

Bayesian population analysis using WinBUGS

Marc Kéry & Michael Schaub



USGS Patuxent Wildlife Research Center, Laurel
8-12 July 2013

Preface to workshop



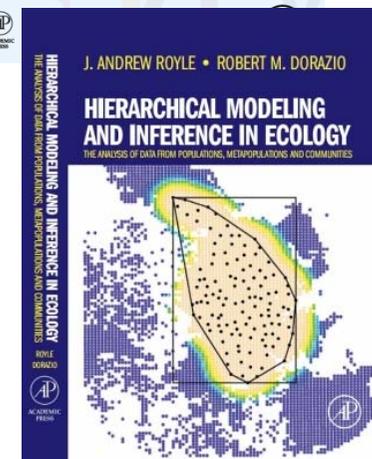
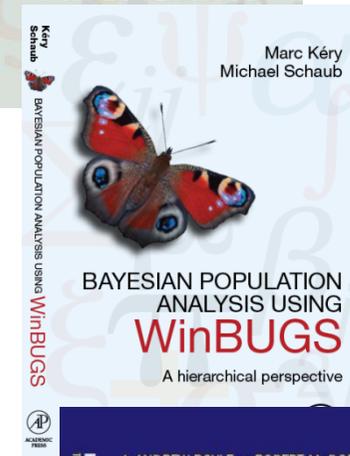
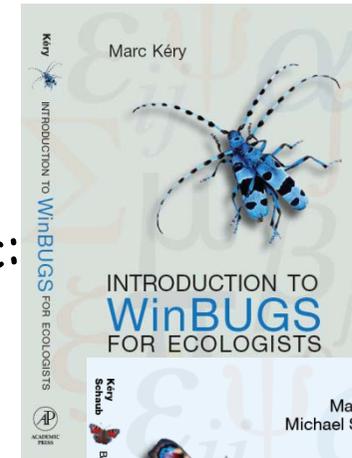
- Introduce Bayesian population analysis using BUGS & JAGS
- Learning by doing for wide range of statistical models relevant to population ecologists
- From ecologists for ecologists: examples from own research
- Key style concepts:
 - Gentle introduction, light reading
 - Lots of fully documented example analyses
 - WinBUGS/OpenBUGS/JAGS run from program R
 - Simulated data sets & real-life data
 - Clear and consistent layout of R/WinBUGS code



Preface to workshop, ctd.



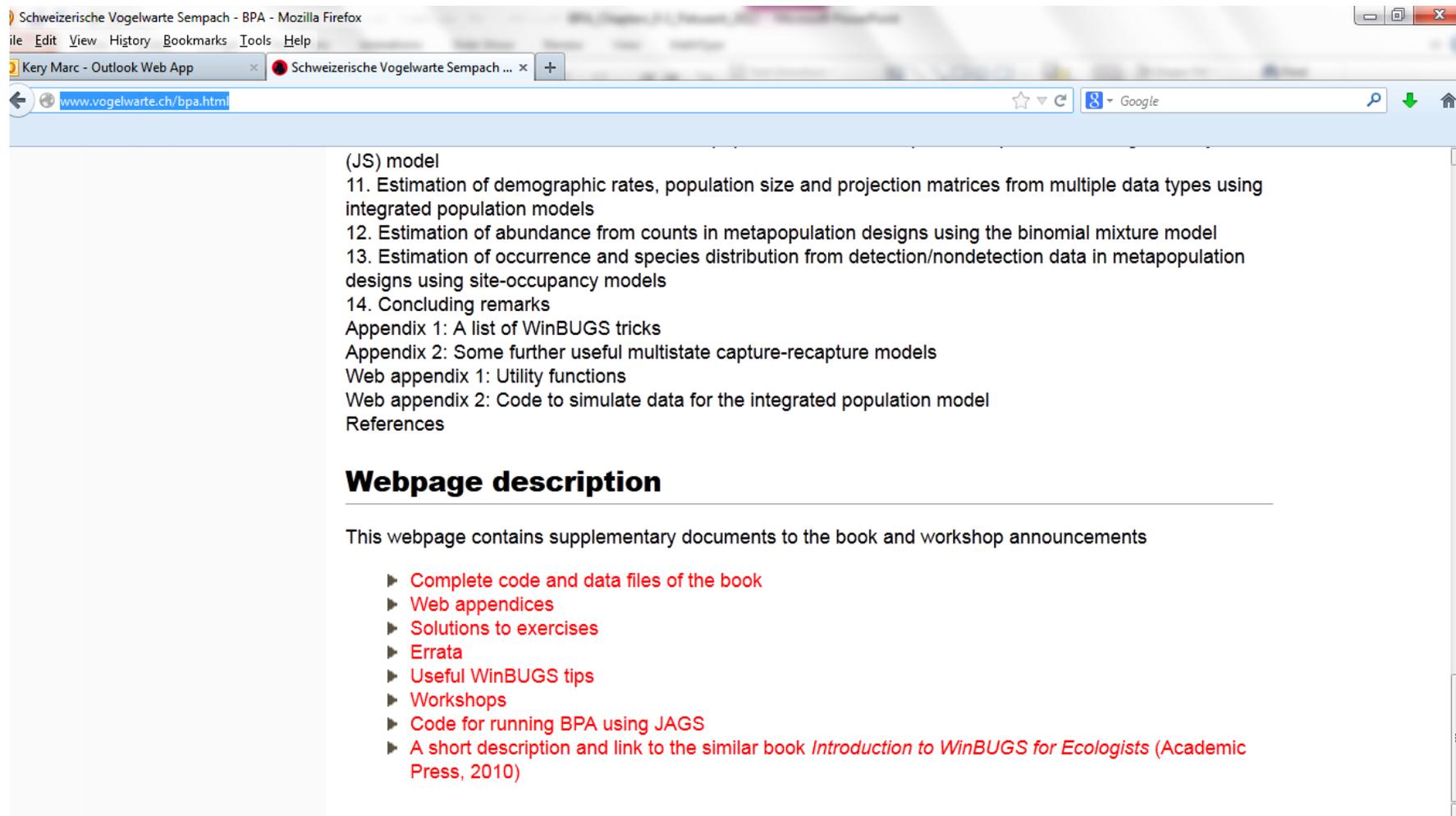
- Course builds a bridge:
 - starts where Kéry (2010) ends
 - leads to more advanced texts on same topic:
 - Royle & Dorazio, AP, 2008
 - King *et al.*, CRC, 2009
 - Link & Barker, AP, 2010
- (Many) Other useful books available
- Prior knowledge of program R required
- Use WinBUGS; JAGS and OpenBUGS should work
- Course documentation (& book & website) contains all code



Preface to workshop, ctd.



Book web site: <http://www.vogelwarte.ch/bpa.html>



Schweizerische Vogelwarte Sempach - BPA - Mozilla Firefox

file Edit View History Bookmarks Tools Help

Kery Marc - Outlook Web App x Schweizerische Vogelwarte Sempach ... x +

www.vogelwarte.ch/bpa.html

(JS) model

11. Estimation of demographic rates, population size and projection matrices from multiple data types using integrated population models
12. Estimation of abundance from counts in metapopulation designs using the binomial mixture model
13. Estimation of occurrence and species distribution from detection/nondetection data in metapopulation designs using site-occupancy models
14. Concluding remarks

Appendix 1: A list of WinBUGS tricks

Appendix 2: Some further useful multistate capture-recapture models

Web appendix 1: Utility functions

Web appendix 2: Code to simulate data for the integrated population model

References

Webpage description

This webpage contains supplementary documents to the book and workshop announcements

- ▶ [Complete code and data files of the book](#)
- ▶ [Web appendices](#)
- ▶ [Solutions to exercises](#)
- ▶ [Errata](#)
- ▶ [Useful WinBUGS tips](#)
- ▶ [Workshops](#)
- ▶ [Code for running BPA using JAGS](#)
- ▶ [A short description and link to the similar book *Introduction to WinBUGS for Ecologists* \(Academic Press, 2010\)](#)

Preface to workshop, ctd.



- Goals of workshop

- Illustrate the Bayesian implementation in WinBUGS and JAGS (via R) of a wide range of useful models for population analysis
- Walk you through gently so you learn by example
- Demystify Bayesian analyses; show how it works in practice
- Demonstrate great value of simulated data sets
- Show importance of four key stats concepts:
 1. linear model,
 2. GLM,
 3. random effects
 4. hierarchical models
- Emphasize explicit hierarchical models

=> **Free the modeller in you !**



Acknowledgements



Workshop

- Courtney Amundson, Erin Roche
- Andy Royle, Jim Hines, Kimmy Gazelski

Book

- Colleagues: Andy Royle, Olivier Gimenez, Bob Dorazio
- Employers: The Vogelwarte



Program



Monday

9.00–10.30	Introduction Population Analyses (Chap. 1)
10.30–11.00	Break
11.00–13.00	Bayesian analysis of a population model (Chap. 2)
13.00–14.00	Lunch Break
14.00–16.00	The GLM as the quintessential statistical model (Chap. 3)
16.00–17.00	Mixed GLMs (Chap. 4)
from 18.30	Group dinner (optional)

Tuesday

9.00–10.30	More mixed GLMs (Chap. 4)
10.30–11.00	Coffee Break
11.00–13.00	State-space models (Chap. 5)
13.00–14.00	Lunch Break
14.00–16.00	Closed population models for abundance estimation (Chap. 6)
16.00–16.30	Coffee break
16.30–18.00	Cormack-Jolly-Seber models (Chap. 7)

Wednesday

9.00–10.30	More Cormack-Jolly-Seber models (Chap. 7)
10.30–11.00	Coffee Break
11.00–13.00	Multistate models (Chap. 9)
13.00–18.00	Lunch Break + Consulting session (incl. Coffee at 16.00 h)



Program



Thursday

9.00–10.30	More multistate models (Chap. 9)
10.30–11.00	Coffee Break
11.00–13.00	Site-occupancy models (Chap. 13)
13.00–14.00	Lunch Break
14.00–15.00	More site-occupancy models (Chap. 13)
15.00–16.00	Jolly-Seber models (Chap. 10)
16.00–16.30	Coffee break
16.30–18.00	More Jolly-Seber models (Chap. 10)

Friday

9.00–10.30	Integrated population models (Chap. 11)
10.30–11.00	Coffee break
11.00–13.00	More integrated population models (Chap. 11)
13.00–14.00	Lunch Break
14.00–17.00	Binomial mixture models (Chap. 12)
17.00	End of workshop



Participants

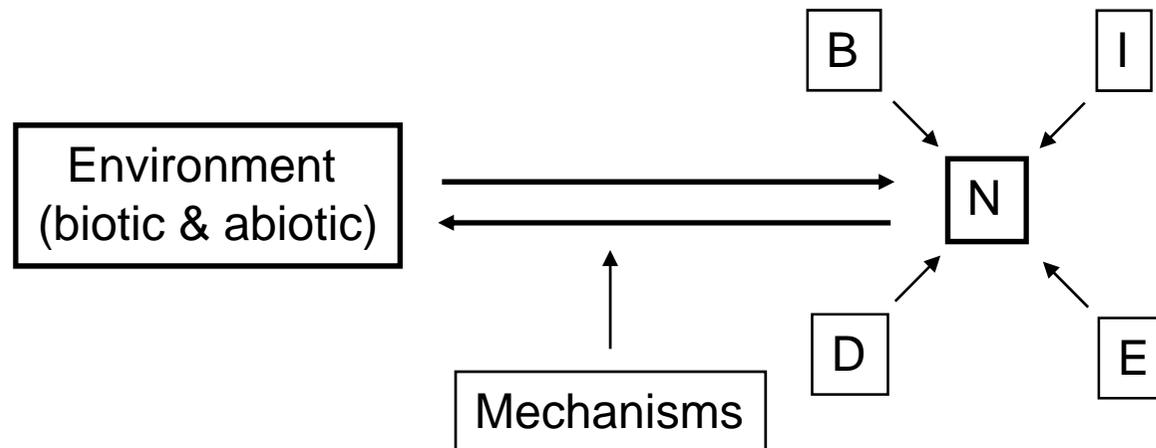




Chapter 1: Population analysis



- Ecology: science of distribution and abundance
- „Population“: key concept; collection of individuals
- Central pillar of ecology: population analysis



Population analysis



- Hierarchical scales of organisation:
genes < individuals < populations < metapopulations,
communities < metacommunities
- Different types of populations
- Only definition of „individual“ varies !

	One site	Multiple sites
One species	„Classic“ population	Metapopulation
Multiple species	Community	Metacommunity



Population analysis



- Analytically almost identical
- Description in terms of state and rate parameters



Population analysis



Demographic analysis of „populations“

Scale of organisation = Type of “population”	Description	
	Static (state variable)	Dynamic (vital rates)
<u>(1) One site, one species:</u> “Classic” population = population of individuals	Abundance N	Survival rate (ϕ), Fecundity rate (γ)
<u>(2) Multiple sites, one species:</u> Metapopulation = population of (local) populations	$N(s)$ $z(s)$ Occupancy $\psi = \Pr(N(s) > 0)$	Extinction rate ($1 - \phi_s$), Colonisation rate (γ_s), Dispersal rates
<u>(3) One site, multiple species:</u> Community = population of species	$N(k)$ $z(k)$ Species richness	Extinction rate ($1 - \phi_k$), Colonisation rate (γ_k), Dispersal rates
<u>(4) Multiple sites, multiple species:</u> Metacommunity = population of multi-species populations	$N(k, s)$ $z(k, s)$ Species richness	Extinction rate ($1 - \phi_{k,s}$), Colonisation rate ($\gamma_{k,s}$), Dispersal rates

Notation: N - abundance; z - presence/absence indicator; k - species index; s - index for site.



Population analysis



BUT: very widely ignored fact about all natural populations:
imperfect detection

- Prevents direct observation of N and rate parameters (latent quantities) measurement error what now ?

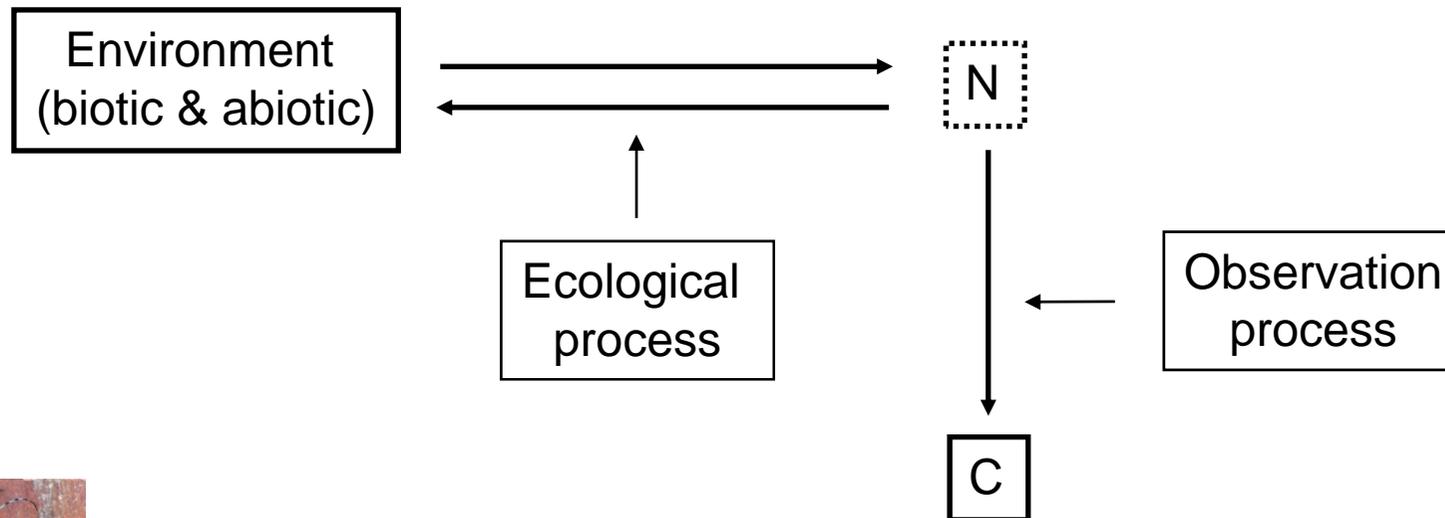


Population analysis



BUT: very widely ignored fact about all natural populations:
imperfect detection

- Prevents direct observation of N and rate parameters (latent quantities) measurement error what now ?
- Hierarchical view of ecology:



Population analysis



- observation process = measurement error process
- 2 components of observation process for counts:
 - „Individuals“ overlooked when present:
 - > false negative errors
 - „Individuals“ misidentified/counted doubly:
 - > false positive errors



Population analysis



Classical way of dealing with observation process:



Population analysis



Classical way of dealing with observation process:

- „Individuals“ overlooked when present:
 - > false negative errors
- „Individuals“ misidentified/counted doubly:
 - > false positive errors



Population analysis



Classical way of dealing with observation process:

- ~~• „Individuals“ overlooked when present:
-> false negative errors~~
- ~~• „Individuals“ misidentified/counted doubly:
-> false positive errors~~
- *„Does not exist“*
- *„Does not affect our study“*
- *„Errors cancel each other out“*



Population analysis



Classical way of dealing with observation process:

- ~~„Individuals“ overlooked when present:
-> false negative errors~~
- ~~„Individuals“ misidentified/counted doubly:
-> false positive errors~~

- *„Does not exist“*
- *„Does not affect our study“*
- *„Errors cancel each other out“*

-- >> Avoid explicit thinking about, or modeling of, observation process, assume away all problems, prefer making fuzzy and implicit assumptions



Population analysis



Our approach to observation process:

- „Individuals“ overlooked when present:
 - > false negative errors
- ~~• „Individuals“ misidentified/counted doubly:
 - > false positive errors~~
- Only assume away false positives (but can deal with them more formally also)
- Eliminate false-positive errors as far as possible:
 - design + analysis
- **Model** false negatives, usually with hierarchical model

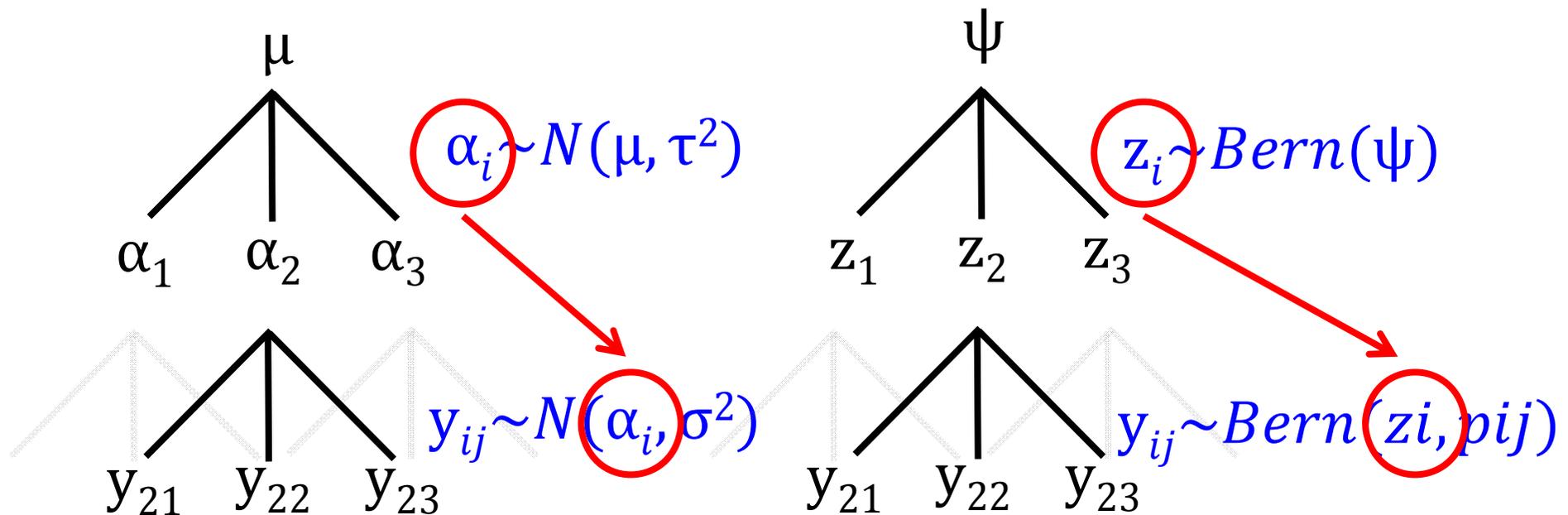


Population analysis



Hierarchical models

- Dependent sequence of random variables (observed & unobserved), e.g., randomised block ANOVA, hierarchical logistic regression

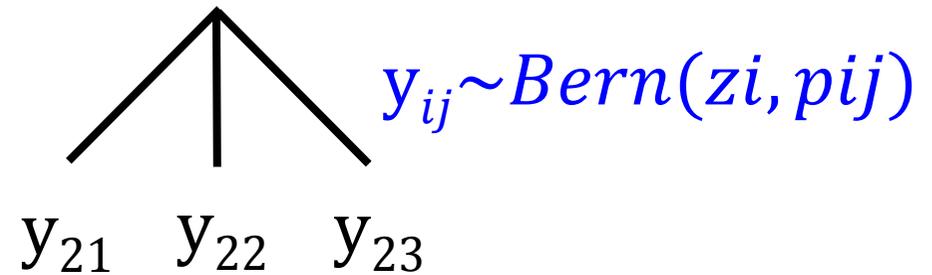
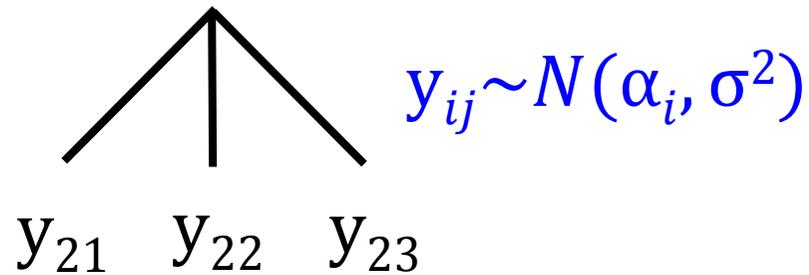


Many other names: state-space, random, mixed-effects, latent variables, unobserved components models

Population analysis



Last level of HM: measurement error = observation process



Qualitatively different:

(1) cancel out

(2) do NOT cancel out



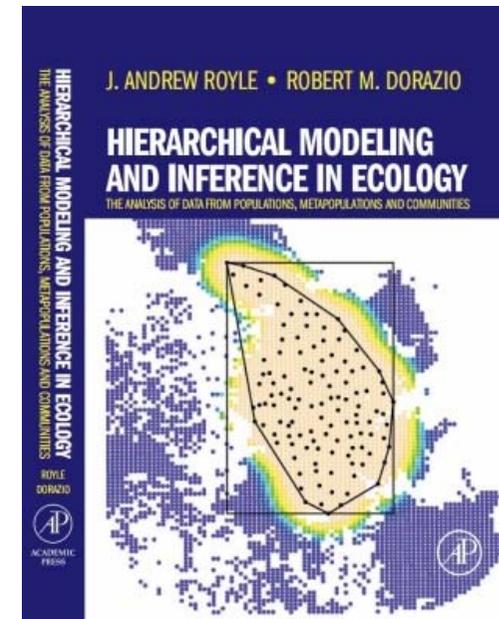
=> Should we care for measurement errors ?

Population analysis



Explicit hierarchical models in ecology:

- e.g., contain 1st submodel for ecological process and 2nd submodel for observation/measurement error process
- model simply reflects hierarchical genesis of ecological data
- parameters with explicit ecological meaning
- foster intellectual clarity in population analysis
- emphasized this week: hierarchical perspective
- see R&D book !



Population analysis



Simple mathematical description of observation process for counts:

- binomial distribution: cf. R code in script
- counts of $N = 16$ sparrows with $p = 0.4$
- features of binomial observation process:
 1. Typical count smaller than N , average $N * p$
 2. Variable counts even under constant conditions
 3. Variation known: $N * p * (1-p)$
 4. May be affected by covariates
- Other descriptions possible: Poisson, beta-binomial, ...



Structure of book/workshop



Quantity modeled	Single site	Multiple sites
Distribution		
Apparent distribution		Logistic regression (ch. 3)
True distribution		Site-occ model (ch. 13)
Abundance		
Apparent abundance	Poisson GLM (ch. 3) Poisson GLMM (ch. 4) State-space model (ch. 5)	~ ~ ~
True abundance	Closed capture-recapture model (ch. 6)	Nmix model (ch. 12)
Vital and other rates		
Survival probability	Cormack-Jolly-Seber model (ch. 7) Ring-recovery model (ch. 8) Multi-state model (ch. 9) Jolly-Seber model (ch. 10)	~ ~ Multi-state model (ch. 9) ~
Fecundity / Recruitment	Poisson GLM (ch. 3) Jolly-Seber model (ch. 10) Integrated pop. model (ch. 11)	Poisson GLM (ch. 3), Poisson GLMM (ch. 4)
Movement rate	-	Multi-state model (ch. 9)
Leslie matrix modeling	Integrated pop. model (ch. 11)	~



Population analysis



The joy of simulating data:

- Lots of advantages !
- Truth is known, i.e., know what solutions should look like
- Check that have not made coding error in BUGS
- Get a grasp for what sampling error is (normally only ever see a single replicate output of the number-generating stochastic machine)
- Check frequentist characteristics of estimates, e.g., bias, precision
- Power analyses, decide on required sample sizes



Population analysis



Joy of simulated data, ctd.

- Check whether params estimable/identifiable
- Check effects of assumption violations
- **and especially:**
 - Prove to yourself that you have understood a model;
build up a data set and then break it down again using
the analysis (data analysis is like fixing a motor-bike ...)



Population analysis



Summary:

- population analysis central in (population) ecology
- variants of „populations“; state and rate variables
- observed data =
 $f(\text{ecological process, observation process})$
- hierarchical model: series of dependent random variables
- explicit hierarchical models (Royle & Dorazio 2008)
- hierarchical models are great
- data simulation is great



Creed for (hierarchical) modeling and BUGS



To make sense of an observation,
everybody needs a model ...
whether he knows it or not.

The most transparent way of describing
a complex stochastic system is by a series of
dependent random variables, i.e., as a hierarchical model.

One of the most transparent ways of building a model
is by describing it in the BUGS language.

~ ~ ~



