boldsymbol{X}_1]$).

Specify a joint utility function

$$u(\tilde{\boldsymbol{y}}_2, \boldsymbol{\theta}, c, \gamma, \phi) = \gamma \cdot u_1(\tilde{\boldsymbol{y}}_2, c) + \phi \cdot u_2(\boldsymbol{\theta})$$

where γ and ϕ are nonnegative constants.

The expected utility is

$$\begin{aligned}
\overline{u}(\tilde{\boldsymbol{X}}_2 \mid \boldsymbol{y}_1, \boldsymbol{X}_1) &= E_{(\boldsymbol{\theta}, \tilde{\boldsymbol{y}}_2 \mid \tilde{\boldsymbol{X}}_2, \boldsymbol{y}_1, \boldsymbol{X}_1)} \left[\gamma \cdot u_1(\tilde{\boldsymbol{y}}_2, c) + \phi \cdot u_2(\boldsymbol{\theta})\right] \\
&= \int \gamma \; u_1(\tilde{\boldsymbol{y}}_2, c) \; p(\tilde{\boldsymbol{y}}_2 \mid \tilde{\boldsymbol{X}}_2, \boldsymbol{y}_1, \boldsymbol{X}_1) \; d\tilde{\boldsymbol{y}}_2 + \\
&\quad \int \int \phi \; u_2(\boldsymbol{\theta}) \; p(\boldsymbol{\theta} \mid \tilde{\boldsymbol{y}}_2, \tilde{\boldsymbol{X}}_2, \boldsymbol{y}_1, \boldsymbol{X}_1) \; p(\tilde{\boldsymbol{y}}_2 \mid \tilde{\boldsymbol{X}}_2, \boldsymbol{y}_1, \boldsymbol{X}_1) \; d\boldsymbol{\theta} \, d\tilde{\boldsymbol{y}}_2 \\
&= E_{(\tilde{\boldsymbol{y}}_2 \mid \tilde{\boldsymbol{X}}_2, \boldsymbol{y}_1, \boldsymbol{X}_1)} \left[\gamma \cdot u_1(\tilde{\boldsymbol{y}}_2, c) + E_{(\boldsymbol{\theta} \mid \tilde{\boldsymbol{y}}_2, \tilde{\boldsymbol{X}}_2, \boldsymbol{y}_1, \boldsymbol{X}_1)} [\phi \cdot u_2(\boldsymbol{\theta})]\right]
\end{aligned}$$

If an optimal set of future management actions $\tilde{\boldsymbol{X}}_2^*$ exists, it is

$$\tilde{\boldsymbol{X}}_2^* = \arg \max_{\tilde{\boldsymbol{X}}_2} \left[\overline{u}(\tilde{\boldsymbol{X}}_2 \mid \boldsymbol{y}_1, \boldsymbol{X}_1)\right].$$

Summary and Research Needs

- Modern methods of Bayesian inference and decision making are capable of solving relatively complex problems of natural resource management.
  - spatial and temporal dependence
  - partial controllability
  - model uncertainty
- We anticipate widespread use of these methods, particularly as software is developed for computing posterior distributions of model parameters and predictions (*see Appendix C of Carlin and Louis 2000*).
- Computationally efficient algorithms are needed for determining optimal *sequences* of future management actions.

(Full) Conditional Posterior Distributions

If inferences are based on 1 year of data:

$$p(\tau \mid \boldsymbol{\beta}, \rho_0, \boldsymbol{y}_1, \boldsymbol{X}_1) \;\propto\; \tau^{n/2+\varepsilon_1-1} \exp\left[-\tau\left(\varepsilon_2 + \frac{d_1(1-\rho_0^2)}{2}\right)\right]$$

$$p(\beta_j \mid \beta_{k(\neq j)}, \tau, \rho_0, \boldsymbol{y}_1, \boldsymbol{X}_1) \;\propto\; \exp\left[-\frac{\tau d_1(1-\rho_0^2)}{2}\right]$$

where $\tau = 1/\sigma^2$ and $d_1 = \sum_{i=1}^{n}(y_{i1} - \boldsymbol{x}_{i1}^T\boldsymbol{\beta})^2$

$$\tau \mid \boldsymbol{\beta}, \rho_0, \boldsymbol{y}_1, \boldsymbol{X}_1 \sim \mathrm{Gamma}\left(\varepsilon_1 + n/2,\; \varepsilon_2 + \frac{d_1(1-\rho_0^2)}{2}\right)$$

(Full) Conditional Posterior Distributions

If inferences are based on 2 years of data:

$$p(\tau \mid \boldsymbol{\beta}, \rho, \boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{X}_1, \boldsymbol{X}_2) \;\propto\; \tau^{n+\varepsilon_1-1} \exp\Big[-\tau\Big(\varepsilon_2 + \frac{d_1(1-\rho^2)}{2} + \frac{d_2}{2}\Big)\Big]$$

$$p(\beta_j \mid \beta_{k(\neq j)}, \mid \tau, \rho, \boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{X}_1, \boldsymbol{X}_2) \;\propto\; \exp\Big[-\tau\Big(\frac{d_1(1-\rho^2)}{2} + \frac{d_2}{2}\Big)\Big]$$

$$p(\rho \mid \boldsymbol{\beta}, \tau, \boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{X}_1, \boldsymbol{X}_2) \;\propto\; (1-\rho^2)^{n/2} \exp\Big[-\tau\Big(\frac{d_1(1-\rho^2)}{2} + \frac{d_2}{2}\Big)\Big]$$

where $\tau = 1/\sigma^2$, $d_1 = \sum_{i=1}^{n}(y_{i1} - \boldsymbol{x}_{i1}^T\boldsymbol{\beta})^2$, and $d_2 = \sum_{i=1}^{n}(y_{i2} - \boldsymbol{x}_{i2}^T\boldsymbol{\beta} - \rho(y_{i1} - \boldsymbol{x}_{i1}^T\boldsymbol{\beta}))^2$

$$\tau \mid \boldsymbol{\beta}, \rho, \boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{X}_1, \boldsymbol{X}_2 \sim \text{Gamma}\left(\varepsilon_1 + n,\ \varepsilon_2 + \frac{d_1(1-\rho^2)}{2} + \frac{d_2}{2}\right)$$

WinBugs code for autoregressive model of waterfowl habitat

```
model {
   SigmaInv <- tau*(1.-rho*rho)
   for (i in 1:n) {
      y1[i] ~ dnorm(beta[x1[i]], SigmaInv)
      mu2[i] <- beta[x2[i]] + rho*(y1[i]-beta[x1[i]])
      y2[i] ~ dnorm(mu2[i], tau)
      for (j in 1:ndesigns) {
         mu[i,j] <- beta[xp[i,j]] + rho*(y2[i]-beta[x2[i]])
         yp[i,j] ~ dnorm(mu[i,j], tau)
         loss[i,j] <- abs(yp[i,j] - ytarget)
      }
   }
   for (j in 1:ndesigns) {
      totalLoss[j] <- sum(loss[,j])
   }
   beta[1]~ dunif(0,1)
   beta[2] ~ dunif(0,1)
   tau ~ dgamma(.1,.1)
   rho ~ dunif(-1,1)
}
```

WinBugs code for autoregressive model of waterfowl habitat

```
DATA
list(n=4, ndesigns=16, y1=c(0.15, 0.55, 0.85, 0.45), y2=c(0.25, 0.50, 0.75,
0.50), x1=c(1,2,2,1), x2=c(1,2,2,1), ytarget=0.5)


xp[,1] xp[,2] xp[,3] xp[,4] xp[,5] xp[,6] xp[,7] xp[,8] xp[,9]
xp[,10] xp[,11] xp[,12] xp[,13] xp[,14] xp[,15] xp[,16]
1 1 2 2 1 1 1 1 1 1 2 2 2 2 2 2
1 2 1 2 1 1 1 2 2 2 1 1 1 2 2 2
1 1 1 1 1 2 2 1 2 2 1 2 2 1 2 2
1 1 1 1 2 1 2 2 1 2 2 1 2 2 1 2


INITS
list(beta=c(.25,.75), tau=25.0, rho=0.0)
```