Sebastes aleutianus (rougheye rockfish), possibly born the same year as Darwin (maximum age of S. aleutianus ~205 yrs)
Natural Resource Modeling in the 21st Century: The Certainty of Uncertainty

Marc Mangel

Olin College Keynote Talk
June 2007

Steve Munch

Thanasis Kottas
• Thank you, RMA

• The nature of environmental problems

• Wicked problems are swathed in uncertainty

• Hypothesis testing

• A retro-classic: competition between species of flour beetles

• Back to the future: Bayesian Non-Parametric methods for stock-recruitment relationships

• 7 principles for ensuring that our work remains relevant
(First?) RMA Meeting in Arcata 1982…then back to Davis
(First?) RMA Meeting in Arcata 1982...then back to Davis

FORAGING AND FLOCKING STRATEGIES: INFORMATION IN AN UNCERTAIN ENVIRONMENT

COLIN W. CLARK AND MARC MANGEL
Department of Mathematics, University of British Columbia, Vancouver V6T 1W5, Canada; Department of Mathematics, University of California, Davis, California 95616

Submitted December 16, 1982; Accepted October 6, 1983
(First?) RMA Meeting in Arcata 1982...then back to Davis

FORAGING AND FLOCKING STRATEGIES: INFORMATION IN AN UNCERTAIN ENVIRONMENT

COLIN W. CLARK AND MARC MANGEL
Department of Mathematics, University of British Columbia, Vancouver V6T 1W5, Canada;
Department of Mathematics, University of California, Davis, California 95616

Submitted December 16, 1982; Accepted October 6, 1983

TOWARDS A UNIFIED FORAGING THEORY

MARC MANGEL
Departments of Mathematics, Agricultural Economics, and Entomology,
University of California, Davis, California 95616 USA

AND

COLIN W. CLARK
Institute of Applied Mathematics, University of British Columbia, Vancouver,
British Columbia V6T 1W5, Canada
(First?) RMA Meeting in Arcata 1982...then back to Davis
(First?) RMA Meeting in Arcata 1982…then back to Davis

FORAGING AND FLOCKING STRATEGIES: INFORMATION IN AN UNCERTAIN ENVIRONMENT

COLIN W. CLARK AND MARC MANGEL

Department of Mathematics, University of British Columbia, Vancouver V6T 1W5, Canada; Department of Mathematics, University of California, Davis, California 95616

Submitted December 16, 1982; Accepted October 6, 1983

Thank you, RMA!

TOWARDS A UNIFIED FORAGING THEORY

MARC MANGEL

Departments of Mathematics, Agricultural Economics, and Entomology, University of California, Davis, California 95616 USA

AND

COLIN W. CLARK

Institute of Applied Mathematics, University of British Columbia, Vancouver, British Columbia V6T 1W5, Canada

Dynamic Modeling in Behavioral Ecology

MARC MANGEL
COLIN W. CLARK
Environmental Problems are Wicked

-- No definitive formulation

-- No stopping rule

-- No final resolution

-- Solutions are “good or bad” not “right or wrong”

-- Plurality of legitimate perspectives

“The best environmental policy depends on how you frame the question”

---John Maddox

Wicked problems are swathed in uncertainty

- Principle of irreducible uncertainty: No matter how much science we do, there will always be remaining a level of uncertainty.

We need to think about data and their interpretation
The Kinds of Uncertainty

Epistemic Uncertainty

Uncertainty due to limited information

Ignorance
Observation Error
Type B
Reducible

Property of analyst
The Kinds of Uncertainty

<table>
<thead>
<tr>
<th>Epistemic Uncertainty</th>
<th>Aleatory Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uncertainty due to limited information</td>
<td>Uncertainty due to randomness in system</td>
</tr>
<tr>
<td>Ignorance</td>
<td>Variability</td>
</tr>
<tr>
<td>Observation Error</td>
<td>Process error</td>
</tr>
<tr>
<td>Type B</td>
<td>Type A</td>
</tr>
<tr>
<td>Reducible</td>
<td>Irreducible</td>
</tr>
<tr>
<td>Property of analyst</td>
<td>Property of system</td>
</tr>
</tbody>
</table>
The Kinds of Uncertainty

Epistemic Uncertainty
Uncertainty due to limited information
Ignorance
Observation Error
Type B
Reducible
Property of analyst

Aleatory Uncertainty
Uncertainty due to randomness in system
Variability
Process error
Type A
Irreducible
Property of system

Both are important for Management of Natural Resources
The classical/frequentist approach to dealing with uncertainty: The Earth is Round (p< 0.05)

“I argue herein that NHST (null hypothesis significance testing has not only failed to support the advance of psychology as a science but also has seriously impeded it” (J. Cohen. 1994. American Psychologist. 49:997-1003)

What is a null hypothesis?

Null hypotheses entertain the possibility that nothing has happened, that a process has not occurred, or that change has not been produced by a cause of interest. They are reference points against which alternatives should be contrasted.

However, since it is often impossible to prove that something has occurred, we construct a null hypothesis that is the complement of the hypothesis of interest and use the accumulated data to assess the probability that the null hypothesis is true.
Examples of Environmental Null Hypotheses That Were Rejected

• The occurrence of sheep remains in coyote scats does not vary across season (p=0.03)

• Duckling body mass does not vary across years (p<0.0001)

• The density of large trees is not greater in unlogged than logged forests (p=0.02)


• Driving cessation [in the elderly] leads to a decline in out-of-home activity (p<0.001)

What’s wrong with NHST?

“Well, among many other things, it does not tell us what we want to know, and we so much want to know what we want to know that, out of desperation, we nevertheless believe that it does.

What we want to know is

*Given these data, what is the probability that the null hypothesis is true?*

But, as most of us know, what it tells us is

*Given that the null hypothesis is true, what is the probability of these (or more extreme) data”*

(Cohen pg 997)

What we want for scientific understanding

<table>
<thead>
<tr>
<th>What we get from NHST</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Pr{H</td>
</tr>
</tbody>
</table>

NHST has “caused scientific research workers to pay undue attention to the results of the tests of significance they perform on their data...and too little to the estimates of the magnitude of the effects they are estimating”

But Sometimes You Might (Jagger and Richards, op. cit.)

Likelihood and Bayesian Methods Show the Way to Get What You Want

“...the discipline of statistics has neglected a key question for which it is responsible: when does a given set of observations support one statistical hypothesis over another?”

The principle fundamental to providing this answer is the law of likelihood, which “provides the explicit objective quantitative concept of evidence that is missing”

Law of Likelihood: If hypothesis $A$ implies that the probability that a random variable $X$ takes the value $x$ is $p_A(x)$, while hypothesis $B$ implies that the probability is $p_B(x)$, then the observation $X=x$ is evidence supporting $A$ over $B$ only if $p_A(x)>p_B(x)$, and the likelihood ratio, $p_A(x)/p_B(x)$, measures the strength of that evidence.

Simple example

H is the hypothesis of interest, with prior probability p; H_0 is the alternative hypothesis, with prior probability 1-p.

We collect data. Then

\[
Pr\{H|\text{data}\} = \frac{p Pr\{\text{data|}H\}}{p Pr\{\text{data|}H\} + (1-p) Pr\{\text{data|}H_0\}}
\]

Bayesian and likelihood methods allow us to deal with uncertainty and information in a consistent framework.
Ecological Detection in Environmental Problem Solving

”Method of multiple working hypotheses”

--T.C. Chamberlain (1890)

The three questions we need to ask

Given two (or more) hypotheses/models and an observation (data) we can ask

• What do I believe, now that I have this observation?

• What should I do, now that I have this observation?

• How should I interpret this observation as evidence regarding the different models/hypotheses?
Tribolium confusum
(The confused flour beetle)


Tribolium castaneum
(The red flour beetle)

(Clemson University USDA Cooperative Extension Slide Series)
A new kind of stat lab
Third Berkeley Symposium on Mathematical Statistics and Probability, Volume IV

1956

Inspired work by

VD Barnett
MS Bartlett
JC Gower
PH Leslie
The key result (for this talk): Zones of Determinate Outcome and of Indeterminate Outcome

Treatments differ by temperature and humidity.
Focal question

Given initial numbers of the two species, what will be the outcome of the competition?
In retrospect, it is perhaps informative that Park initially was not motivated by questions of stochastic events but, rather, with the demonstration that certain reasonably realistic ecological factors could affect the survival of one species over another when both were competing. It is also amusing to record, again in retrospect, that Park’s growing appreciation of indeterminacy was nurtured by the data themselves, by his association with the statisticians PH Leslie, J Neyman and MS Bartlett, who were interested, each in his own way, in the Tribolium experiments; and, finally, that Park had an increasing awareness of the fact that ecological systems are indeed stochastic -- that a population’s survival, even in the absence of environmental stochasticity, is related intimately to a component which has recently, and meaningfully, been described as ‘gambling for existence’

D Mertz, DA Cawthon and T Park. 1975. PNAS 73:1368
A variety of models

Model 1: There are no underlying deterministic dynamics, it is all stochastic (Neyman, Park and Scott; Leslie and Gower; Bartlett, Leslie and Gower; Barnett)

\[ X_1(t) = \text{Adults of } T.\text{castaneum} \]

\[ X_2(t) = \text{Adults of } T.\text{confusum} \]

Set \( \Delta X_i = X_i(t + \Delta t) - X_i(t) \) Then, given \( X_i(t) = x_i \)

\[
\Pr\{\Delta X_1 = 1, \Delta X_2 = 0 \mid \text{change}\} = \frac{x_1 \lambda_1}{k}
\]

\[
\Pr\{\Delta X_1 = -1, \Delta X_2 = 0 \mid \text{change}\} = \frac{x_1 \mu_1}{k}
\]

\[
\Pr\{\Delta X_1 = 0, \Delta X_2 = 1 \mid \text{change}\} = \frac{x_2 \lambda_2}{k}
\]

\[
\Pr\{\Delta X_1 = 0, \Delta X_2 = -1 \mid \text{change}\} = \frac{x_2 \mu_2}{k}
\]
Pr\{\Delta X_1 = 1, \Delta X_2 = 0 \mid \text{change}\} = \frac{x_1 \lambda_1}{k}

Pr\{\Delta X_1 = -1, \Delta X_2 = 0 \mid \text{change}\} = \frac{x_1 \mu_1}{k}

Pr\{\Delta X_1 = 0, \Delta X_2 = 1 \mid \text{change}\} = \frac{x_2 \lambda_2}{k}

Pr\{\Delta X_1 = 0, \Delta X_2 = -1 \mid \text{change}\} = \frac{x_2 \mu_2}{k}

\lambda_1 = 0.11 - 0.0007x_1 - 0.01x_2
\lambda_2 = 0.08 - 0.0007x_1 - 0.007x_2
\mu_1 = 0.01
\mu_2 = 0.005

k = x_1(\lambda_1 + \mu_1) + x_2(\lambda_2 + \mu_2)

Suitable for rapid Monte Carlo simulation even 50 years ago

Deterministic dynamics as an emergent property (Mangel. 1994. Barrier transitions driven by fluctuations, with applications to ecology and evolution. Theoretical Population Biology 45:16-40)
Model 2. Deterministic dynamics with a saddle point perturbed by noise

\[
\frac{dx}{dt} = x(1 + a - x - ay) \\
\frac{dy}{dt} = y(1 + a - y - ax)
\]

\[ dX_1 = x_1(\lambda_1 - \mu_1)dt + \sum_j \varepsilon a_{1,j}dW_j \]

\[ dX_2 = x_2(\lambda_2 - \mu_2)dt + \sum_j \varepsilon a_{2,j}dW_j \]

Here \( \varepsilon \sim 1/\text{Steady state population size} \). With the previous stochastic assumptions

\[
\varepsilon a = \begin{bmatrix}
\frac{\varepsilon x_1(\lambda_1 + \mu_1)}{k} & 0 \\
0 & \frac{\varepsilon x_2(\lambda_2 + \mu_2)}{k}
\end{bmatrix}
\]
Surround the separatrix with a band and set

$$u(x_1, x_2) = \Pr\{\text{exit the band towards the } y\text{-axis} | X_i(0) = x_i\}$$

Derive the Kolmogorov backward equation, construct its solution

Contours of probability (indeterminacy)
Practice problem for Model 2

\[ dX = xdt + \varepsilon \sqrt{dt} \]

\[ dW \sim N(0, dt) \]

\[ u(x) = \Pr\{X(t) \text{ first exits } [-1,1] \text{ through 1}\mid X(0)=x\} \]

Show that

\[ 0 = \frac{\varepsilon}{2} \frac{\partial^2 u}{\partial x^2} + x \frac{\partial u}{\partial x} \]

Determine the appropriate boundary conditions and solve the equation. What is the behavior of the solution for small \( \varepsilon \)?
Model 3 It’s deterministic dynamics but initial conditions are unknown (R.V. Wiederkehr, personal communication 1977)

\[
\frac{dX_1}{dt} = x_1(\lambda_1 - \mu_1); X_1(0) = x_{10} \exp\{Z_{\sigma_1}\}
\]

\[
\frac{dX_2}{dt} = x_2(\lambda_2 - \mu_2); X_2(0) = x_{20} \exp\{Z_{\sigma_2}\}
\]
Model 4 It is deterministic dynamics with random effects (a 21st century problem)

\[
\frac{dX_1}{dt} = x_1(\lambda_{10} - \lambda_{11}x_1 - \lambda_{12}x_2 - \mu_1)
\]

\[
\frac{dX_2}{dt} = x_2(\lambda_{20} - \lambda_{21}x_1 - \lambda_{22}x_2 - \mu_2)
\]

\[\lambda_{ij} \sim ??\]

\[\mu_i \sim ??\]

We might

- Draw from the distributions and use the solution of the differential equations to compute the zones of indeterminancyn and update the distributions

- Construct a GLM for the rates in terms of environmental and genetic variables

- BNP the right hand side to let the data tell us the functional response
Back to the Future: Stock-Recruitment Relationships

Ricker

\[ R(S) = aSe^{-bS} \]

Beverton Holt

\[ R(S) = \frac{aS}{1 + bS} \]

Shepherd

\[ R(S) = \frac{aS}{1 + b\left(\frac{S}{K}\right)^c} \]
The data are noisy.

Sometimes, the choice of stock-recruitment relationship does not matter

…..but sometimes it does
We have mechanistic descriptions for some of these models

Beverton-Holt

\[ \frac{dR}{d\tau} = -m_1 R - m_2 R^2 \]
\[ R(0) = \phi S \]

Ricker

\[ \frac{dR}{d\tau} = -m_1 R - m_2 RS \]
\[ R(0) = \phi S \]

And could thus argue which is ‘right’
Which model is
• Best?
• Right?

Figure 9. Pollock stock recruitment plot with estimated Ricker curves for Models A (constant natural mortality) and E (with predation).

Which model is

• Best?

• Right?

Is this a

21st Century

Question?

Figure 9. Pollock stock recruitment plot with estimated Ricker curves for Models A (constant natural mortality) and E (with predation).

A Wish List for Stock Recruitment Functions. A framework that....

- Gives a single, unified framework for considering stock-recruitment functions

\[ S(t + 1) = fS(t) + R(S(t)) \]

- Allows incorporation of prior biological information

- Is flexible enough to allow the data to dictate the shape of the function

- Has easily computable confidence intervals, and management reference points

\[ R'(S_{MSY}) = 1 - f \]
BNP (Probability Distributions on Functions) Delivers

Bayesian approaches
- Explicitly model uncertainty in parameters.
- Allows posterior evaluation of the odds of competing models
- Forces honesty

Nonparametric approaches
- Allow data to dictate shape of curve (or dispense w/ curve altogether)
- Inference about recruitment at low stock size is insensitive to points far away
- Functional form is fixed
- Inferences about recruitment at low stock sizes affected by data at all stock sizes
- No biological underpinnings
- Overfitting

Components of a Bayesian Analysis

- Prior for an unknown
- Data that depend upon the unknown
- A probability model (the likelihood) that links the unknown and the data
- Posterior for the unknown using Bayes Theorem, called ‘updating’ the prior given the data
- For multiple, sequential observations, the current posterior becomes the next prior
In the context of stock-recruitment relationships
We can think of correlations between two values of recruitment depending upon spawning stock size.

\[
R(S_1) \quad R(S_2)
\]
Some technical details.

\[ X = \log(S) \]
\[ Y = \log(R) \]

have a relationship

\[ y = \mathcal{I}(x) + \varepsilon \]

Where \( \varepsilon \) is a standard normal random (observation error) variable but

\[ \mathcal{I}(x) \]

is unknown and described by a probability model.
We adopt a Gaussian prior:

\( I(x) \) is completely defined by a mean function \( m(x) \) and a covariance \( C(x,x') \) which determines how much wiggle is possible between values of \( x \).

We begin with a prior mean function for the possible points, for example for the Ricker:

\[
\mu(x) = \ln(a) + x - be^x
\]
The covariance function determines how values of $\mathcal{I}(x)$ are related for different values of $x$. We use

$$C(x, x') = \tau^2 \exp\left[-\phi|\alpha|^{\alpha}\right] \quad \tau > 0, \phi > 0, \alpha \in (0,2]$$

With $t^2$, $f$ and $a$ parameters (to be determined).

- $a$ determines the fine scale variability (graininess)
- $f$ determines the large scale variability (range of dependence)
Our model is thus essentially one in which the observed log-recruitment is multivariate normal

\[
\{y_1, \ldots, y_n\} \mid \theta \sim N_n(\mu^n(\theta), C^n(\theta) + I^n\sigma^2_\varepsilon)
\]

Where we simultaneously learn about parameters and functional forms
A “few” details remain:

--- specifics of updating (modern MCMC methods)
--- choosing priors
--- posterior prediction for new stock sizes
A “few” details remain:

-- specifics of updating (modern MCMC methods)

-- choosing priors

-- posterior prediction for new stock sizes
Testing the approach

– Can we predict correct shape when underlying model is known?

– Is there a tendency to over-fitting?

– Can we use the method to obtain estimates of management reference points?
The contestants are....

**Ricker**  \[ R = aSe^{-bS} \]

**Beverton Holt**  \[ R = \frac{aS}{1 + bS} \]

**Maynard Smith-Shepherd**  \[ R = \frac{aS}{1 + (bS)^c} \]

**Saila-Lorda**  \[ R = aS^c e^{-bS} \]
Open mixture

\[ R = faS^c e^{-bS} + (1 - f) \frac{\alpha S}{1 + \beta S} + m \]

Procedure: Generate data with one of the models, fit it with all and compare
Example: Ricker (Low Observation Error)

Recruits

Spawners

Blue-true model
Green-Ricker fit
Magenta-BevHolt fit
Red-Lowess
Black-NP Bayes
Summary -- Residual Sum of Squares Tells the Story

Low observation error
High observation error

![Graph showing residual sum of squares of fitted model for different generating models: Ricker, Beverton Holt, Shepherd, Saila-Lorda, and Open-Mixture. The graph indicates that the BNP model has a significantly higher residual sum of squares compared to the others, suggesting high observation error.]
And now for some data --- Sockeye salmon

Fry

Females (x10^4)
Bayesian methods force us to be honest about the unknown unknowns
What about reference points?

• Need numerical definitions in order to implement reference points in the BNP framework
  • Unfished biomass
  • Steepness (recruitment at 20% of unfished biomass)
  • Biomass at maximum sustainable yield

• Obtain posterior predictive distributions for each reference point.

• Compare posterior predictive mean to analytical solutions for Ricker and Beverton-Holt models
We generate distributions for these quantities

- Unfished biomass
- MSY Stock Size
- Steepness

Filled triangle - true value
Open triangle - estimated value
Seven principles for ensuring that our work is relevant

1) The highest function of ecology is understanding consequences (Frank Herbert, *Dune*)
Seven principles for ensuring that our work is relevant

1) The highest function of ecology is understanding consequences (Frank Herbert, *Dune*)

2) Variation is not noise. It is the core of biology.
Seven principles for ensuring that our work is relevant

1) The highest function of ecology is understanding consequences (Frank Herbert, *Dune*)

2) Variation is not noise. It is the core of biology.

3) The simpler the mathematical tool, the more likely it is to deliver the goods (John Hammersley)
Seven principles for ensuring that our work is relevant

1) The highest function of ecology is understanding consequences (Frank Herbert, *Dune*)

2) Variation is not noise. It is the core of biology.

3) The simpler the mathematical tool, the more likely it is to deliver the goods (John Hammersley)

4) I would not give a fig for simplicity this side of complexity, but I would give my life for simplicity on the other side of complexity (Oliver Wendell Holmes)
5) Ecological modeling is iterative: The form of the model depends on the question and analysis should inform what to do next.

(For a good example, see Ch 4 of CW Clark and M Mangel. 2000. *Dynamic State Variable Models in Ecology. Methods and Applications*. Oxford University Press. Also see Ch 3 on “Using The Model”)

5) Ecological modeling is iterative: The form of the model depends on the question and analysis should inform what to do next

(For a good example, see Ch 4 of CW Clark and M Mangel. 2000. *Dynamic State Variable Models in Ecology. Methods and Applications*. Oxford University Press. Also see Ch 3 on “Using The Model”)

6) Know your data -- a tool may be mathematically correct and biologically irrelevant

From M-J Fortin
7) Hierarchical models link science & signal and statistics & sampling; they work for Bayesian and frequentist alike (you just have to ‘believe’ in conditional probability)

\[ P[y, z, \theta] = f(y \mid z)g(z \mid \theta)h(\theta) \]
The location of our future

Nature IS variable.
The location of our future

Nature IS variable. AND much of that variability can be understood
Nature IS variable. AND much of that variability can be understood