International Statistical Ecology Conference 2016

June 28 - July 1, 2016
Seattle, Washington - USA

ABSTRACT BOOKLET

ISEConference.net

International Statistical Ecology Conference 2016

June 28 - July 1, 2016
Seattle, Washington - USA

TABLE OF CONTENTS

TUESDAY, 28 JUNE

Welcome and Opening Session, 0830-1000

CONTRIBUTED A - CONCURRENT SESSIONS, 1030–1200

Fisheries Research 1

Abundance Estimation 1

Movement Ecology 1

Plenary Session, 1300–1350

CONTRIBUTED B - CONCURRENT SESSIONS, 1400–1530

Population Dynamics & Integrated Modelling

Multi-species Modelling

Occupancy Modelling

CONTRIBUTED C - CONCURRENT SESSIONS, 1600–1715

Population Dynamics/Abundance Estimation

Fisheries Research 2

Citizen Science/Big Data/Individual-based Models

Poster Session, 1715–1800

WEDNESDAY, 29 JUNE

Plenary Session, 0830–1000

CONTRIBUTED D - CONCURRENT SESSIONS, 1030–1200

Abundance Estimation 2

Species Distribution & Spatial Modelling

Population Dynamics

CONTRIBUTED E - CONCURRENT SESSIONS, 1300–1400

Big Data

Statistical Theory 1

Abundance Estimation 3

CONTRIBUTED F - CONCURRENT SESSIONS, 1400–1530

Movement Ecology 2

Survey Design

Occupancy Modelling 2

CONTRIBUTED G - CONCURRENT SESSIONS, 1600–1730

Capture-Recapture 1

Spatial Modelling 1

Community Dynamics

THURSDAY, 30 JUNE

CONTRIBUTED & INVITED H - CONCURRENT SESSIONS, 0830–1000

Invited: Ecology, Statistics, and Marine Mammals of the Pacific

Invited: Statistical approaches to the analysis of marine ecosystems

Disease Ecology/Population Genetics/Spatial Ecology

CONTRIBUTED I CONCURRENT SESSIONS, 1030–1200

Capture-Recapture 2

Distance Sampling

Multi-species Modelling/Species Distribution Modelling

FRIDAY, 1 JULY

Plenary Session, 0830–1000

CONTRIBUTED SESSION J CONCURRENT SESSIONS, 1030–1200

Movement Ecology 3

Biodiversity

Movement and Spatial Modelling

CONTRIBUTED K CONCURRENT SESSIONS, 1300–1430

Capture-Recapture 3

Spatial Modelling 2

Statistical Theory 2

CLOSING SESSION

The merging of mathematical ecology and statistics: A retrospective look from the 1970s to present

Ray Hilborn [a]

[a] School of Aquatic and Fishery Sciences
University of Washington
Seattle, Washington, USA

Keywords: mathematical ecology, fisheries, Bayesian statistics.

Abstract: In this talk I will review my personal experiences in the unification of mathematical ecology and statistics in natural resource management. In the 1970s mathematical ecology emerged as a significant force in the fields of fisheries, forestry, pest control and disease. But at that time there was little if any connection between the models used and data. It was not until the 1990s that there was a melding of statistical tools and the models used in these fields. With the declaration of the 200 mile exclusive economic zones in the late 1970s scientists in fisheries needed to provide management advice, and were generally at the forefront of using statistical tools to estimate the parameters of their models. During the 1990s advances in computation enabled Bayesian statistics to emerge as a major force in such analysis.

References

Hilborn, R. and M. Mangel. 1997. The Ecological Detective: confronting models with data. Princeton University Press, Princeton, N.J. 315 pps

Punt, A.E. and R. Hilborn. 1997. Fisheries stock assessment and decision analysis: the Bayesian approach. Reviews in Fish Biology and Fisheries. 7: 35-63

On the Accumulation of Evidence in Programs of Conservation and Science

Jim Nichols
Patuxent Wildlife Research Center, USGS
jnichols@usgs.gov
301-497-5660

Statistical analyses of ecological data are usually focused on one of two kinds of endeavor, conservation and science. In both cases, analyses focus on discrimination of competing hypotheses by assessing the degree to which ecological data correspond to differing model-based predictions. In conservation, hypotheses typically concern effects of management actions on focal systems, whereas scientific hypotheses deal more generally with how a system “works”. In ecology, definitive discrimination of hypotheses is seldom achieved with a single key experiment, analysis, or set of observations. Instead, multiple experiments and observations lead to the accumulation of evidence for, and confidence in, specific hypotheses. Despite this reality, a substantive fraction of ecological research has the appearance of one-off studies. Research funding structures are likely responsible for this phenomenon to some degree, but even well-funded, long-term studies often seem to shift focus among different questions and corresponding sets of hypotheses. This emphasis on one-time studies and analyses has been accompanied by a focus of statistical methodology on discriminating among competing hypotheses based on a single study or data set.

Here, I consider programs of conservation and science in which inferences based on each experiment or analysis are used to update an *information state*, defined as a vector of model “weights” reflecting the relative accumulated evidence corresponding to each model in the focal set. I show examples of the evolution of information state across time, exclusively from programs of conservation that employ adaptive management. It is suggested that the concept of an evolving information state has implications for study design that go beyond the usual considerations employed for single studies. The possibility is considered that devotion of more attention to evolving information states for sets of carefully constructed hypotheses might increase our rate of learning about ecological systems.

What Are We: Statistical Ecologists or Ecological Statisticians?

Bob O'Hara
Senckenberg Biodiversity and Climate Research Centre [BiK-F]
bob.ohara@senckenberg.de
Tel. +49 (0)69 7542 1863

Statistics is ultimately a service subject: it exists to help other people, whether they be ecologists, politicians, or gamblers. But statistics has also developed as a separate discipline. Coming from other direction, ecologists have to analyse their data, and this sucks them in to learning about statistics.

Somewhere in the middle is either ecological statistics or statistical ecology. I will try to work out what this means, and ask how ecology is best served by statistics - to what extent should we be developing new methods, or educating ecologists so that they use old techniques better? And how much of ecological statistics (or statistical ecology) should be ecological modelling? I will pick out some examples from my own work to suggest some different approaches to these problems.

Citizen Science: Trick or Treat?

Byron Morgan
University of Kent
B.J.T.Morgan@kent.ac.uk

Members of the public have been producing data for scientific analysis for decades. Citizen science is used in a wide range of different areas, including mapping the surface of Mars, recording the incidence and abundance of wild animals, and identifying animals that have triggered cameras in Africa. In this talk the focus will be on ecological applications, where different data sets are often collected at different scales. We shall consider whether and how citizen science data, potentially varying in quality, might be combined with information from more stringently designed surveys. As can be seen from the references below, this is a research area of current interest. Illustrations will be provided from studies of British wildlife, in particular British butterfly surveys. This is joint work with Emily Dennis, and colleagues from Butterfly Conservation, Dorset, UK, and the Centre for Ecology and Evolution, Wallingford, UK.

References

Dorazio, R. M. (2014) Accounting for imperfect detection and survey bias in statistical analysis of presence-only data. Global Ecology and Biogeography, 23, 1472-1484.

Fithian, W., Elith, J., Hastie, T. and Keith, D.A. (2015) Bias correction in species distribution models: pooling survey and collection data for multiple species. Methods in Ecology and Evolution, 6, 424-438.

Giraud, C., Calenge, C., Coron, C. and Juillard, R. (2015) Capitalizing on opportunistic data for monitoring relative abundance of species. Biometrics, DOI: 10.1111/biom.12431.

Isaac, N.J.B., van Strien, A. J., August, T.A., de Zeeuw, M. P. and Roy, D. B. (2014) Statistics for citizen science: extracting signals of change from noisy ecological data. Methods in Ecology and Evolution, 5, 1052-1060.

Pagel, J., Anderson, B., O'Hara, R., Cramer, W., Fox, R., Roy, D., Thomas, C. and Schurr, F. (2014) Quantifying range-wide variation in population trends from local abundance data and large-scale presence records. Methods in Ecology and Evolution, 5, 751-760.

Pocock, M.J.O., Roy, H.E., Preston, C. D. and Roy, D. B. (2015) The Biological Records Centre: a pioneer of citizen science, Biological Journal of the Linnean Society, 115, 475-493.

The Past, the Present and the Future of Capture-Recapture

Ruth King
University of Edinburgh
Ruth.King@ed.ac.uk
Web: http://www.maths.ed.ac.uk/~rking33
Tel: 0131-650-5947

Capture-recapture studies and associated statistical analyses have a long history. I will discuss some of this history, focusing on key statistical developments and their impact on the research area. This will lead to a discussion of recent research in this area, before concluding with future challenges.

Disclaimer: this talk will be a biased perspective.

Spatial, temporal and demographic connectivity: How novel network algorithms can help investigating species connectivity in fragmented landscapes

Marie-Josée Fortin
University of Toronto
mariejosee.fortin@utoronto.ca
http://labs.eeb.utoronto.ca/fortin

Given current climate and land use change, species movement in intensively human-modified landscapes is impeded by landscape cover types as well as the distances among remnant habitat patches. It is important to determine to what degree habitat patches scattered throughout the landscape may function as stepping stones facilitating dispersal among otherwise isolated habitat. To this end connectivity metrics can be used; however, such metrics do not account for species abundance and population dynamics that also affect species dispersal and persistence. Hence one needs to model the spatio-temporal dynamic of population dynamics (demography) and dispersal (connectivity). Here I introduce (1) a new generalized network model of habitat connectivity that accounts for the number of dispersing individuals and for long-distance dispersal processes across generations and (2) how to measure spatio-temporal connectivity. I illustrate how this generalized network model can be used to test how stepping-stones are important to promote (i) species range expansion (Black Woodpecker, Spain) and (2) the spread of vector-borne disease (Lyme disease, Ontario) in fragmented landscapes based on wildlife-host movement.

For the Black Woodpecker range expansion example, I show that the loss of intermediate and large stepping-stone habitat patches can cause a sharp decline in the distance that can be traversed by species (critical spatial thresholds) that cannot be effectively compensated by other factors previously regarded as crucial for long-distance dispersal (fat-tailed dispersal kernels, source population size). Then for the Lyme disease spread example, I compare the probability of infected-tick *(I. scapularis)* spread for a suite of hosts (white-footed mice, American robins and white-tailed deer) in a Lyme-endemic island landscape in Thousand Islands National Park. I show that stepping-stone habitat is critical for short- and long-distance invasion of both the tick vector and the pathogen by mice and deer. Last, I show that the anticipating the impact of land use changes on species dispersal can be modelled using novel spatio-temporal network algorithms.

The Case of the Missing Model:
The Modernisation of Multivariate Analysis in Ecology

David Warton
University of New South Wales
david.warton@unsw.edu.au
+61 2 9385 7123

For the best part of four decades, multivariate analysis in ecology has diverged substantially from mainstream statistics, perhaps because state-of-the-art in 1980's statistics was not capable of handling the complexity frequently seen in multivariate abundance data simultaneously collected across many species. But the methods developed in the ecological literature, still widely used today, have some serious shortcomings that suggest they are fast approaching their use-by date.

The statistical literature appears to be "catching up" with ecology, in part through technologies to fit quite flexible hierarchical models capable of accommodating key date structure. There is a significant movement now to reunify multivariate analysis in ecology with modern statistical practices. Some key developments on this front will be reviewed, and some immediate challenges identified.

Contributed Sessions are grouped by topic and in chronological order.

Fitting integrated growth models for bigeye tuna in a Bayesian context

**M.S. Chambers[a,c], L.A. Sidhu[a], B. O'Neill[a], S. McKechnie[b],
J. Hampton[b] and B. Leroy[b]**

[a]School of Physical, Environmental and Mathematical Sciences
University of New South Wales at the Australian Defence Force Academy
Canberra, Australia

[b]Oceanic Fisheries Programme
Secretariat of the Pacific Community
New Caledonia, Noumea

[c]email: mark.stanley.chambers@gmail.com

Keywords: fisheries; growth; integrated models; model selection; data uncertainty.

Abstract: Estimates of growth rates of members of harvested fish populations are needed to calculate optimum harvest rates. While tag recapture data are an important source of growth information for many species, growth models fitted to tagging data have been shown to be biased when individual variability in growth is not accounted for. Estimating individual variability in growth from tagging data is difficult because the ages of the tagged fish are unknown and typically each fish is recaptured at most once. Laslett *et al.* (2002) describe an approach whereby the length-at-tagging and length-at-recapture of each fish are jointly modelled, ages-at-tagging are assumed to be random variables and variability in growth between individuals is accommodated by allowing a unique maximum length for each fish. Although this approach allows approximately unbiased estimation of the growth curve, Laslett *et al.* (2004) demonstrate that plots of length-at-tagging versus maximum likelihood estimates of age-at-capture do not follow the growth curve, even when the fitted data are simulated from the assumed model. Another issue is that multiple sources of uncertainty and hierarchical assumptions are potentially problematic in the application of model selection criteria. We fit a range of integrated growth models to tag-recapture and direct length-age observations of bigeye tuna (*Thunnus obesus*) and highlight issues with model selection. We also contrast the equivalent Bayesian result with the anomaly described by Laslett *et al.* (2004) in a maximum likelihood context.

References

Laslett, G.M., Eveson, J.P. and Polacheck, T. (2002) A flexible maximum likelihood approach for fitting growth curves to tag-recapture data. *Canadian Journal of Fisheries and Aquatic Sciences*, 59:976-986.

Laslett, G.M., Eveson, J.P. and Polacheck, T. (2004) Estimating the age at capture in capture-recapture studies of fish growth. *Australian and New Zealand Journal of Statistics*, 46:59-66.

An integrated state-space stock assessment model for Atlantic Cod (*Gadus morhua*) off the northeast coast of Newfoundland, and short-term prospects for the fishery

N. Cadigan

Centre for Fisheries Ecosystems Research
Marine Institute of Memorial University of Newfoundland
St. John’s, NL, Canada
noel.cadigan@mi.mun.ca

Keywords: censored likelihood, compositional data, integrated likelihood, process error.

Abstract: I present a state-space assessment model for northern cod, off the northeast coast of Canada, that utilizes much of the existing information on the productivity of this stock: offshore trawl survey indices, inshore gillnet indices, inshore acoustic biomass estimates, fishery catch age-composition information, partial fishery landings information, and tagging information. By integrating much of the information in one model I can address problems that are difficult to deal with using data sources individually, such as estimating fishing and natural mortality rates separately and accounting for changes in spatial distribution. I use the model to provide projections of the impacts of various levels of future fishery catches on the continued recovery of this stock. Results indicate that, although the stock size in 2014 was still low relative to levels in the 1980’s, status has improved since 2005 and is projected to continue to improve in 2015-2019 under a range of catch options.

References

Cadigan, N.G. (2015) A state-space stock assessment model for northern cod, including under-reported catches and variable natural mortality rates. *Canadian Journal of Fisheries and Aquatic Sciences* 72.999: 1-13

Models for forecasting sustainable world tuna catches using natural splines with auto-correlated errors

B. Lee [a] and D. McNeil [b]
[a] Faculty of Science and Technology, Prince of Songkla University, Pattani 94000, Thailand
[b] Faculty of Science, Macquarie University, Sydney 2109, Australia

Corresponding presenter: bm_lee2001@yahoo.com

Keywords: world tuna catches, time series forecasting, autoregressive model, natural splines, sustainable tuna fishery

Abstract: Tuna is an important global traded commodity but its stocks are at risk of being overexploited. Fish statistics from the UN's Food and Agriculture Organization (FAO) shows that annual tuna catches have dramatically increased since 1950 while its growth rates have slightly decreased over past two decades (1991-2011). This trend is investigated and used to develop a practical model for forecasting world tuna catches. An approach involves fitting models to tuna catches time series data that focus on both the signal (using a natural quintic spline function) and the noise (using autoregressive model). A quintic spline creates a quadratic curve interpolating a changing point in the rate and volume of tuna catches. However, positions of knots are free parameters and they are chosen based on plausibility of forecasts. This integrated autoregressive-spline model predicts that the world tuna catches will slightly expand to 5.65 million tonnes in 2025 and remain approximately constant thereafter. Such forecasts are important for planning to ensure tuna sustainability. This same method could be applied to fit catches of separate tuna species and catches in separate fishing areas.

References

Havice, Elizabeth. 2010. The structure of tuna access agreements in the Western and Central Pacific Ocean: Lessons for Vessel Day Scheme planning. Marine Policy 34 (2010) 979–987.

McNeil, N., Odton, P., and Ueranantasun, A. 2011. Spline interpolation of demographic data revisited. Songklanakarin Journal of Science and Technology. 33 (1), 117-120.

A Hierarchical Bayesian framework to merge different ecological processes and data sources for modelling the life cycle of nursery dependent fish species

E. Rivot [a], B. Archambault [a], S. Rochette [b], Y. Vermard [c], M. Savina-Rolland [d], O. Le Pape [a]

[a] Agrocampus Ouest, UMR 0985 ESE INRA, Agrocampus Ouest, Ecologie et Santé des Ecosystèmes, Rennes, France
etienne.rivot@agrocampus-ouest.fr
[b] Ifremer, Département Dynamiques de l'Environnement Côtier, Plouzané, France
[c] Ifremer, Ecology and models applied to fisheries resources, Nantes, France
[d] Ifremer, Channel and North Sea Fisheries Department, Boulogne sur Mer, France

Keywords: Spatial Population dynamics; Life cycle models; Fish stock assessment; Connectivity; Solea solea

Abstract: Integrated life cycle models are key tools for an ecosystem approach to fish population dynamics and stock assessment. Modelling fish life cycles still remains challenging as it requires flexible tools to merge different pieces of ecological and biological knowledge and data sources. Hierarchical Bayesian Models (HBMs) offer a comprehensive framework for such a synthesis, with the potential to embed complex demographic models within statistical models for various sources of data.

We present a body of research aiming at modeling the life cycle of a nursery-dependent flatfish species, the common sole in the Eastern English Channel. A HBM is used to combine different processes and sources of data: (i) Outputs of an individual-based Lagrangian model that provided estimates of the dispersion and mortality of eggs and larvae, from spawning grounds to settlement in several coastal nurseries; (2) A habitat suitability model, based on trawl surveys, to estimate juvenile densities and surface areas of suitable juvenile habitat in different coastal nursery sectors; (3) A statistical catch-at-age model for the estimation of the numbers-at-age and the fishing mortality on adults. The framework enables us to disentangle the effects of multiple stress factors on population renewal. Results revealed contrasted estimates of density-dependent mortality of juveniles across nursery grounds and provide a better understanding of the contribution of nursery grounds toward recruitment and population renewal. The model is used to explore how alternative hypotheses on the degree of mixing between three subpopulations affect the dynamics and stock assessment. While assessment obtained when hypothesizing a single spatially homogeneous population in the Channel indicates catches near MSY, considering a metapopulation with three loosely connected subpopulations supported by local nurseries revealed contrasted status, with two subpopulations being exploited above MSY and one below MSY.

Ongoing developments concern the acquisition of new data on the movement of fish (mark-recapture and otoliths) and their integration within the model to improve our knowledge on the connectivity between subpopulation.

References

Rochette, S., Le Pape, O., Vigneau, J., & Rivot, E. (2013). A hierarchical Bayesian model for embedding larval drift and habitat models in integrated life cycles for exploited fish. Ecological Applications, 23(7), 1659–1676.

Archambault, B., Le Pape, O., Baulier, L., Vermard, Y., and Rivot E. (under review). Adult-mediated connectivity affects inferences on population dynamics and stock assessment of harvested nursery dependent fish populations. Fisheries Research.

Identifying drivers of change in size and weight of Atlantic herring populations using gradient boosting regression trees

O. Lyashevska[a*], **C. Harma**[a], **D. Brophy**[a], **C. Minto**[a], **M. Clarke**[b]

[a] Marine and Freshwater Research Centre (MFRC)
Galway-Mayo Institute of Technology (GMIT)
Dublin Road
Galway, Ireland

[b] Fisheries Science Services Department (FSS)
The Marine Institute (MI)
Rinville, Oranmore
Co. Galway, Ireland

*Corresponding author
olga.lyashevska@gmit.ie

Keywords: Celtic Sea herring; population dynamics; abundance; supervised machine–learning, drivers of change

Abstract:

A decline in size and weight of Atlantic herring (*Clupea harengus*) in the Celtic Sea has been observed since the mid-1980's (ICES 2015). Herring are one of the most important pelagic species exploited by fisheries in the northeast Atlantic. Reductions in growth have consequences for stock productivity. The cause of the decline remains largely unexplained but is likely to be driven by the interactive effect of various endogenous (phenotypic plasticity) and exogenous (environmental forcing) factors. In this study we interrogate a long time-series of biological data (individual lengths, weights and ages) obtained from commercial landings from 1959 to 2012. We use gradient boosting regression trees (Freund 1997) to identify important variables underlying changes in growth from various potential drivers (e.g. Atlantic multi-decadal oscillation, sea surface temperature, North Atlantic oscillation, salinity, wind, zooplankton abundance and fishing pressure). This learning algorithm quantifies the influence of the potential drivers of change with lower test error when compared to other supervised learning techniques. The predictor variables importance spectrum allows us to identify the underlying patterns and potential tipping points while resolving the external mechanisms underlying observed changes in size and weight of herring. The outputs of the analysis are of relevance to conservation efforts and sustainable fisheries management which promotes species resistance and resilience.

References

ICES, 2015. Report of the Herring Assessment Working Group South of 62 degrees N (HAWG). ICES Advisory Committee; ICES Headquarter, Copenhagen, Denmark. 10–19 March 2015. ICES CM 2015/ACOM:06. 850 pp.

Freund Y., Schapire R.E., A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting, Journal of Computer and System Sciences, Volume 55, Issue 1, August 1997, Pages 119-139

Spatio-temporal modelling of fishing effort pattern after displacement due to offshore wind developments using INLA

A. Kafas [a,c] , J. B. Illian [b], I. M. Davies [a], and B. E. Scott [c]

[a] Marine Laboratory
Marine Scotland Science
Aberdeen, Scotland, UK
andronikos.kafas@gov.scot, ian.davies@gov.scot

[b] Centre for Research into Ecological & Environmental Modelling
University of St. Andrews
St. Andrews, Scotland, UK
jbi@st-andrews.sc.uk

[c] Institute of Biological and Environmental Sciences
University of Aberdeen
Aberdeen, Scotland, UK
b.e.scott@abdn.ac.uk

Keywords: Integrated Nested Laplace Approximation, Gaussian Markov Random Fields, fisheries, displacement, offshore wind.

Abstract:

Spatial overlap between emerging offshore wind farms and commercial fishing activity may impede access to traditional fishing grounds. Consequently, fishermen may re-allocate (displace) their fishing effort to alternative sea areas with lower profits and/or less reliability in catches. This study develops a hierarchical spatio-temporal model for fishing effort distribution. The model involves a Gaussian Markov Random Field (GMRF) through the Stochastic Partial Differential Equations (SPDE) approach and adopts the Integrated Nested Laplace Approximation (INLA) algorithm. Multi-year spatio-temporal data for Scottish fishing vessels are sourced through the Vessel Monitoring System (VMS) and used in the model. The effect of relevant covariates as well as time and space dependence is modelled in a Bayesian framework. Validation of the model outputs is undertaken with current fishing patterns. The model evaluates and predicts changes in fishing effort and income patterns in space and time in response to new constraints from offshore wind farms.

Estimating species abundance from environmental DNA

T. Chambert [a], C.S. Goldberg [b], D.S. Pilliod [c], H. Doi [d], T. Takahara [e]

[a] Department of Ecosystem Science and Management
Pennsylvania State University
University Park, PA, USA
thierry.chambert@gmail.com

[b] School of the Environment
Washington State University
Pullman, WA, USA
caren.goldberg@wsu.edu

[c] Forest and Rangeland Ecosystem Science Center
U.S. Geological Survey
Boise, ID, USA
dpilliod@usgs.gov

[d] Graduate School of Simulation Studies
University of Hyogo
Kobe, Japan
hideyuki.doi@icloud.com

[e] Faculty of Life and Environmental Science
Shimane University
Matsue, Japan
ttakahara@life.shimane-u.ac.jp

Keywords: abundance, eDNA, survey design

Abstract: Detection of species using environmental DNA (eDNA) is becoming a very promising method for wildlife monitoring, especially for aquatic organisms. So far, eDNA data has been used mainly to estimate binary metrics of species distribution, such as occupancy of a sampling area, but there is growing interest in using eDNA to infer species abundance and density. Previous studies have found significant levels of correlations between eDNA quantity (e.g., via qPCR or ddPCR) in field samples and species abundance at sites where samples were collected, but no statistical model has yet been developed to directly estimate abundance from raw eDNA data. We present a new statistical model that uses both eDNA and observed abundance data from a few sites to infer species abundance at many other sites, where only eDNA data are available. Our approach does not rely on *ad-hoc* abundance estimation based on extrapolation of a correlation coefficient. Instead, abundance is directly estimated in an integrative modeling approach that uses all data at once and accounts for overdispersion, a common issue with eDNA data. To assess if the model provides reliable estimates of abundance, we used a cross validation approach, applied on four different datasets (3 amphibians, 1 fish) for which both eDNA and abundance data were available. Preliminary findings indicate that the model performed well on all datasets, but some uncertainty and imprecision in the method remains. Our initial results suggest a promising future for this approach. We propose future research directions, regarding sampling designs and sample replication, to further improve the reliability of this method.

Modelling pin-point cover data of complementary vegetation classes

C. Damgaard

Bioscience
Aarhus University
Silkeborg, Denmark
cfd@bios.au.dk

Keywords: Joint distribution of plant abundance; plant cover; hierarchical models; plant ecology; point-intercept; structural equation modelling

Abstract: It is often ecologically meaningful to divide the vegetation into a number of complementary vegetation classes or functional types. Here, a method for modelling pin-point plant cover data for such complementary classes is presented. The joint distribution of pin-point cover data of complementary vegetation classes is modelled using a mixture distribution of the multinomial distribution and the Dirichlet distribution, where the Dirichlet distribution is used to model the effect of spatial aggregation.

In order to demonstrate the method, the variation in cover with space or time is modelled using a hierarchical Bayesian approach, where the mean cover of each site at a specific time is modelled by a latent variable.

The statistical modelling procedure is exemplified in a case-study of pin-point cover data of the two dominating species *Calluna vulgaris* and *Deschampsia flexuosa*, and the abundance of the complement species class of all other higher plants on Danish dry heathlands. The cover of *C. vulgaris* increased significantly with annual precipitation and the cover of *D. flexuosa* decreased significantly with annual precipitation. Furthermore, the mean cover of *C. vulgaris* and *D. flexuosa* within-sites was negatively correlated. There were no significant changes in the cover of the three complementary dry heathland vegetation classes from 2007 to 2012.

The presented model allows information of complementarity to be incorporated and whereby increasing the statistical power. Furthermore, the spatial aggregation of the vegetation is modelled so that statistical inference tests will not be deflated due to pseudo-replication.

Open models for removal data

E. Matechou, **R. S. McCrea**, **B. J. T. Morgan**, **D. Nash** and **R. A. Griffiths**

University of Kent

Keywords: common lizards, depletion, great crested newts, RJMCMC, stopover model

Abstract: Individuals of protected species, such as amphibians and reptiles, often need to be removed from sites before development commences. Usually, the population is considered to be closed. All individuals are assumed to i) be present and available for detection at the start of the study period and ii) remain at the site until the end of the study, unless they are detected. However, the assumption of population closure is not always valid. We present new removal models which allow for population renewal through birth and/or immigration, and population depletion through sampling as well as through death/emigration. When appropriate, productivity may be estimated and a Bayesian approach allows the estimation of the probability of total population depletion. We demonstrate the performance of the models using data on common lizards, *Zootoca vivipara*, and great crested newts, *Triturus cristatus*.

Confronting preferential sampling in population surveys

P. B. Conn[a], J. T. Thorson[b], and D. S. Johnson[a]

[a] Marine Mammal Laboratory, Alaska Fisheries Science Center, NOAA National Marine
Fisheries Service, 7600 Sand Point Way NE, Seattle, WA 98115 USA
paul.conn@noaa.gov, devin.johnson@noaa.gov

[b] Fisheries Resource Assessment and Monitoring Division (FRAM), Northwest Fisheries
Science Center, National Marine Fisheries Service (NMFS),
NOAA, 2725 Montlake Boulevard E, Seattle, WA 98112, USA
james.thorson@noaa.gov

Keywords: count data; preferential sampling; spatial autocorrelation; species distribution model.

Abstract: Population surveys are often used to estimate the density, abundance, or distribution of animal populations. Recently, model-based approaches to analysing survey data (i.e. species distribution models) have become popular because one can more readily accommodate departures from pre-planned survey routes and construct more detailed maps than one can with design-based procedures. Model-based analysis often makes use of species-covariate relationships and/or spatially autocorrelated random effects to help predict density or occurrence in unsampled locations. Such species distribution models often make the implicit assumption that locations chosen for sampling and the state variable of interest are conditionally independent given modelled covariates. However, this assumption is likely violated in many cases when survey effort is non-randomized, leading to preferential sampling. We develop a hierarchical statistical modelling framework for detecting and alleviating the biasing effects of preferential sampling in species distribution models. The approach works by jointly modelling animal abundance/occurrence and the locations selected for sampling, and specifying a dependent correlation structure between the two models. Using simulation, we show that under moderate levels of preferential sampling, our modelling approach reduces bias from ≈40\% (under a naive species distribution model) to ≈ 5\%. We also demonstrate our approach using a bearded seal count dataset. In this case, incorporating relevant explanatory covariates also helped to alleviate bias. When animal population surveys are conducted without a well-defined sampling frame, ecologists should recognize the potential biasing effects of preferential sampling. Joint models, such as those described in this paper, can be used to test and correct for such biases.

Extensions of recent models for butterfly abundance

E.B. Dennis [a, b] and B.J.T Morgan [a]

[a] National Centre for Statistical Ecology
School of Mathematics, Statistics and Actuarial Science
University of Kent
Canterbury, UK
E.B.Dennis@kent.ac.uk, B.J.T.Morgan@kent.ac.uk

[b] Butterfly Conservation
Manor Yard, East Lulworth
Wareham, UK

Keywords: abundance; citizen science; concentrated likelihood; dynamic models; generalised abundance indices; phenology; stopover models.

Abstract: Modelling butterfly abundance presents specific challenges, but is important to improve our understanding of population changes and provide biodiversity indicators. It is necessary to account for seasonal fluctuations in butterfly count data due to their multi-stage life cycles. Many species are also multivoltine, where two or more broods of adults emerge each year. Fitting models to data for many species from extensive, long-term monitoring schemes, such as the UK Butterfly Monitoring Scheme, also presents a computational challenge.

A number of recent developments have been made for modelling butterfly monitoring data. Generalised additive models (Dennis et al. 2013) are currently used for reporting abundance trends, but a generalised abundance index proposed by Dennis et al. (2015a) is very efficient in comparison, and provides new parametric descriptions of seasonal variation. Novel dynamic models explicitly describe dependence between broods and years to produce indices and estimated productivities separately for each brood when appropriate (Dennis et al. 2015b).

We describe several new developments and extensions for these models. These include wider application of dynamic models to explore trait-based species' variation, and the extension of dynamic models to describe the arrival of migrant species to the UK, such as Painted Lady *Vanessa cardui*. A combined analysis of transect and larval web counts for Marsh Fritillary *Euphydryas aurinia* is also explored. We finish with a discussion of potential future developments. We describe work undertaken in collaboration with colleagues from Butterfly Conservation, UK, and the Centre for Ecology and Hydrology, UK.

References

Dennis, E.B., Freeman, S.N., Brereton, T. and Roy, D.B. (2013) Indexing butterfly abundance whilst accounting for missing counts and variability in seasonal pattern. *Methods in Ecology and Evolution*, **4**, 637-645.

Dennis, E.B., Morgan, B.J.T., Freeman, S.N., Brereton, T. and Roy, D.B. (2015a) A generalised abundance index for seasonal invertebrates. Revision submitted to *Biometrics*.

Dennis, E.B., Morgan, B.J.T, Freeman, S.N., Roy, D.B. and Brereton, T. (2015b) Dynamic models for longitudinal butterfly data. *Journal of Agricultural, Biological and Environmental Statistics*, DOI: 10.1007/s13253-015-0216-3.

Untangling Species Distribution and Abundance Patterns from Ordinal Data

K.M. Irvine[a] and **T.J. Rodhouse**[b]

[a]Northern Rocky Mountain Science Center
US Geological Survey
Bozeman, Montana USA
kirvine@usgs.gov

[b]Upper Columbia Basin Network
National Park Service
Bend, Oregon USA
Tom_Rodhouse@nps.gov

Keywords: Beta regression; Grouped continuous; Hurdle Model; Midpoint regression; Non-proportional odds

Abstract: Ecology is often summarily described as being the study of the distribution and abundance of organisms (Andrewartha and Birch 1954). However, a common complication in ecological data analysis is an over-abundance of zeros (Martin et al. 2005; Wenger and Freeman 2008). Zero-inflated and zero-augmented models are available for binary, count, and continuous data (Martin et al. 2005). However, there is a lack of methodology and applications for ordinal data that arise when observations of ecological phenomena are discretized into pre-defined classes, e.g., percent cover. We introduce a two-part or hurdle model framework for the case of ordinal responses with a true zero indicating, in our case, absence of a species as the lowest category (an anchor). This approach deepens ecological inferences by providing the capacity to explore whether distribution is governed by potentially different ecological processes from those influencing abundance. We develop an ordinal regression that directly models ordinal category probabilities by way of a biologically realistic latent variable. This allows for statements in terms of an effect of an explanatory variable of interest on the mean of the latent variable. Our model is applicable to any ordered categorical data with a meaningful zero category and non-zero categories that arise from discretizing or grouping a continuous variable measured on any interval. We motivate our work with empirical plant species abundances in the sagebrush steppe biome in western North America.

References

Andrewartha, H. G., and Birch, L. C. (1954) The distribution and abundance of animals. University of Chicago Press, Chicago, Illinois, USA.

Martin, T. G., Wintle, B. A., Rhodes, J. R., Kuhnert, P. M., Field, S. A., Low-Choy, S. J., Tyre, A. J., and Possingham, H. P. (2005) Zero tolerance ecology: improving ecological inference by modeling the source of zero observations. *Ecology Letters* **8,** 1235–1246.

Wenger, S. J. and Freeman, M. C. (2008) Estimating species occurrence, abundance, and detection probability using zero-inflated distributions. *Ecology* **89,** 2953–2959.

Using hidden Markov models (HMMs) to analyse whale call rates

Valentin Popov[a] and **Roland Langrock**[b] and **Stacy L. DeRuiter**[c] and **Fleur Visser**[d]

[a]School of Mathematics and Statistics
University of St Andrews
The Observatory, Buchanan Gardens, St Andrews, Scotland
vmp@st-andrews.ac.uk

[b]Department of Business Administration and Economics
Bielefeld University
Postfach 10 01 31, 33501 Bielefeld, Germany
roland.langrock@uni-bielefeld.de

[c]Mathematics and Statistics Department
Calvin College
3201 Burton SE Grand Rapids, MI 49546, USA
sld33@calvin.edu

[d]Faculty of Science
Leiden University
Sylviusweg 72, 2333 BE Leiden, Netherlands
f.visser.4@umail.leidenuniv.nl

Keywords: animal communication, pilot whale, call, hidden Markov models

Abstract: Vocalisations by cetaceans form a key component of their social interactions, with vocalisation rates influenced by latent motivational states. Modeling these unobservable states together with the observable call counts poses a challenging problem, which can be successfully solved using hidden Markov models (HMMs). We applied HMMs to call production data from long-finned pilot whales (Globicephala melas) from 7 independent data sets gathered over 3 years in the Vestfjord basin of Lofoten, Norway. Baseline models motivated the use of 3 states to describe calling behaviour, while more complex models were fit to study the influence of behavioural covariates on the between-state transition probabilities. The best model included as a covariate a synthetic variable, formed by compacting information from a cohesion cluster of several variables related to social group cohesion using Multiple Component Analysis (MCA). Model results indicate that pilot whale call production rates vary widely; the influence of group cohesion on call production rates explains some of this variation, with more tightly associated groups producing fewer calls.

Bayesian inference for continuous time movement based on steps and turns

A. Parton[a] and **P.G. Blackwell**[b]

[a]School of Mathematics and Statistics
University of Sheffield
Sheffield, England
AParton2@sheffield.ac.uk

[b]School of Mathematics and Statistics
University of Sheffield
Sheffield, England
P.Blackwell@sheffield.ac.uk

Keywords: movement modelling; continuous-time; observation error; correlated random walk; Ornstein-Uhlenbeck process; Bayesian statistics; GPS data; Markov Chain Monte Carlo

Abstract: Although animal locations gained via GPS, etc. are typically observed on a discrete time scale, movement models formulated in continuous time are preferable in order to avoid the struggles experienced in discrete time when faced with irregular observations or the prospect of comparing analyses on different time scales. A class of models can be defined by representing movement as a combination of stochastic processes describing both speed and bearing. Such models are able to emulate the range of movement ideas described by a variety of popular discrete time approaches based on step lengths and turning angles, such as that of Morales et al. (2004).

This talk will outline ongoing work in methods for Bayesian inference for such models using a Markov chain Monte Carlo approach. Such inference relies on an augmentation of the animal's locations in discrete time that have been observed with error, with a more detailed movement path gained via simulation techniques. The incorporation of behavioural state switching, with heterogeneity both in space and time is then possible using the method described in Blackwell et al. (2015).

Real data on an individual reindeer *Rangifer tarandus* (courtesy of Anna Skarin, Swedish University of Agricultural Sciences, Uppsala) will be used to illustrate the presented methods.

References

Blackwell, P.G., Niu, M., Lambert, C. and LaPoint, S.D. (2015) Exact Bayesian inference for animal movement in continuous time. *Methods in Ecology and Evolution*, doi: 10.1111/2041-210X.12460.

Morales, J.M., Haydon, D.T., Frair, J., Holsinger, K.E. and Fryxell, J.M. (2004) Extracting more out of relocation data: Building movement models as mixtures of random walks. *Ecology*, 85(9):2436–45.

Simplifying and visualizing large-scale movement patterns using the Latent Dirichlet Allocation model

D. Valle [a], R. Fletcher [b], E. Robertson [b], B. Reichert [b], S. Cvetojevic [a], H. Hochmair [a]

[a] School of Forest Resources and Conservation
University of Florida
Gainesville, FL, United States

[b] Wildlife Ecology and Conservation
University of Florida
Gainesville, FL, United States

Keywords: latent dirichlet allocation; mixed membership model; clustering; movement; endangered species; tourism; Florida.

Abstract:

Background / question / methods:

Understanding movement patterns of individuals is critical in multiple disciplines but paradoxically most of the methods to analyze movement data focus on locations, not individuals. Here we describe a novel statistical method that clusters individuals with similar movement patterns, based on the Latent Dirichlet Allocation model. This method enables the identification of sites that are used by multiple groups (i.e., hubs), in sharp contrast with most of the existing network analysis methods, where locations can only belong to a single group. These hubs are likely to play a critical role in several contexts. For instance, these hubs may be important to ensure population connectivity for an endangered species. Furthermore, we show how our method can be used to analyze social media data, generating important insights regarding origin of tourists and their seasonal movement patterns.

Results / Conclusion:

Using simulated data, we show that our method reliably recovers the underlying groups of individuals, regardless of the presence or absence of mixed-membership locations, whereas the more traditional clustering algorithms used in network analysis struggle to determine the relationship between these locations. Our first case study utilizes long-term data from an endangered raptor, the snail kite, and show how this bird species has radically changed landscape use through time. We find that older individuals rely predominantly on southern lakes in Florida while younger individuals consistently use northern lakes. Furthermore, despite heavy reliance on southern lakes, our analysis reveals that older individuals exhibit a shift towards greater utilization of northern lakes in recent years. We attribute this change in landscape use to the exotic snail invasion of several northern lakes beginning in 2005, which substantially enhanced resources for snail kites. Our second case study illustrates the wealth of information that our algorithm can generate from social media data. In particular, we analyze Twitter data from people that visited Florida in 2014. We find a large two-way flow of people from Florida to Brazil, highlighting the historically strong links between these

locations. Both international and domestic tourists tended to visit Florida more often during the winter than the summer. Finally, the major hub in Florida was Orange county (i.e., Orlando), being highly visited by international, national and local tourists, highlighting the pivotal role of this county for the tourism industry in Florida. We believe that the proposed method is likely to be an important addition to the tool kit of researchers interested in understanding movement patterns.

Contact information

D. Valle: drvalle@ufl.edu

moveHMM: an R package for the statistical analysis of animal movement

T. Michelot[a], **R. Langrock**[b], **T. Patterson**[c]

[a]University of Sheffield, UK
tmichelot1@sheffield.ac.uk

[b]University of Bielefeld, Germany
roland.langrock@uni-bielefeld.de

[c]CSIRO, Australia
toby.patterson@csiro.au

Keywords: behavioural state; maximum likelihood; random walk; state-switching model; telemetry data

Abstract: Recent years have seen an explosion in the amount of movement data being collected. It has led to a huge demand for statistical tools that allow ecologists to draw meaningful inference from large tracking data sets. The class of hidden Markov models (HMMs) matches the intuitive understanding that animal movement is driven by switches in underlying behavioural modes and has proven to be very useful for analyzing movement data (Langrock et al., 2012; McKellar et al., 2015). These models are often sufficiently flexible to capture the complex correlation structure found in movement data, yet are computationally inexpensive compared to their competitors. We describe the R package moveHMM (Michelot et al., 2015), which implements HMMs for applications to animal movement. The package allows ecologists to fit an HMM to movement data, in particular allowing for the incorporation of environmental covariates, and includes assessment and visualization tools for the fitted model. We illustrate the use of moveHMM using (simulated) movement of the legendary wild haggis. Our findings illustrate the role our software, and movement modelling in general, can play in conservation and management by illuminating habitat selection constraints.

References

- Langrock, R., King, R., Matthiopoulos, J., Thomas, L., Fortin, D., and Morales, J. M. (2012). Flexible and practical modeling of animal telemetry data: hidden Markov models and extensions. *Ecology*, 93(11):2336–2342.
- McKellar, A., Langrock, R., Walters, J., and Kesler, D. (2015). Using mixed hidden Markov models to examine behavioral states in a cooperatively breeding bird. *Behavioral Ecology*, 26(1):148–157.
- Michelot, T., Langrock, R., Patterson, T., and Rexstad, E. (2015). *moveHMM: Animal Movement Modelling using Hidden Markov Models.* R package version 1.1.

Modelling animal accelerometer data using hidden Markov models

V. Leos-Barajas[a], **R. Langrock**[b],**T. Photopoulou**[c],**T.A. Patterson**[d],**Y.P. Papastamatiou**[e]

[a]Department of Statistics
Iowa State University
Ames, IA, USA
vianey@iastate.edu

[b]Department of Business Administration and Economics
Bielefeld University
Bielefeld, Germany
roland.langrock@uni-bielefeld.de

[c]Centre for Statistics in Ecology, Environment and Conservation
Department of Statistical Sciences
University of Cape Town
Cape Town, South Africa
theoni.photopoulou@uct.ac.za

[d]CSIRO, Oceans and Atmosphere
Hobart, Tasmania, Australia
Toby.Patterson@csiro.au

[e]Department of Biological Sciences
Florida International University
North Miami, Fl, USA
ypapastamatiou@gmail.com

Keywords: animal behaviour; time series; machine learning

Abstract: Use of accelerometers is now widespread within animal biotelemetry as they provide a means of measuring an animal's activity in a meaningful and quantitative way where direct observation is not possible. In sequential acceleration data there is a natural dependence between observations of movement or behaviour – once initiated, particular animal behaviours often last for periods longer than the sampling frequency. This fact has been largely ignored in most analyses of acceleration data or is treated as an afterthought. The studies where serial dependence has been explicitly modelled have largely relied on hidden Markov models (HMMs). HMMs can be implemented in both a supervised or unsupervised learning context with data availability and aim dictating which approach to take. A supervised learning approach can be taken when the objective is to do classification and there is a set of pre-defined behaviours of interest, such that the model's ability to correctly predict and categorize behaviours is of main interest. Alternatively, if the objective is to extract more general information on animal behaviour and make corresponding inference, then the unsupervised approach provides an excellent framework. In this talk I will cover key differences in assessment and implementation of HMMs in both contexts, as well as what we can learn under either approach.

An empirical comparison of models for the phenology of bird migration

A. Lindén [a,b], K. Meller [c] and J. Knape [d]

[a] Novia University of Applied Sciences
Raseborgsvägen 9, FI-10600, Ekenäs, Finland
andreas.linden@iki.fi

[b] Center for Ecological and Evolutionary Synthesis (CEES),
Department of Biosciences, University of Oslo,
P.O. Box 1066 Blindern, NO-0316 Oslo, Norway

[c] The Helsinki Lab of Ornithology (HelLO)
Finnish Museum of Natural History
P.O. Box 17. FI-00014 University of Helsinki, Finland
kalle.meller@helsinki.fi

[d] Department of Ecology,
Swedish University of Agricultural Sciences,
P.O. Box 7044, SE-75007 Uppsala, Sweden
jonas.knape@slu.se

Keywords: citizen sciences, movement ecology, phenology

Abstract: Bird migration phenology shows strong responses to weather and climate change. Trends and patterns in phenology are typically described based on summarizing metrics, but modelling the whole phenological distribution throughout the season may be more relevant when studying trophic interactions or evolutionary change. Here we use count data of migration passage from a Finnish bird observatory to compare different statistical modelling approaches for the whole phenological distributions of spring migration (27 species) and autumn migration (57 species). We assess model parsimony and goodness-of-fit in a set of models, with various phenological functions of different complexity, optionally with covariates accounting for day-to-day variability. The covariates describe migration intensities of related species, or relative migration intensities the previous day (autocovariates). We found that parametric models are often preferred over the more flexible GAMs with constrained degrees of freedom, perhaps because of their better mechanistic basis. Models corresponding to the passage of two distinct populations were frequently preferred over simpler ones, perhaps due to age or sex specific migration phenology. Usually no more complex models were needed. Slightly more complex models were favoured in spring compared to autumn. Related species' migration activity effectively improves model fit by accounting for part of the large day-to-day variation. Autocovariates did not prove particularly helpful for improving model fit, although there were existing effects. Hence, autocorrelation is generally not a major concern as long as the phenological trends are modelled properly. We suggest that parametric models are relatively good for studying single-population migration phenology, especially if daily variation in migration intensity can be controlled for. GAMs may be more useful when migration patterns are irregular and complex. The choise of model should be done case-by-case considering quesitons studied as well as the character of the data.

What Are We: Statistical Ecologists or Ecological Statisticians?

Bob O'Hara
Senckenberg Biodiversity and Climate Research Centre [BiK-F]
bob.ohara@senckenberg.de
Tel. +49 (0)69 7542 1863

Statistics is ultimately a service subject: it exists to help other people, whether they be ecologists, politicians, or gamblers. But statistics has also developed as a separate discipline. Coming from other direction, ecologists have to analyse their data, and this sucks them in to learning about statistics.

Somewhere in the middle is either ecological statistics or statistical ecology. I will try to work out what this means, and ask how ecology is best served by statistics - to what extent should we be developing new methods, or educating ecologists so that they use old techniques better? And how much of ecological statistics (or statistical ecology) should be ecological modelling? I will pick out some examples from my own work to suggest some different approaches to these problems.

Integrated population modeling of Chinook salmon (*Oncorhynchus tshawytscha*) migration on the West Coast of Vancouver Island

A. Béliveau[a], R. Dunlop[b] and C. Schwarz[a].

[a] Department of Statistics and Actuarial Science
Simon Fraser University
Burnaby, Canada
abelivea@sfu.ca
cschwarz@stat.sfu.ca

[b] Fisheries Department
Nuu-chah-nulth Tribal Council
Gold River, Canada
Roger.Dunlop@nuuchahnulth.org

Keywords: Capture-recapture; Integrated population modelling; *Oncorhynchus tshawytscha*.

Abstract: We develop a Bayesian integrated population model to study the return of Chinook salmon (Oncorhynchus tshawytscha) from the ocean to the spawning grounds in Burman River, on the west coast of Vancouver Island, Canada. Chinook salmon on the west coast of Vancouver Island return to their natal stream in the fall after reaching maturity to spawn and die. When entering Burman River, fish stop for at least some time at a stopover pool, where a capture-recapture survey takes place, then move upstream where they spawn and die. The upstream portion of the river is surveyed periodically by snorkelers that count the number of marked and the total number of fish seen (alive). Carcass surveys also take place periodically, during which marked and unmarked carcasses are picked. Our integrated population model integrates the capture-recapture data, carcass data and snorkel data all in a single analysis without rely on an independence assumption. This is, to our knowledge, the first use of explicit integrated population modeling applied to salmon migration. Our explicit integrated population model uses latent variables to follow explicitly the movement and state of fish throughout the migration. In this work, we also implement a Bayesian version of the Jolly-Seber model (Schwarz and Arnason, 1996) to the capture-recapture data alone and compare estimates between the integrated method and the Jolly-Seber method.

References

Schwarz C.J. and Arnason A.N. (1996). A General Methodology for the Analysis of Capture-Recapture Experiments in Open Populations. *Biometrics*, 52:860-873.

When Models Get Too Large:
Determining Estimability in the Stationary Gompertz and Gompertz State Space Density Dependence Models with a Covariate

C. Leatherman [a] , B. Dennis [b] , M. Wiest [c] , and R. Long [d]

[a] Department of Statistical Science
University of Idaho
Moscow, Idaho
caral@uidaho.edu

[b] Department of Fish and Wildlife Sciences
University of Idaho
Moscow, Idaho
brian@uidaho.edu

[c] Department of Statistical Science
University of Idaho
Moscow, Idaho
mwiest@uidaho.edu

[d] Department of Fish and Wildlife Sciences
University of Idaho
Moscow, Idaho
ralong@uidaho.edu

Keywords: estimability; Gompertz; Gompertz State Space; density dependence; covariates.

Abstract: We find that the limits of estimability of the Stochastic Gompertz (Gompertz) and Gompertz State Space (GSS) models of density dependence occur with many fewer parameters than commonly occur in current useage. When adding a single covariate to both models, simulation trials and maximum likelihood estimates of the parameter values show us that while the Gompertz model performs relatively well in most situations, when sample size is low (n=10) the addition of the covariate creates very imprecise estimates. In nearly all cases, the GSS model with a covariate has extremely high estimate variance, estimates often covering the entire range of values that the parameter of interest is constrained to. These results represent severe limitations to the use of covariates with the Gompertz and GSS model and do not bode well for larger state space and other hierarchical models used in modern statistics.

References

Dennis, B. et al. (2006) Estimating Density Dependence, Process Noise, and Observation Error. *Ecological Monographs*, 76(3):323-341.

Joint Inference from Multiple Data Sources: Estimating the Source Stock of Dolphins Stranded After the Deepwater Horizon Oil Spill

L. Thomas [a], C.G. Booth [b], A.A. Hohn [c], P.E. Rosel [d], J. Litz [e], L. Schwacke [f].

[a] Centre for Research into Ecological and Environmental Modelling, University of St
Andrews, The Observatory, St. Andrews, KY16 9LZ UK.
len.thomas@st-andrews.ac.uk

[b] SMRU Consulting, New Technology Centre, St Andrews, KY16 9SS UK.
cgb@smruconsulting.com

[c] NOAA National Marine Fisheries Service, Southeast Fisheries Science Center, 101 Pivers
Island Road, Beaufort, North Carolina 28516. USA.
aleta.hohn@noaa.gov

[d] NOAA National Marine Fisheries Service, Southeast Fisheries Science Center, 646
Cajundome Boulevard, Lafayette, Louisiana 70506. USA.
patricia.rosel@noaa.gov

[e] NOAA National Marine Fisheries Service, Southeast Fisheries Science Center, 75 Virginia
Beach Drive Miami, FL 33149. USA.
Jenny.Litz@noaa.gov

[f] NOAA National Centers for Coastal Ocean Science, 331 Fort Johnson Road, Charleston,
South Carolina, USA 29412. USA.
Lori.Schwacke@noaa.gov

Keywords: Bayesian hierarchical model, data fusion, precision weighting, Deepwater Horizon oil spill programmatic damage assessment.

Abstract: Imagine we wish to make inferences about a population parameter and that different types of relevant measurements have been made on overlapping subsets of the individuals within the population, with each data source being analyzed using a separate method. We present an approach for making joint inference that is designed to impose the minimum of assumptions. Uncertainty from each data source is quantified using a nonparametric bootstrap, mixture models are fit to the bootstrap distributions and these are used to estimate the Effective Sample Sizes (ESSs) associated with each measurement. The combined estimate is then an ESS-weighted average of the distributions from each data source, with an additional component accounting for extrapolation to the individuals for which there are no measurements. This method is applied to common bottlenose dolphin strandings data as part of the Deepwater Horizon oil spill damage assessment (NOAA 2015). The inference goal is to estimate the proportion of reported strandings that come from the coastal vs estuarine populations. Up to two data sources were available on stranded individuals: genetic assignment and stable isotope. We conclude that the stock composition of reported strandings varies by region, but is mostly estuarine animals. We show where this information fits into the overall dolphin damage assessment.

References

NOAA (2015) Deepwater Horizon Oil Spill Draft Programmatic Damage Assessment and Restoration Plan and Draft Programmatic Environmental Impact Statement. http://www.gulfspillrestoration.noaa.gov/restoration-planning/gulf-plan/

Bayesian hierarchical modeling: application to the babassu palm tree population dynamics in Brazil

N. Sirakov[a], **P. Loisel**[a], **B. Fontez**[a], **D. Mitja**[b], **T. Libourel**[b]
A. M. d. Santos[c] and **I. Souza Miranda**[c]

[a]UMR 0729 MISTEA
SupAgro - INRA
2 place Viala 34060 Montpellier, France
nikolay.sirakov@supagro.inra.fr
patrice.loisel@supagro.inra.fr
benedicte.fontez@supagro.fr

[b]UMR 228 ESPACE-DEV
IRD, Université de Montpellier
500, Rue Jean François Breton 34093 Montpellier, France
danielle.mitja@ird.fr
therese.libourel@univ-montp2.fr

[c]Instituto Socio Ambiental e de Recursos Hídricos, Programa de Pós-Graduação em Ciências Florestais
Universidade Federal Rural da Amazônia (UFRA)
CP. 917, CEP 66077-530, Belém, Pará, Brasil
alessiomsag@gmail.com
izildinha.miranda@ufra.edu.br

Keywords: population dynamics; babassu palm tree; bayesian modeling

Abstract: The babassu palm tree (*Attalea speciosa* Mart. ex Spreng.) is an endemic species of the amazonian forests. For decades, the progress of the pioneer front highlights this palm tree in the anthropogenic open areas: pastures and cultivated fields. The babassu is one of the "extractive" resources in Brazil: gathering activity followed by marketing of non-timber products. This activity involves people among the most disadvantaged in the country. Despite this fact, the knowledge of long term functioning of the palm tree is sorely lacking. Our goal is to understand its population dynamics in the pastures and to guide local populations towards a sustainable management of the species. Our modelization process is based on field data series collected between 2013 and 2015. This field research work was conducted *in situ* in the community of Benfica, state of Para, Brazil. First we selected transects inside pastures randomly, then we counted systematically all the individuals inside those transects. We had analyzed the babassu life cycle and we decided to modelize each one of the three major biological processes: mortality, growth and recruitment. We actually develop a multinomial hierarchical model approach which relies on aggregated data: sum of the individuals per stage and per transect. In this talk we present a Bayesian modeling framework analysis with MCMC algorithms for estimate mortality, growth and recruitment rates. Mortality and growth rates are estimated for all six biological stages. We detect and highlight a critical transition between stages 3 and 4. Biologicaly this transition corresponds to the transfer of the bud from the soil to the surface. The vulnerability of babassu thus increases enormously. Our results also underline differences between transects - important location variability.

Different sex-structure in individual quality drive the dynamics of a population of hoopoe

F. Plard[a], R. Arlettaz[b], and M. Schaub[a]

[a] Swiss Ornithological Institute, CH–6204 Sempach, Switzerland
floriane.plard@vogelwarte.ch

[b] Division of Conservation Biology, Institute of Ecology and Evolution,
University of Bern, Bern, Switzerland
raphael.arlettaz@iee.unibe.ch

Keywords: bill length, body condition, individual heterogeneity, integral projection model, population dynamics

Abstract: Models of population dynamics often describe the successive population sizes based on average reproductive and survival rates without taking account of differences at the individual level. However, population dynamics is directly linked to individual trajectories and interactions among. Thus, this is of main importance to understand how much each individual or groups of individuals contribute to shape population dynamics. Here, we present a case study on a population of hoopoe (*Upupa epops*) where we show how the interactions between males and females individual quality- and condition-structures influence demographic rates and population dynamics. We used an Integral projection model that has been developed to include individual sex-, condition- and -quality to understand how these structures and their interaction influenced the demographic rates and the population dynamics. We have addressed four questions: i) We have analyzed the influence of individual characteristics on survival and reproductive rates to disentangle the influence of individual condition and quality on demographic rates in each sex. ii) We have built an IPM to show how the distribution of the two sexes interacted to influence the demographic rates iii) we have estimated individual fitness according to individual quality and condition to measure variability in individual contribution to the population. iv) We have conducted a sensitivity analysis to understand how and how much each sex distribution influenced population growth rate. The sex ratio and the quality of both sexes had main consequences on population growth rate. However males and females qualities influenced differently population dynamics as female quality was directly linked to annual reproductive success whereas male quality influenced annual survival. Males and females seemed thus to display different strategies in the population. If both sexes were limiting at the population level, at the individual level, males can contribute more to the population than females as high quality male reached higher fitness than high quality females. This study showed how including male and female individual trait dynamics in population model would ameliorate our understanding of both population dynamics and the evolution of sex specific strategies as they are directly related.

Detecting the effects of environmental covariates in coarsely surveyed population time series

Jake M. Ferguson

Center for Modeling Complex Interactions
University of Idaho
Moscow, Idaho
troutinthemilk@gmail.com

Keywords: Time series; interaction; inference.

Abstract: Inferring the effects of environmental variables on animal population growth has proven to be an essential approach for understanding the factors that determine population numbers. Here we show that the ability of standard time series methods to detect these interactions depends on a timescale defined by the underlying vital rate. When an interaction occurs on timescales that are shorter than the census period, standard time series methods may not be able to detect environmental interactions. We develop a strategy for building covariates from data measured at high frequencies that can be used when the timescale of the underlying process is short. This approach summarizes highly resolved environmental variables over the characteristic timescale of the process, rather than the census period. The approach is tested using simulated population time series data. Our results indicate that it can reliably detect population-environment interactions. However, simulations show coarsely censused population abundances are not effective at detecting niche responses to environmental variables under natural conditions. One implication of these results are that the low of predictability of population dynamics by weather and climate variables found in previous studies may be in part due to how predictor variables are constructed.

Predicting Fish Biomass in a Tropical Estuarine Habitat

S. Chesoh [a], C. Lueangthuwapranit [a] & D. McNeil [b]

[a] Faculty of Science and Technology, Prince of Songkla University, Pattani 94000, Thailand
[b] Faculty of Science, Macquarie University, Sydney 2109, Australia

Corresponding presenter: chesoh.s@hotmail.com

Keywords: Fish standing crop, factor analysis, statistical model, estuarine aquaculture

Abstract: The UN has forecast that tropical estuarine environments will provide most of the marine food for the world's future population. It is thus essential that aquaculture is sustainably managed in these environments, and that appropriate scientific studies are undertaken based on routine ongoing data sampling. In this paper we focus on methods for predicting fish biomass for specified taxa communities as a function of upstream location, season, year, and corresponding fish larvae taxa density. These methods are illustrated using data from an ongoing 11-year study in the Na Thap River that flows into the Gulf of Thailand. Fish standing crop weights and corresponding larvae densities for 58 different taxa (mostly genera) were sampled monthly from ten sites.

To substantially reduce the number of variables, a common factor model was fitted by maximum likelihood to the correlation matrix of standing crop weights for these taxa after taking square roots to approximate normality assumptions, yielding a major group comprising 35 marine species, a smaller group of 15 preferring freshwater, and a small group of eight ubiquitous species. We then fitted simple linear regression models to predict crop weights from larvae densities (also square root-transformed) within each taxon, obtaining r-squared values ranging from 0.30 (Mantis Shrimp) to 0.86 (Black Tiger Shrimp) for the major group, from 0.58 (Sumatran Tiger Barb) to 0.92 (Common Silver Barb) for the minor group, and from 0 (Lanchester's Freshwater Prawn) to 0.18 (Dwarf Goby) for the ubiquitous group. To assess effects of season, annual trend and upstream location, we fitted separate multiple regression models with these three factors together with larvae density to the square roots of aggregated crop weights for each group. Results gave different interpretable patterns for the three groups. In summary, we believe that these methods could be usefully applied more generally to improve sustainable management of tropical estuarine fish farming and resource rehabilitation.

Community occupancy modeling of alpha, beta, and gamma bird diversity during the decade following forest fire

Morgan W. Tingley[a], Viviana Ruiz-Gutiérrez[b], Robert L. Wilkerson[c], Christine A. Howell[d], and Rodney B. Siegel[c]

[a]Ecology & Evolutionary Biology, University of Connecticut, Storrs, CT
morgan.tingley@uconn.edu

[b]Cornell Lab of Ornithology, Ithaca, NY

[c]The Institute for Bird Populations, Point Reyes Station, CA

[d]USDA Forest Service, Pacific Southwest Region, Vallejo, CA

Keywords: multi-species occupancy, colonization-extinction dynamics, diversity metrics, imperfect detection, hierarchical Bayesian.

Abstract: An emerging hypothesis in fire ecology is that pyrodiversity supports increased species diversity. This principle has primarily been tested across landscapes with mosaics of burned and unburned areas, rather than within the footprints of mixed-severity fires. We evaluate how multiple scales of bird community diversity (alpha, beta, and gamma) respond to burn severity in the first 10 years following fire. We use a Bayesian dynamic community occupancy model to generate robust diversity metrics while accounting for imperfect detection. We apply the model to a spatially and temporally extensive dataset of bird surveys in post-fire ecosystems in montane California, comprising 2334 surveys at 1106 survey points across 97 fires. Our results provide strong support for a positive relationship between pyrodiversity and bird diversity. Forest patches that burned at medium or high severity recruited or retained more species during the decade after fire than did unburned forest or patches of forest that burned at low severity. Much of this increased species richness was explained by the greater richness of cavity- and shrub-nesting birds occupying more severely burned patches. Consequently, when evaluated at the scale of individual fires, fires with a greater heterogeneity of burn severities had significantly more species.

Niche Overlap and Clustering in Ecological Communities

S. Pledger[a] and **D. Fernandez**[a]

[a]School of Mathematics and Statistics
Victoria University of Wellington
Wellington, New Zealand
shirley.pledger@vuw.ac.nz

Keywords: clustering; community structure; dimension reduction; finite mixture model; niche separation.

Abstract: Ecological community data often comes as an n by p matrix Y of abundances (counts, presence-absence or biomass) of p species over n sites. Various environmental variables, either numerical or categorical, may also be recorded at each site; these are coded as an n by m model matrix X. Can niche overlap models use the same data? For a given species, K environmental variables may be used to define a niche (a volume in K-dimensional hyperspace) within which the species is viable. The different dimensions may be categorical, e.g. soil type, or numerical, e.g. altitude (Geange et al., 2011). Models for the environmental niche overlap (or its complement, niche separation) between two species require (i) relevant environmental variables to be chosen and recorded, (ii) the sites to cover a wide range of these variables, and (iii) each species to have reasonably high total counts or incidences. Typically ecological community data has many rare species, for which niche overlap models fail under criterion (iii). Hence we cluster the p species into a few clusters ("species archetypes") with similar occurrence or abundance patterns within cluster (Pledger and Arnold, 2014). In this way, the rare species gain strength by being associated with more common species, and the counts per cluster are higher than for unclustered data. We show by examples how biclustering by finite mixtures provides a broad overview of patterns in the data and identifies the major variables for predicting niche separation.

References

Geange, S., Pledger, S., Burns, K. and Shima, J. (2011) A unified analysis of niche overlap incorporating data of different types. *Methods in Ecology and Evolution*, 2: 175-184.
Pledger, S. and Arnold, R. (2014) Multivariate methods using mixtures: Correspondence analysis, scaling and pattern-detection. *Computational Statistics and Data Analysis*, 71: 241-261.

Niche Overlap for Ecological Ordinal Data Incorporating RJMCMC for Mixture Models

D. Fernández[a] and **S. Pledger**[a]

[a]School of Mathematics and Statistics
Victoria University of Wellington
Wellington, New Zealand
dfdez23@hotmail.com

Keywords: Clustering; Community structure and dynamics; Dimension reduction; Finite mixture model; Braun-Blanquet scale; Multispecies models; Niche Overlap; Ordinal data; Reversible-jump MCMC; Species distribution models; Stereotype model

Abstract: A niche may be visualized as the volume in multidimensional hyperspace within which a species can maintain a viable population (Hutchinson, 1957), where each axis of the hypervolume may represent, for instance, an environmental variable at each site (e.g. soil type, rainfall). Niche overlap between species is a measure used in the analysis of ecological community data. Its quantification has become an important tool for investigating invasive species, relative abundance distributions, global change, species coexistence, and evolutionary diversification. Additionally, the use of ordinal data in ecological community data is very common (e.g. Braun-Blanquet scale). Recent research has developed a set of likelihood-based finite mixture models for a data matrix of ordinal data (Fernández *et al.*, 2016). This approach applies fuzzy clustering via finite mixtures to the stereotype model. We obtain the fuzzy allocation of species into corresponding clusters (called "species archetypes") by performing a Reversible-Jump MCMC sampler, the main steps of which are presented in this talk. We use this mixture approach to (a) detect species with similar patterns of abundance over the different sites and (b) determine niche overlaps between two groups of species. Examples with ordinal data sets will be shown to illustrate the application of this approach.

References

Fernández, D., Arnold, R. and Pledger, S. (2016). Mixture-based Clustering for the Ordered Stereotype. Computational Statistics and Data Analysis 93: 46-75.

Hutchinson G. (1957). Concluding Remarks. Cold Spring Harbor Symposia on Quantitative Biology 22: 415-427.

Modelling species interactions for multivariate abundance data with Gaussian copulas

G. Popovic [a], D.I. Warton [b], and F.K.C Hui [c]

[a] School of Mathematics and Statistics
University of New South Wales
Sydney, Australia
g.popovic@unsw.edu.au

[b] School of Mathematics and Statistics, Evolution & Ecology Research Centre, Centre for Ecosystem Science
University of New South Wales
Sydney, Australia
david.warton@unsw.edu.au

[c] Mathematical Sciences Institute
Australian National University
Canberra, Australia
fhui28@gmail.com

Keywords: multispecies models ; abundance; co-occurrence ; species interactions.

Abstract: Data for a community of species often arise as discrete correlated responses representing abundance or presence of individuals of each species. Interest is often in understanding species interactions, co-occurrence and inter-species dynamics. Multispecies models allow us to study species interactions and make inferences about shared responses to environment variables. We propose to build models for multivariate abundance data by combining discrete marginal distributions with covariance modelling techniques, using Gaussian copulas. These models respect key properties of abundance data, like overdispersion, and covariance modelling techniques enable estimation even when there are a large number of species relative to sample size. We use these to carry out inference on environmental and experimental variables as well as to investigate species interactions. Our algorithm makes use of standard covariance modelling algorithms developed for Gaussian data, but extends them via an iteratively reweighted scheme for maximum likelihood estimation of discrete models with Gaussian copulas. We illustrate our approach with overdispsered count data of hunting spiders, using factor analysis models and graphical models to parsimoniously model species co-dependence patterns.

Model-based approaches for vegetation classification and mapping

M. Lyons[a*], S. Foster[b], D. Keith[a], R. Kingsford[a] and D. Warton[a]

[a] UNSW Australia, Sydney, Australia
[b] CSIRO, Hobart, Australia
[*]Corresponding author: mitchell.lyons@unsw.edu.au

Keywords: clustering, optimal partition, distribution modelling, community ecology

The recent development of model-based approaches in multivariate ecology (e.g. Warton et al. 2015) provides a basis for also considering model-based approaches for traditional vegetation classification and mapping activities. We present two uses where we have found model-based approaches to be particularly valuable.

Assessing vegetation clustering solutions: a 'good' vegetation classification should provide information about the composition and abundance of the species within communities, and predictive models are a natural way of formalising this idea. We used generalised linear models fit separately to each species, taking categorical cluster membership as the predictor, and the multivariate species composition matrix as the response. We used the estimated predictive performance (in our case, summation of AIC across all species) of alternative models to: i) identify optimal partitioning of sites among multiple competing clustering solutions (that is, the one that minimises AIC); ii) identify species that contribute most to compositional differences between clusters (i.e. characteristic species); and iii) rank clusters by floristic 'distinctiveness' to guide classification refinement. We demonstrate this procedure on various clustering solutions for a real data set with around 500 sites and around 700 species, as well as simulated data sets.

Simultaneous classification and mapping: Multivariate mixture models offer the ability to streamline the traditionally two-step process of vegetation classification and mapping into a single analytical step, tightening the link between description of biodiversity patterns and their depiction in space. Yet little work has been done to incorporate multivariate species data and environmental data into the classification process. Using a special type of mixture model (from Foster et al. 2013), we simultaneously model both species and environmental data in a probabilistic classification framework. Model parameters are estimated for sites that have been sampled for species composition, as then that that classification can be predicted out into the rest of the (unsampled) environmental space. We demonstrate this model on a data set of around 10,000 sites and around 500 species, across an area of about 810,000 km^2 (about the size of Washington, Oregon and California combined).

References

Foster, S. D., Givens, G. H., Dornan, G. J., Dunstan, P. K., & Darnell, R. 2013. Modelling biological regions from multi-species and environmental data. Environmetrics, **24**:489-499.

Warton, D. I., Foster, S. D., De'ath, G., Stoklosa, J., & Dunstan, P. K. 2015. Model-based thinking for community ecology. Plant Ecology, **216**:669-682.

Bringing Hierarchical Occupancy Models to Species Interaction Networks

Ben. G. Weinstein [a] and Catherine H. Graham [b]

[a] Department of Ecology and Evolution
Stony Brook University
Stony Brook, NY
benjamin.weinstein@stonybrook.edu

[b] Department of Ecology and Evolution
Stony Brook University
Stony Brook, NY
catherine.graham@stonybrook.edu

Keywords: Hummingbirds, Trait-matching, Ecuador, Occupancy Models, Bayesian, Hierarchical, N-mixture

Abstract: Recent advances in network ecology have provided insights into the mechanisms that maintain biodiversity. While many studies focus on computing network properties, few studies have focused on how to generate networks from empirical data. For example, the importance of detection bias has been ignored in estimating network properties, despite the danger of confusing non-detection with non-interaction. By using N-mixture occupancy models adapted from wildlife ecology, we differentiate the probability of detecting a species from the intensity of species interactions. We combine simulations and an empirical dataset on Andean hummingbirds to show that ignoring detection bias can underestimate the importance of traits in determining species interactions. Models including detection bias accurately recovered simulated parameters and better fit the empirical dataset than models that assumed perfect detection or another commonly used approach using multinomial likelihood. Our results suggest that ignoring detection bias in networks can lead to misleading interpretations. Our approach reduces reliance on assumptions made by previous approaches, and can be extended to any type of ecological interaction. The increased flexibility of separately modelling process and observation error should increase the biological relevance ecological networks.

Influence of climate drivers on occupancy dynamics of amphibians in the Greater Yellowstone Area

W.R. Gould[a] A. Ray[b], B.R. Hossack[c], A. Sepulveda[d], D. Thoma[b], D. Patla[e], R. Daley[b] and R. Al-Chokhachy[d]

[a] Applied Statistics Program
New Mexico State University,
Las Cruces, NM, USA
wgould@nmsu.edu

[b] National Park Service
Greater Yellowstone Inventory and Monitoring Network
Bozeman, MT, USA
Andrew_Ray@nps.gov
Rob_Daley@nps.gov
Dave_Thoma@nps.gov

[c]U.S. Geological Survey
Aldo Leopold Wilderness Research Institute
Missoula, MT, USA
Blake_Hossack@usgs.gov

[d]U.S. Geological Survey,
Northern Rocky Mountain Science Center
Bozeman, MT, USA
Adam_Sepulveda@usgs.gov
Robert_Al-Chokhachy@usgs.gov

[e]Northern Rockies Conservation Cooperative
Jackson, WY, USA
dpatla@hughes.net

Keywords: climate; dynamics; Grand Teton National Park; occupancy; wetlands; Yellowstone National Park.

Abstract: We integrated seven years of amphibian monitoring data and high resolution gridded climate data from Grand Teton and Yellowstone national parks to explicitly model how variation in climate drivers (e.g., evapotranspiration, runoff, etc.), and flooding conditions affect the occurrence and breeding dynamics (i.e., annual extinction and colonization rates) of amphibians. Our results showed that models incorporating climate drivers outperformed models of amphibian breeding dynamics that were exclusively habitat-based. Moreover, climate-driven variation in extinction rates, but not colonization rates, disproportionately influenced amphibian occupancy in monitored wetlands. Wetland depth was also positively associated with breeding occupancy for all species. By better understanding how breeding dynamics are influenced by external factors, park managers can make science-based decisions that will increase the likelihood of long-term persistance of wetland species.

A Variational Bayes approach to the analysis of spatial site-occupancy models.

A. E. Clark[a,*], **R Altwegg**[a] and **J.T. Ormerod**[b]

[a]Department of Statistical Sciences
University of Cape Town
Rondebosch, 7701
Cape Town, South Africa

[b] School of Mathematics and Statistics
University of Sydney
ARC Centre of Excellence for Mathematical & Statistical Frontiers

[*]Contacting author: Allan.Clark@uct.ac.za

Keywords: logistic regression; species distribution model; variational Bayesian modelling

Abstract: Detection-nondetection data are often used to investigate species range dynamics using Bayesian occupancy models which rely on the use of Markov chain Monte Carlo (MCMC) methods to sample from the posterior distribution of the parameters of the model. In this article we develop a Variational Bayes (VB) approximation to the posterior distribution of the parameters of a single-season spatial site occupancy model. This task is accomplished through the development of iterative algorithms that do not use MCMC methods. Simulations and small practical examples demonstrate the effectiveness of the proposed technique.

References

Johnson, D.S., Conn, P.B., Hooten, M.B., Ray, J.C. and Pond, B.A., (2013) Spatial occupancy models for large data sets. *Ecology*, 94(4), pp.801-808.

Species range dynamics when detection is not perfect

Res Altwegg[a], **Guillaume Péron**[a], **Kristin Broms**[b], **Florent Bled**[a] and **James D Nichols**[c]

[a]Statistics in Ecology, Environment and Conservation
Department of Statistical Sciences
University of Cape Town, South Africa
Res.Altwegg@uct.ac.za

[b]Department of Fish, Wildlife, and Conservation Biology
Colorado State University
Fort Collins, Colorado, United States

[c]Patuxent Wildlife Research Center
US Geological Survey
Laurel, Maryland, United States

Keywords: dynamic occupancy models; global change; macroecology; Southern African Bird Atlas Project

Abstract: Some of the most pressing problems in nature conservation play out at large geographic scales, which has led to the new field of Conservation Biogeography. At this scale, incomplete detection and the difficulty of studying dynamic processes are two key issues. We used dynamic occupancy models adapted to bird atlas data for analysing patterns and processes in the distributions of bird species in southern Africa between 1987 and 2015. We found both clear legacies of past climatic conditions in the current species assemblages, and evidence that species ranges dynamically changed in apparent reaction to recent bush encroachment and changes in land use. Range changes led to opportunities for new ecological and evolutionary interactions between brood parasites and their host species. We also found evidence that pairs of closely related species competively limit the range of one another. Finally, we found that protected areas are critical for the persistence of some species. Our work shows the value of citizen science projects for our understanding of large-scale biodiversity processes but also highlights the need to account for the variable observation process that affects such data sets.

Estimating salmon escapement across the Snake River basin: a novel approach using PIT tags

K.E. See [a], J. White [a] and B.X. Semmens [b]

[a] Quantitative Consultants, Inc.
Seattle, WA
Kevin@QCInc.org
Jody@QCInc.org

[b] Scripps Institution of Oceanography
UC San Diego, CA
semmens@ucsd.edu

Keywords: patch-occupancy model, mark – recapture, salmon escapement, animal movement

Abstract: Estimates of salmonid adult escapement are crucial to evaluating the status of threatened populations, the success of recovery programs and the productivity of populations. In the Columbia Basin, these estimates are often made using redd counts or mark-recapture models from weir data, but these methods have limitations. Redd counts have (potentially) unknown observation error, and weirs can only be placed in certain locations. As an alternative method, we developed a branching patch-occupancy model (Royle & Dorazio, 2008) within a Bayesian framework that relies on adult salmon being PIT tagged at Lower Granite dam and subsequently re-detected by the extensive array of instream PIT tag antennas across the Snake River basin. This state-space model estimates the probability of fish moving past each detection point, as well as the probability of detection at each site and can then be used to estimate escapement at a variety of spatial scales across the entire basin. Estimates of steelhead and spring/summer Chinook from spawning years 2010 – 2015 compared favourably with independent estimates at a variety of locations, validating the model results.

References

Royle, J. A. and Dorazio, R. M. (2008). Hierarchical modeling and inference in ecology. Academic Press.

A new assessment tool for occupancy models when revisits are serially correlated

W.J. Wright[a], K.M. Irvine[b], and T.J. Rodhouse[c]

[a] Department of Mathematical Sciences
Montana State University
Bozeman, Montana
wilson.wright@msu.montana.edu

[b] Northern Rocky Mountain Science Center
US Geological Survey
Bozeman, Montana
kirvine@usgs.gov

[c] Upper Columbia Basin Network
National Park Service
Bend, Oregon
Tom_Rodhouse@nps.gov

Keywords: presence-absence, detection probability, goodness-of-fit, monitoring, bats.

Abstract: Occupancy models are an important tool for exploring the ecological processes that generate species distribution patterns and are increasingly important for large-scale conservation monitoring programs. These models are a type of zero-inflated binomial model, with explicit attention given to species detection probabilities estimated from replicate surveys to sample units within each field season. A central assumption is that these replicate surveys are independent Bernoulli trials with outcomes either detected or not detected. However, this assumption may be untenable when ecologists serially deploy remote cameras and acoustic recording devices over days and weeks to survey rare and elusive animals. Solutions to address this issue include occupancy models that incorporate a first-order Markov structure detection-level covariate (Hines et al. 2010). A way to evaluate the fit of this model versus the standard independent model is needed for practitioners; only an ominibus goodness-of-fit test based on Pearson's χ^2 test (MacKenzie and Bailey 2004) is currently available. We propose a novel join count χ^2 test adapted from spatial statistics to directly assess serial correlation after fitting a model. We motivate our work with a dataset of multi-night bat call recordings contributed as a pilot study for the North American Bat Monitoring Program. Using simulations, we compared these two tests and found that our join count χ^2 test was more reliable for detecting inadequacy of the standard independent model than the omnibus goodness-of-fit test. When applied to two bat species, our test illustrates that sophisticated models do not guarantee adequate fit to real data, underscoring the importance of model assessment. Our join count χ^2 test extends the capacities of occupancy modelling to serially-correlated data. This provides much-needed flexibility to ecological studies by enabling serial deployments of survey equipment that can result in substantial cost savings.

References

Hines, J.E., Nichols, J.D., Royle, J.A., MacKenzie, D.I., Gopalaswamy, A.M., Samba Kumar, N. & Karanth, K.U. (2010) Tigers on trails: occupancy modelling for cluster sampling. *Ecological Applications*, 20, 1456-1466.

MacKenzie, D.I. & Bailey, L.L. (2004) Assessing the fit of site occupancy models. *Journal of Agricultural, Biological, and Environmental Statistics*, 9, 300-318.

A dynamic state model to assess the population consequences of disturbance on migratory baleen whales

E. Pirotta [a], M. Mangel [b], D.P. Costa [c], E.A. McHuron [c], L. Schwarz [c], B. Mate [d],
L. Huckstadt [c] and L. New [a]

[a] School of Mathematics
Washington State University
Vancouver, WA, USA
enrico.pirotta@wsu.edu; leslie.new@wsu.edu

[b] Department of Applied Mathematics and Statistics
University of California,
Santa Cruz, CA, USA
msmangel@ucsc.edu

[c] Department of Ecology and Evolutionary Biology
University of California
Santa Cruz, CA, USA
costa@ucsc.edu; emchuron@ucsc.edu; lisa.schwarz@ucsc.edu; lahuckst@ucsc.edu

[d] Department of Fisheries and Wildlife
Oregon State University
Newport, OR, USA
bruce.mate@oregonstate.edu

Keywords: baleen whales, migration, state dependent behavioural models, stochastic dynamic programming, Monte Carlo simulations.

Abstract: Human activities at sea can cause sub-lethal changes in animals' behaviour, potentially affecting their energy budget. However, the relevance of these behavioural changes for the dynamics of marine populations is often difficult to predict, complicating management decisions. As capital breeders, baleen whales acquire most of the energy to sustain migration, gestation and lactation costs before the birth of their calf. Disturbance from human activities along the migratory route and in the summer feeding grounds can therefore have long-term repercussions on the probability of calf survival via a reduced condition of the mother. Insufficient energy reserves stored in her blubber may lead to the abortion of the foetus while pregnant, or the inability to deliver enough milk while lactating. We developed a dynamic state model, implemented via stochastic dynamic programming and forward Monte Carlo simulations (Mangel and Clark 1988), to predict the effects of disturbance on the survival probability of a calf. The model covers one reproductive cycle and identifies optimal behavioural decisions of the female based on the point in time along the cycle, her location and her blubber reserves. We applied the framework to eastern North Pacific blue whales, using tracking data and other existing ecological and physiological knowledge to inform model parameters, and explored the effects of a simulated gradient of disturbance. Our modelling effort provides a general approach to assess the population consequences of disturbance on migratory baleen whale species. Moreover, it helps characterising the sensitivity of predictions to parameter uncertainty, highlighting the key knowledge gaps that should be prioritized by future data collection.

References: Mangel, M. and Clark, C. W. (1988) *Dynamic Modeling in Behavioral Ecology*. Princeton University Press, Princeton, N.J.

Implementation of a Hierarchical Logistic Exposure Model for Nest Survival

J.E. Thornton and A.J. Kroll

Timberlands Technology Department
Weyerhaeuser Company
PO Box 9777
Federal Way, WA 98063
jamie.thornton@weyerhaeuser.com

Keywords: logistic exposure; nest survival; hierarchical models;

Abstract: Nest survival is often the focus of ecological research and is a critical component of many conservation and management programs. However, nest survival models are rarely developed within a coherent framework to account for important differences between experimental and observational research, including design factors and fixed vs. random effects. Here, we extend the logistic-exposure nest survival model into a multilevel structure. The hierarchical Bayesian implementation allows covariates and random effects to be included at different levels of interest (interval, nest, study plot) and to address the inherent split-plot design in nest survival studies. Model flexibility supports a range of study designs from standard observational studies to estimation of experimental treatments applied at multiple scales. We use an experimental study of Chestnut-sided Chickadee nest survival to examine model advantages and limitations. Additionally, we use simulation to examine effects of interval length, daily survival rate, total period, and sample size on survival probability estimation when using the logistic-exposure model.

References

Shaffer, T. (2004) A Unified Approach to Analyzing Nest Success. *The Auk*, 121(2):526-540.

Spatial and temporal variation of water temperature regimes

E. Ashley Steel [a]

[a] PNW Research Station
USDA Forest Service
asteel@fs.fed.us

Keywords: water temperature, spatial stream network models, wavelets

Abstract: Riverine thermal regimes cannot be captured, understood, or managed through mean temperatures or single-point measurements. We now understand that thermal regimes include minute-by-minute, day-to-night, and season-to-season fluctuations in water temperature at every location within the stream network. They also include meter-to-meter and reach-to-reach variability at every minute in time. Human actions and other disturbances, e.g., dams, land-use, climate change, and wildfire, impact not only mean and maximum temperatures but also the variability and complexity of water temperature regimes. Such changes have the potential for dramatic sub-lethal effects on food web dynamics, species phenology, and community composition. We summarize, for example, experimental results linking degree-day delivery to Chinook salmon phenology and explore the behavior of a mathematical model of Chinook salmon egg development when water temperatures are allowed to fluctuate. Understanding the drivers and patterns of thermal complexity is therefore essential for effective management of aquatic systems. Such efforts require a combination of extensive empirical data collection and new modeling approaches. Using data collected every 30-min across the Snoqualmie River, WA, we apply spatial stream network models (SSNMs) and wavelet decomposition to map variability at multiple time scales; explore spatial pattern across metrics and over time; and conceptualize which aspects of temporal variance are driven by within network versus out-of-network processes. We demonstrate, for example, that while the spatial distribution of mean temperature over the stream network varies little between summer and winter, the spatial distribution of variance varies tremendously between summer and winter.

References

Steel, E. A., Sowder, C. and E.E. Peterson. (In press) Spatial and temporal variation of water temperature regimes on the Snoqualmie River network. *Journal of the American Water Resources Association*.

A hierarchical Bayesian approach to understand Atlantic salmon population dynamics in two Canadian catchments.

G. Dauphin [a] and R. Cunjak [a]

[a] Canadian Rivers Institute
University of New Brunswick
Fredericton, New Brunswick, CANADA
gdauphin@unb.ca
cunjak@unb.ca

Keywords: Atlantic salmon, Hierarchical Bayesian model, Population dynamics.

Abstract:

Atlantic salmon (*Salmo salar*) populations have been decreasing since the 1980's. Consequently, there has been a widespread call to address the decline and devise a comprehensive conservation plan. Key to such an effort is a thorough analysis of the freshwater data to identify temporal and spatial trends in stock recruitment dynamics, and potential bottlenecks to production. In this study we attempt to understand at the mechanisms affecting the population dynamics of Atlantic salmon populations in two Canadian catchments (Miramichi and Restigouche). To understand these mechanisms we i) use long-term (>40 years) electrofishing data collected by the Department of Fisheries of Oceans in a hierarchical Bayesian model to obtain juvenile production estimates at the watershed scale and ii) use these estimates and adult returns counts in a state-space model to estimate density dependent relationships and factors affecting these relationships.

Statistical modeling of forest stand dynamics using inhomogeneous Markov chains.

J. Liénard[a] and **N. Strigul**[b] *

Department of Mathematics and Statistics
Washington State University Vancouver
WA, USA
[a] jean.lienard@wsu.edu

[b] nick.strigul@wsu.edu

Keywords: inhomogeneous Markov chains; patch mosaic concept; forest stand dynamics; climate change scenarios

Abstract: Forecasting forest dynamics under evolving environmental conditions is a challenging problem, as it involves considering together two interdependent components, climatic regimes and forested ecosystems. We propose here a novel method to model progressive dynamic changes in forest biomass resulting from variations in natural disturbance regimes and land-use practices using inhomogeneous Markov Chain models. First we estimate transition matrices based on forest inventories using a Bayesian approach, then we incorporate dynamic disturbance and forest growth scenarios and simulate transient dynamics and stationary states. We apply this method to investigate the consequences of existing global warming scenarios, which predict changes in fire rate in Quebec hardwood forests as well as possible growth enhancements due to increasing CO_2 and temperature. We found that none of the considered scenarios was able to counterbalance the currently observed trend of increasing biomass, in the next 30 to 40 years. At horizon 2090, the scenarios predict diverging but overall limited changes of about ±5% mean biomass. The developed modeling approach is capable of unifying predictions from mechanistic models of disturbance and growth, and provides a useful tool to reduce uncertainties associated with managing complex ecosystems.

References

Lienard, J. F. and Strigul, N. S. (2016). Modeling of hardwood forest in Quebec under dynamic disturbance regimes: a time-inhomogeneous Markov chain approach. *Journal of Ecology*, in press.

Lienard, J. F., Gravel, D., and Strigul, N. S. (2015). Data-intensive modeling of forest dynamics. *Environmental Modelling and Software*, 67: 138-148

Strigul, N. (2012). Individual-based models and scaling methods for ecological forestry: implications of tree phenotypic plasticity. In Garcia, J. and Casero, J., editors, *Sustainable Forest Management*, pages 359-384. InTech, Rijeka, Croatia.

Strigul, N., Florescu, I., Welden, A. R., and Michalczewski, F. (2012). Modelling of forest stand dynamics using Markov chains. *Environmental Modelling and Software*, 31:64-75

Reducing variance and bias in indices of abundance based on commercial catch and survey data using a Bayesian hierarchical model for zero-inflated spatial biomass data

J.B. Lecomte [a], R. Forrest [a] and M.P. Etienne [b]

[a] Pacific Biological Station,
Fisheries and Oceans Canada,
Nanaimo, British Columbia, Canada
Jean-baptiste.lecomte@dfo-mpo.gc.ca
Robyn.forrest@dfo-mpo.gc.ca

[b] Laboratory for Risk Management in Environmental Sciences,
UMR AgroParisTech/INRA 518,
Paris, France
marie.etienne@agroparistech.fr

Keywords: abundance index; zero-inflated data; spatial model; stock assessment

Abstract: A large proportion of quota groundfish species in British Columbia (BC) Canada are data-limited for stock assessment. The greatest challenge for assessing data-limited species is the lack of reliable, long-term abundance indices, where often the only long-term index comes from commercial or recreational fishing data; or from very patchy survey observations. Fishery-dependent catch-per-unit-effort (CPUE) indices are subject to biases arising from non-random fishing effort and from incorrect spatial assumptions that have the implicit assumption that average abundance in unobserved areas is the same as in observed areas. In addition, data-limited species often present a high proportion of zero observations, which require specific zero-inflated modelling approaches to avoid further bias.

We propose a Bayesian hierarchical model for zero-inflated spatial biomass data to provide less biased abundance indices for the quota groundfish species in BC. Due to the excess of zero values, we use a Compound Poisson-Gamma model for the observation layer. The spatial structure is modeled in the latent layer by convolving an independent process with an exponential kernel on a spatial grid. This class of models offers a number of advantages over the standard classes of parametric variogram/covariance functions commonly used in geostatistics. In addition, ecological factors such as temperature, depth and sediment type are included as predictors in the latent model. Due to the flexibility of the approach, we also investigate methods to extend to a spatio-temporal model. This is the first time this model has been applied to groundfish species in BC and represents a novel method to improve indices of abundance for data-limited stocks.

References

Lecomte, J.B., et al. 2013a. Compound Poisson-Gamma versus Delta-Gamma to handle zero-inflated continuous data under a variable sampling volume. Methods Ecol Evol 4: 1159-1166.

Robust statistical method in a single species age-structured state-space assessments model

A. Nielsen[a] and **C.W. Berg**[a]

[a]Technical University of Denmark
National Institute for Aquatic Ressources
Charlottenlund, Denmark
an@aqua.dtu.dk

Keywords: Fish stock assessment; age-structured; state-space; robust method

Abstract: State-space fish stock assessment models are increasingly used as alternative to (semi-) deterministic procedures and to full parametric statistical catch-at-age models. The distribution of change between time steps is part of the model formulation in state-space models and the model parameters are estimated from the entire data set, so if an unusual large change occur (e.g. in recruitment or mortality) then the distribution should be flexible enough to allow it. Robust statistical methods are normally used to reduce the influence of extreme outlying (potentially wrong) data points. A common choice is to substitute the Gaussian distribution with a robust mixture between a Gaussian and a heavy tailed distribution. Here it is investigated if a robust mixture is also useful in the process model. The mixture distribution is applied in the recruitment process to allow for sudden large recruitment events and in the fishing mortality process to allow for sudden large changes in the fishing pattern or the fishing level. A simulation study is conducted to validate the approximation used for estimation in the state-space assessment model. Finally, the different configurations of the assessment model are applied to North Sea Haddock, where previous assessments have estimated extreme recruitment events.

Reference

Nielsen, A. and Berg, C.W. (2014). Estimation of time-varying selectivity in stock assessments using state-space models. *Fisheries Research*, 158:96-101.

A stochastic surplus production model in continuous-time incorporating seasonal harvesting

M.W. Pedersen[a] and **C.W. Berg**[a] *

[a]Technical University of Denmark
National Institute for Aquatic Ressources
Charlottenlund, Denmark
mawp@dtu.dk

Keywords: Data-limited methods, fisheries management, maximum sustainable yield, seasonal population dynamics, stock assessment, Pella-Tomlinson model

Abstract: Surplus production modelling has a long history as a method for managing fish stocks. Recent advancements have cast surplus production models as state-space models that separate random variability of population dynamics from error in observed indices of biomass. We present a stochastic surplus production model in continuous-time (SPiCT), which in addition to stock dynamics also models the dynamics of the fisheries. This enables error in the catch process to be reflected in the uncertainty of estimated model parameters and management quantities. Benefits of the continuous-time state-space model formulation include the ability to accommodate data sampled at arbitrary and possibly irregular intervals and missing values. However, such sub-annual fisheries data often contain seasonal variability that if unmodelled leads to systematic bias. We present two approaches to model seasonal harvesting. The simpler model uses a cyclic spline to represent seasonality while the more complex alternative model relies on a coupled system of equations that allows for shifts in intensity and timing of the seasonal fishing pattern. We show in a simulation that the ability to analyse sub-annual data can increase the effective sample size and improve estimation of model parameters relative to discrete-time analysis of aggregated annual data. We furthermore illustrate the flexibility of the model by analysing sub-annual fisheries data from the North Sea.

Quantifying temporal turnover in biodiversity, and how it varies spatially for the North Sea fish community

Y. Yuan* and **S. T. Buckland**

Department of Mathematics and Statistics
University of St Andrews
The Observatory, Buchanan Gardens, St Andrews, UK
*yy84@st-andrews.ac.uk

Keywords: biodiversity, turnover measures, spatial-temporal analysis, Bayesian Hierarchical model, INLA, stochastic partial differential equations

Abstract: Quantifying species turnover is an important aspect of biodiversity monitoring. Turnover measures are usually based on species presence/absence data, reflecting the rate at which species are replaced. However, if available, it is more informative to use species abundance estimates to measure the compositional change over time, and how it varies spatially. Turnover measures that reflect the rate at which individuals of a species are replaced by individuals of another species are far more sensitive to change. We perform Bayesian spatial-temporal data analysis using the stochastic partial differential equation (SPDE) approach with integrated nested Laplace approximations (INLA), and then use the predicted density surface for each species over time to quantify the rate at which individuals of one species are replaced by individuals of other species. We use a dataset of North Sea fish from the International Bottom Trawl Survey (IBTS) to illustrate the methods. We apply several turnover measures to explore how temporal trends in the individual-based species turnover vary spatially, which helps to identify how climate change is affecting the North Sea fish community.

Adaptive nonparametric smoothing for capture-recapture models with application to survival estimation for migrating juvenile salmon

J.R. Faulkner[a,b] and **V.N. Minin**[c]

[a]Quantitative Ecology and Resource Management
University of Washington
Seattle, Washington
[b]National Oceanic and Atmospheric Administration
Northwest Fisheries Science Center
Seattle, Washington
jfaulknr@uw.edu

[c]Departments of Statistics and Biology
University of Washington
Seattle, Washington
vminin@uw.edu

Keywords: Bayesian; shrinkage prior; data augmentation; Hamiltonian Monte Carlo

Abstract: We extend a locally-adaptive nonparametric curve-fitting method known as the Bayesian trend filter for use in capture-recapture modeling. We use this method to estimate nonlinear functions of time or covariates for survival and capture probabilities within a Cormack-Jolly-Seber capture-recapture modeling framework. This fully Bayesian approach places a horseshoe prior on the kth-order differences in the discretized latent trend function. This formulation allows adaptation to local changes in smoothness of a function, including abrupt changes or jumps, without compromising smoothness across the remaining function. We use Hamiltonian Monte Carlo to approximate the posterior distribution of model parameters because this method provides superior performance in the presence of the high dimensionality and strong parameter correlations exhibited by our models. We use simulated data to assess the performance of the method, and we demonstrate its application on data from migrating juvenile salmon.

A novel individual-based forest model, LES, demonstrates that gap dynamics are sufficient to explain shade tolerance changes along soil moisture gradient.

N. Strigul[a] and **J. Liénard**[b]

Department of Mathematics and Statistics
Washington State University Vancouver
WA, USA
[a] nick.strigul@wsu.edu

[b] jean.lienard@wsu.edu

Keywords: individual-based forest model; forest succession; shade tolerance; gap dynamics; crown competition; root competition; moisture gradient

Abstract: Competition of individual trees occurs above- and below- ground simultaneously and results in forest self-organization that is poorly understood at the community level. We performed statistical analyses of two North American forest inventory databases, and report a substantial decrease of shade tolerance index along mesic—xeric soil transition. We hypothesize that gap dynamics can cause these transition patterns. We examine this hypothesis using a spatially-explicit individual-based model, called LES, simulating simultaneous crown and root competitions. In contrast with previous studies, our model does not assume any trade-off between shade- and drought-tolerances. Computer simulations demonstrate that the transition from mesic to xeric soils is accompanied by a decrease of shade tolerant index, similarly to the pattern recorded in forest inventories. This transition is also accompanied by a change from water to sunlight competitions, and happens at three successive stages: (1) mostly water-limited parkland, (2) simultaneously water- and sunlight-limited closed canopy forests featuring a very sparse understory, and (3) mostly sunlight-limited forests with a populated understory. This pattern is caused by contrasting effects of gap dynamics which favor either shade tolerant or shade intolerant species, depending on soil moisture and understory density. Gap dynamics may explain the observed differences in distributions of shade tolerant and shade intolerant species along latitude and moisture gradients. This work demonstrates that forest patterns along environmental gradients can evolve from tree spatial competition without physiological trade-offs. Mechanistic understanding of population processes involved in the forest-parkland-desert transition will improve our ability to explain species distributions and predict forest responses to climatic changes.

References

Lienard, J. F. and Strigul, N. S. (2016). An individual-based forest model links canopy dynamics and shade tolerances along a soil moisture gradient. *Royal Society Open Science*, in press.
Lienard, J. F. and Strigul, N. S. (2015). Linking forest shade tolerance and soil moisture in North America. *Ecological Indicators*, 58: 332-334.
Strigul, N. (2012). Individual-based models and scaling methods for ecological forestry: implications of tree phenotypic plasticity. In Garcia, J. and Casero, J., editors, *Sustainable Forest Management*, pages 359-384. InTech, Rijeka, Croatia.

Big data challenges of full life-cycle abundance modeling with citizen science data

D. Fink[a], T. Auer[a], N.E. Bruns[a], W.M. Hochachka[a], and S. Kelling[a]

[a] Cornell Lab of Ornithology
Ithaca, NY, USA
daniel.fink@cornell.edu

Keywords: abundance, big data, citizen science, species distribution models.

Abstract:

Effective management of bird species across their life cycles requires knowledge of how distributions and abundances change throughout the year. However, obtaining this knowledge requires overcoming two main data-analysis challenges: spatial and temporal non-stationarity, and the logistics of working with big data. In our presentation, we describe the analytical approach that we have taken in order to create year-round abundance models for bird species from data collected by the citizen science project eBird (http://www.ebird.org). For each species, we created an ecological niche model, predicting abundances of birds based on associations of species with environmental features. These associations vary through space and time in ways that are unique to each species. As a generalized solution to the non-stationarity challenge, we have employed a divide-and-recombine strategy (Fink et al. 2010), in which an ensemble of independent sub-models was fit to data within restricted time periods and spatial regions with results subsequently aggregated. Each sub-model was a zero-inflated abundance model, needed to account for the potentially large proportions of observations in which recorded abundances were zero. The large volumes of data used for these analyses (over 5 million species checklists at over 800,000 unique locations) necessitated using high performance computing systems. We used the divide-and-recombine structure of the model to implement a simple parallel computational strategy within the Hadoop MapReduce framework on the Microsoft Azure cloud. From these models we calculated weekly species abundance estimates across North America, used for the 2016 State of The Birds Report, a tri-national (Canada, USA, Mexico) conservation report issued by the U.S. Department of the Interior. Based on our experiences, we believe that both statistical and computational efficiency will become important issues in the analyses of ecological data collected across large spatial extents.

References

Fink, D., Hochachka, W. M., Zuckerberg, B., Winkler, D. W., Shaby, B., Munson, M. Arthur, Hooker, G. J., Riedewald, M., Sheldon, D., Kelling, S. Spatiotemporal Exploratory models for Large-scale Survey Data. *Ecological Applications*, 20(8), 2131-2147.

Using citizen science to monitor avian communities in the Arizona Sky Islands, USA

J.S. Sanderlin[a,b], W.M. Block[a,c] , J.L. Ganey[a,d], and J.M. Iniguez[a,e]

[a] Rocky Mountain Research Station, U.S. Forest Service
2500 S. Pine Knoll Dr.
Flagstaff, AZ 86001, USA

[b] jlsanderlin@fs.fed.us

[c] wblock@fs.fed.us

[d] jganey@fs.fed.us

[e] jiniguez@fs.fed.us

Keywords: Bayesian hierarchical models; bird communities; citizen science; local extinction probability; local colonization probability; multi-species, multi-season occupancy models; species richness.

Abstract: The avifauna within southeastern Arizona Sky Islands includes species found nowhere else in the U.S. Thus, birdwatchers from across the globe visit the region, providing a vibrant state and local ecotourism industry. Rocky Mountain Research Station (RMRS) scientists initiated a bird study across montane forest and woodland types in 5 Coronado National Forest mountains from 1991 to 1995. Since then, the region has been under increased stress from ongoing drought and wildfires. We know little about fire effects on and habitat associations of Neotropical migratory birds in this region. Southern Arizona is also unique in having skilled citizens able to identify birds by sight and sound. These citizens are eager to assist with monitoring bird populations, but often are unfamiliar with design and implementation of rigorous monitoring programs. At the request of local citizens, RMRS scientists initiated a study in 2012 and 2013 in the Chiricahua Mountains to sample original and new trail transects in coordination with a local citizen group. For data calibration, a professional bird crew sampled trail and original transects in 2014. To implement a long-term monitoring plan, our objectives were to 1) evaluate occupancy and detection differences between avian communities along trails versus original transects, which cut across the terrain along a random direction, and 2) evaluate detection differences between professional bird crew and citizen scientist observers. We used multi-species, multi-season and single season occupancy models in a Bayesian hierarchical framework to estimate occupancy, species richness, and local extinction/colonization, while accounting for imperfect detection. We show how this effort allows for inexpensive and statistically rigorous long-term monitoring, and fosters greater local involvement in science and conservation.

Separating signal from noise in citizen-science data: Using hierarchical models to estimate relative abundance from camera-trap data

T. Arnold[1], M. Palmer[2], M. Kosmala[2], A. Swanson[2], E. Erhart[1], J. Forester[1], and C. Packer[2]

[1]Department of Fisheries, Wildlife, and Conservation Biology
[2]Department of Ecology, Evolution, and Behavior
University of Minnesota
St. Paul, MN, USA
arnol065@umn.edu

Keywords: camera traps, citizen science, N-mixture models, relative abundance.

Abstract: We analysed data on occurrence and abundance of Serengeti wildlife based on images classified by researchers, technicians and citizen scientists via the Snapshot Serengeti website. Previous analyses suggest that citizen-scientist data can be used to reliably estimate species occurrence (~97% accuracy; A. Swanson et al., unpubl. data), but estimates of abundance are more biased and less precise. We used zebra and wildebeest images to develop alternative formulations of N-mixture models to provide less biased and more precise estimates of abundance while continuing to utilize highly efficient crowd-sourcing for image classification. Models using negative binomial and beta-binomial distributions explained some of the excess heterogeneity in abundance and detection probability, respectively, but we also identified additional sources of variation that required more complex model formulations. A large source of heterogeneity in count data was associated with proximity of animals to the camera, which we approximated by measuring dimensions of individual animals in standardized images. Detection probability declined from 0.965 (SE 0.009) to 0.611 (SE 0.035) for zebra designated as close versus far, whereas wildebeest had lower overall detection and a greater decline in detection with increasing distance ($p = 0.880 \pm 0.013$ for close versus 0.271 ± 0.0261 for far). For citizen-scientist classified images that lacked data on proximity of animals to the camera, we were nevertheless able to approximate this source of variation by hypothesizing latent mixtures of close animals (which were accessible to all observers and detected with high probabilities) versus distant animals (which were not present in all images, were not identified by all observers, and were detected at lower probabilities). Martin et al. (2011) extended N-mixture models to include site- and survey-specific sources of latent variation in detection probability, and our analyses further extend this concept to among-individual variation occurring within a single survey.

References

Martin, J., Royle, J.A., Mackenzie, D.I., Edwards, H.H., Kéry, M., and Gardner, B.. (2011). Accounting for non-independent detection when estimating abundance of organisms with a Bayesian approach. *Methods in Ecology and Evolution*, 2:595-601.

The second Southern African Bird Atlas Project: causes and consequences of geographical sampling bias

S. Hugo [a] and R. Altwegg [b]

[a,b] Centre for Statistics in Ecology, Environment and Conservation
Department of Statistical Sciences
University of Cape Town
Rondebosch 7701
South Africa

[a] sanethug@gmail.com
[b] res.altwegg@gmail.com

Keywords: citizen science; generalized least squares; sampling effort; southern Africa; species distribution atlas; spatial bias.

Abstract: The second Southern African Bird Atlas Project (SABAP2) aims to record bird species within eight southern African countries at a relatively fine spatial resolution (i.e. grid cells of five minutes latitude by five minutes longitude, colloquially termed ‘pentads’). Since commencing in 2007, volunteers and citizen scientists have collected all of the data comprising more than 144000 surveys with more than 7 million records (http://sabap2.adu.org.za/). However, SABAP2 is characterised by a pronounced geographical bias in sampling effort that may lead to distorted views of species distributions. Focusing on South Africa, we aimed to examine the possible determinants and effects of spatial variation in sampling effort for SABAP2. First we observed that sampling effort (i.e. the number of surveys per pentad) tends to decrease with distance from a few extremely well-sampled pentads (that we termed ‘sampling hubs’) corresponding to several major cities and popular bird watching locations scattered across South Africa. We use generalized least squares regression, taking spatial autocorrelation into account, to show that spatial variation in sampling effort is best explained by smallest distance from a ‘sampling hub’, smallest distance from a main road, and protected area cover. Further, based on the amount of effort spent within distinct environmental zones (i.e. subsets of pentads defined by similarities in climate and vegetation), we demonstrate that most sampling effort, and repeatedly-surveyed pentads, were concentrated in regions that are not representative of the average or range of environments within South Africa. Large parts of South Africa remain under-represented, which casts doubt on the effectiveness of species distribution and occupancy modelling techniques (Araújo and Guisan 2006). However, by building on our findings of the determinants of sampling effort, atlas coordinators may improve representativeness by directing volunteers toward undersampled regions while taking into account volunteer preferences (Robertson *et al*. 2010).

References

Araújo, M.B. and Guisan, A. (2006) Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, 33:1677–1688.

Robertson, M.P., Cumming, G.S. and Erasmus, B.F.N. (2010) Getting the most out of atlas data. *Diversity and Distributions*, 16:363–375.

Selecting relevant predictors for presence-only species distribution modelling. A case study from the marine environment missing documented absences and uncertain distributions

S. Bosch[a,b], L. Tyberghein[b] and O. De Clerck[a]

[a] Phycology Research Group, Ghent University
Ghent, Belgium
samuel.bosch@ugent.be
olivier.declerck@ugent.be

[b] Flanders Marine Institute (VLIZ)
Ostend, Belgium
lennert.tyberghein@vliz.be

Keywords: benchmark dataset; SDM; marine; model selection.

Abstract: Ideally, datasets for species distribution modelling contain data covering the entire distribution of the species, with evenly sampled records, confirmed absences and auxiliary data (e.g. physiological experiments) allowing informed decisions on predictor selection. Unfortunately, these criteria are only rarely met in real world datasets. This is even more the case for marine organisms for which distributions are too often only scantly characterized and absences generally not recorded. In order to evaluate predictor selection (see Barbet-Massin & Jetz 2014) we compiled a dataset of marine species suitable for benchmarking species distribution modelling, MarineSPEED. We selected 530 well studied and well identifiable species from all major marine taxonomic groups with different range sizes and from different ecoregions. Over three million distribution records where compiled from public sources (e.g. OBIS, GBIF, EMODNET, Reef Life Survey) and linked to environmental and biological data enabling refinement of the analytical results by species group, ecology, geography or life history characteristics.
Using this dataset, predictor selection was performed under different variations of numbers of predictor variables and sampling bias correction. Distributions were modelled with combinations of 3, 4, 7 and 8 environmental variables (calcite, chlorophyll a, diffuse attenuation, nitrate, photosynthetically active radiation, pH, phosphate, salinity, silicate, sea surface temperature, bathymetry and distance to shore) selected from Bio-ORACLE and MARSPEC. Sampling bias correction was performed using spatial thinning and a target group background. Performance of the models was evaluated using random as well as spatial cross-validation (Hijmans 2012). From the results we can conclude that the model setup has a significant impact on predictor selection. The MarineSPEED dataset can serve to evaluate the performance of modelling techniques aimed at predicting distributions of species under current and future climatic scenarios and under a wide array of parameter settings.

References

Barbet-Massin, M., & Jetz, W. (2014). A 40-year, continent-wide, multispecies assessment of relevant climate predictors for species distribution modelling. *Diversity and Distributions*, 20(11), 1285–1295.
Hijmans, R. J. (2012). Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. *Ecology*, 93(3), 679–688.

Evaluation of Some Weibull Parameter Estimation Methods for Characterizing Stem Distribution and Site Quality in a Tropical Mixed Forest of Southern Nigeria

A.A. Adeyemi

Department of Forestry and Wildlife Technology, Federal University of Technology, Owerri, Nigeria
adesoji.adeyemi@futo.edu.ng

Keywords: density; diameter distribution; species richness; slenderness coefficient; model.

Abstract: Weibull distribution is the most suitable function used in characterizing diameter distribution, especially in planted forests. However, the suitability of this function in natural forests has rarely been tested. There is also paucity of information on the most appropriate Weibull parameter estimation method(s) in natural forests in Nigeria. Similarly, assessment of site quality in natural forest has remained a challenge because of difficulty in determining tree ages. Moreover, the applicability of site form (SF) method for site quality evaluation in natural forests has not been sufficiently proven elsewhere. Therefore, this study evaluated parameter estimation methods for tree diameter characterization in a natural forest. Site quality of the different locations in Oban Forest was also assessed using the SF method. Systematic sampling technique was adopted for plot locations in the four study sites (*Aking*, *Ekang*, *Erokut* and *Ekuri*). Four transects of 2 km long at 600 m apart were laid in each of the four study sites. Five 50 m × 50 m plots were demarcated alternately along each transect at 400 m intervals. A total of 16 transects and 80 sample plots were used for the study. Tree diameters, at breast height ≥10 cm, base, middle and merchantable limit and total height (TH) were measured to compute quadratic mean diameter (D_q), stand density, stem volume (SV) and slenderness co-efficient (SC). Trees were identified and classified into genera and families. Soil samples were collected from 25 sub-plots of size 0.01ha within each of the sample plots at 0-15 and 15-30 cm depths using soil auger. The samples were air-dried and a composite soil sample was then formed per plot and analyzed using standard laboratory procedures. Data were statistically analyzed using descriptive statistics, correlation and regression analyses. The fitted models were assessed using co-efficient of determination (R^2), root mean square error (RMSE), ANOVA and t-test. Moment, percentile, hybrid and maximum likelihood Weibull parameter estimation methods were evaluated using negative log-likelihood (mLogL) statistic. A total of 131 tree species distributed among 113 genera and 41 families were identified. *Strombosia postulata* had the highest relative density (RD) of 4.9%. *Andira inermis*, *Baphia obanensis*, *Drypetes gossweileri*, *Hymenodictyon pachyantha*, *Oxystigma mannii*, *Scottellia coriacea*, *Strephonema mannii*, *Zenkerella citrina* and *Ziziphus espinosa*, were sighted only in *Ekuri* site with RD of 0.06%, 0.45%, 0.11%, 0.17%, 0.06%, 0.17%, 0.17%, 0.28% and 0.17%, respectively. About 76% of the species were locally threatened, 21.4% were rare, while 2.3% were occasional with only *S. postulata* being frequent. The species richness index was 2.38. For pooled data, there were 152±9 stems/ha with a mean basal area of 24.27±11.87m^2/ha. On site basis, *Ekuri* had highest stems/ha of 173±51 while *Erokut* had the least (143±39). *Ekang* had the highest mean SV/ha of 3847.14±2.16 m^3 with a least of 2127.71±1382 m^3/ha in *Erokut*. *Ekuri* had the highest of 62 stems/ha with high SC (SC: >80). The least of 37 stems/ha with high SC were recorded in *Erokut*. About 45, 50, 50 and 43 stems/ha at *Aking*, *Ekang*, *Erokut* and *Ekuri*, respectively had moderate SC (SC: 70-80). There were 66, 52, 56 and 84 trees/ha with low SC (SC: <70) in the four sites. The SF of 23.18 m, 26.52 m, 21.70 m and 23.69 m were obtained for *Aking*, *Ekang*, *Erokut* and *Ekuri*, respectively. Tree BA, SV and SF had positive and significant correlation with most of the soil physico-chemical parameters. For instance, SV correlated with soil organic matter (r = 0.995*), calcium (r = 0.99*) and silt (r = 0.95*). The most suitable height-diameter model was logarithmic (TH = -14.759 + 11.339lnDbh) with R^2 and RMSE of 0.51 and 4.462, respectively. The most suitable SV model was polynomial (SV = -2.4585 + 0.0063Dbh + 0.0117Dbh^2) with R^2 and RMSE of 0.85 and 2.207, respectively. The percentile-based method was the best for Weibull parameters estimations. For this method, the 25th and 95th was the most appropriate percentile pair (with least mLogL of 53.71±6.09), with D_q as the only suitable predictor for Weibull parameters.

Varying spatial correlation structure in grey seal movement data

Haakon Bakka[a], **Esther L. Jones**[b,c*], Janine Illian[a,b], Daniel Simpson[d] and Håvard Rue[a]

[a] Institute of Mathematical Sciences
Norwegian University of Science and Technology
Trondheim, Norway
haakon.c.bakka@gmail.com

[b] Centre for Research into Ecological and Environmental Modelling
University of St Andrews
St Andrews, UK
el298@st-andrews.ac.uk

[c] Sea Mammal Research Unit, Scottish Oceans Institute
University of St Andrews
St Andrews, UK

[d] Department of Mathematical Sciences
University of Bath
Bath, UK

Keywords: spatial ecology, spatial autocorrelation, grey seals, activity budget, second order structure, non-stationary random field

Abstract: Spatial ecology is dependent on our ability to construct spatial abundance maps that do not break with ecological assumptions. For some species, environmental covariates (e.g. terrain) have a strong influence on animal movement, both in the abundance mean and in the spatial autocorrelation. We use a case study of grey seal movements around the UK to account for spatial clustering in different areas, reflecting the daily activities of the species: land, resting near-shore, travelling, and foraging areas. By having one spatial correlation range for each area we produce a model accounting for varying spatial autocorrelation (Lindgren *et al.* 2011). Activity budgets combine information on animal movement with other sensor data (e.g. angular momentum) and are often used to define animal behavior (McClintock *et al.* 2013). We also test whether the second order clustering structure of the point pattern can be used as a proxy for defining activity budgets.

References

Lindgren, F., Rue, H. & Lindstrom, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields : the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **73**, 423–498.

McClintock, B., Russell, D.J.F., Matthiopoulos, J. & King, R. (2013). Combining individual animal movement and ancillary biotelemetry data to investigate population-level activity budgets. *Ecology*, **94**, 838–849.

A spatio-temporal model of catch and effort for snow crab (*Chionoecetes opilio*) in the Southern Gulf of St. Lawrence, Canada : A precursor to an integrated spatial stock assessment model

N. Cadigan

Centre for Fisheries Ecosystems Research
Marine Institute of Memorial University of Newfoundland
St. John's, NL, Canada
noel.cadigan@mi.mun.ca

Keywords: depletion model, errors in variables, spatial model, maximum likelihood.

Abstract: Fish stock size declines throughout the fishing season as fishery catches increase and this relationship can sometimes be used to estimate stock size based on the rate of decline in a catch per unit effort (CPUE) time-series. However, this decline can also be affected by changes in the spatial distribution of the fishing fleet. I propose a spatio-temporal depletion model that adjusts for bias caused by changes in the fleet distribution. This model is applied to a time-series of catch and effort data for the economically important snow crab fishery in the Southern Gulf of St. Lawrence off the east coast of Canada. I illustrate how the model can be extended to integrate other survey data.

Point process models for spatio-temporal distance sampling data

Y.Yuan[a*], **F.E.Bachl**[b], **D.L.Borchers**[a], **F.Lindgren**[b], **J.B.Illian**[a,c]
S.T.Buckland[a], **H.Rue**[c] and **T. Gerrodette**[d]

[a]Department of Mathematics and Statistics
University of St Andrews
The observatory, Buchanan Gardens, St Andrews, UK
*yy84@st-andrews.ac.uk

[b]Department of Mathematical Sciences
University of Bath
Claverton Down, Bath, UK

[c]Department of Mathematical Sciences
Norwegian University of Science and Technology
Trondheim, Norway

[d]Southwest Fisheries Science Center
NOAA National Marine Fisheries Service
California, USA

Keywords: distance sampling, spatio-temporal modelling, stochastic partial differential equations, INLA, spatial point process

Abstract: Distance sampling is a widely used method for estimating wildlife population abundance. Conventional design-based distance sampling approaches allow us to estimate animal abundance in the study region without modelling the spatial distribution of animals, in the process of which the spatial information is often ignored. More sophisticated model-based approaches are in demand to study spatially referenced questions, such as the effect of spatially-varied covariates on the spatial distribution of animals and the underlying spatial structure of the surveyed population. We formulate distance sampling data as a thinned spatial point process and propose a model-based inference method that involves fitting a spatial log-Gaussian Cox process model to the survey data. Our method uses a flexible stochastic partial differential equation (SPDE) approach to account for the spatial autocorrelation, and the Integrated Nested Laplace Approximation (INLA) approach for Bayesian inference. We illustrate the method using distance sampling data from a series of shipboard line transect surveys in the eastern tropical Pacific (ETP).

A Bayesian hierarchical modelling approach to unravel large scale patterns of decline in the marine productivity of A. salmon population in the North Atlantic Ocean

M. Olmos [1], F. Massiot-Granier [1,2,3], E. Prévost [2], G. Chaput [4], E. Rivot [1]

[1] Agrocampus Ouest, UMR 0985 ESE INRA, Agrocampus Ouest, Ecologie et Santé des Ecosystèmes, Rennes, France.
etienne.rivot@agrocampus-ouest.fr

[2] INRA, UMR 1224 EcoBiop, Aquapôle, St Pée sur Nivelle, France

[3] IRSTEA, UR EABX, Aquatic Ecosystems and Global Changes, Cestas, France

[4] Fisheries and Oceans, Moncton, Canada

Keywords: Atlantic salmon, Bayesian Hierarchical Model, integrated population models, life cycle, population dynamics

Abstract: Integrated life cycle models are key tools for an ecosystem approach to fish population dynamics and stock assessment. They allow analysing ecological processes underlying the spatio-temporal variability of different life stages, together with the integration of multiple interacting sources of environmental and anthropogenic stressors in a hierarchy of spatial and temporal scales along the life cycle.

We developed a life cycle model for Atlantic salmon (*Salmo salar*) that captures the population dynamics of 13 regional and national stock units from the European (7 stock units from France to Iceland) and the North American coast (6 stock units from USA to Labrador) of the Atlantic Ocean.

The model is developed in the Bayesian Hierarchical Modelling framework and assimilates a 42-years time series of data (1970-2012) compiled by the International Council for the Exploration of the Sea (ICES). Its hierarchical structure provide a tool for separating out signals at different temporal (*e.g.*, year, decades) and spatial (*e.g.*, specific or shared by all the 8 stock complexes) scales in demographic traits. Hence, it improves the capacity to identify responses to key influential stressors associated with different scales.

Results show that both the survival during the first months at sea and age at maturity exhibit common trends shared by all units of stocks. The common trends explain more than 50% of the total variance. These results support a global response of salmon populations to broad scale ecosystem changes. It also supports the hypothesis of a synchronous collapse of marine survival with ecosystem changes observed in the North Atlantic in the early 1990s. Simultaneously, the proportion of early maturing salmon has increased, suggesting a change in life-history strategies in those populations. Ecological interpretations of these trends are discussed with regards to ecological literature on recent ecosystem changes in the Atlantic Ocean.

References

Massiot-Granier, F., Prévost, E., Chaput, G., Potter, T., Smith, G., White, J., Rivot, E. (2014). Embedding stock assessment within an integrated hierarchical Bayesian life cycle modelling framework: an application to Atlantic salmon in the Northeast Atlantic. ICES Journal of Marine Science, 71(7), 1653–1670.

Spatial prediction models for the probable biological condition of streams and rivers in the USA

E.W. Fox[a], **R.A. Hill**[b], **S.G. Leibowitz**[a], **A.R. Olsen**[a], and **M.H. Weber**[a]

[a]U.S. Environmental Protection Agency
National Health and Environmental Effects Research Laboratory
Western Ecology Division, Freshwater Ecology Branch
Corvallis, OR, USA
fox.ericw@epa.gov; leibowitz.scott@epa.gov; olsen.tony@epa.gov; weber.marc@epa.gov

[b]Oak Ridge Institute for Science and Education (ORISE)
Post-Doctoral Participant
Corvallis, OR, USA
hill.ryan@epa.gov

Keywords: National Rivers and Streams Assessment; random forests; Stream-Catchment Dataset; macroinvertebrate Multimetric Index

Abstract: The National Rivers and Streams Assessment (NRSA) is a probability-based survey conducted by the US Environmental Protection Agency and its state and tribal partners. It provides information on the ecological condition of the rivers and streams in the conterminous USA, and the extent to which they support healthy biological condition. An important problem is the prediction of stream integrity at new, unsampled locations. Using random forests (Brieman, 2001) we develop a model to predict the probability that a stream is in good (or conversely poor) biological condition. The model is fit to categorical response data consisting of 1365 NRSA survey sites and their designation as being in good or poor condition according to the macroinvertebrate Multimetric Index (MMI). The predictor data consist of over 200 GIS-level catchment and watershed variables from the EPA's Stream-Catchment Dataset (Hill et al., 2015). The out-of-bag performance of the random forest classifier is evaluated with classification rates, the area under the curve, and other graphical summaries. We find that random forest performs remarkably well according to these diagnostics, and is well-suited for modeling stream condition classes with a large predictor set. We also address issues with variable selection and model stability, and compare other statistical modeling approaches for modeling MMI with random forests.

The application of building the random forest model is the creation of maps displaying the predicted probability of good biological condition for all catchments within the NRSA sampling frame. The maps may have very useful applications towards watershed restoration and conservation efforts.

References

Breiman, L. (2001) Random forests. *Machine Learning*, 45(1), 5-32.

Hill, R.A., Weber M.H., Leibowitz S.G., Olsen A.R., and Thornbrugh D.J. (2015) The Stream-Catchment (StreamCat) Dataset: A Database of Watershed Metrics for the Conterminous United States. *Journal of the American Water Resources Association*, 1-9.

A Study to Determine Seasonal Effects of Transporting Fish from the Snake River to Optimize a Transportation Strategy

B.P. Sandford*, S.G. Smith and T.M. Marsh

Fish Ecology Division
Northwest Fisheries Science Center
NOAA Fisheries
Seattle, WA, USA
ben.sandford@noaa.gov

Keywords: salmon, barge transportation, seasonal effects, Poisson regression

Abstract: Studies have shown that the benefit of smolt transportation varies within migration seasons. Typically, transportation has been less beneficial for earlier migrating steelhead and spring/summer Chinook salmon. The objective of this study is to investigate within-season patterns in smolt-to-adult return rates (*SAR*s) of transported ("T") and in-river migrant fish ("B"), and patterns in the *SAR* ratio (T:B).

We considered all PIT-tagged wild and hatchery yearling Chinook salmon and steelhead released upstream from Lower Granite Dam from 1998 through 2012, along with fish collected and PIT-tagged at Lower Granite Dam. Because known dates of juvenile passage were required, inriver migrant groups were formed from PIT-tagged fish that were bypassed (i.e., detected and returned to the river) at Lower Granite Dam. We explored adjusting standards of comparison to account for lower *SAR*s generally observed for detected fish vs. never-detected fish. We used the statistical method of Poisson log-linear regression to model *SAR*s for daily groups of transported and migrant fish, and used AIC-based model averaging for multi-model inference.

Though we observed a wide variety of temporal patterns in *SAR*s and T:B ratios, transported fish had similar or higher *SAR*s in most cases with steelhead benefitting most. This benefit was generally reduced when compared to never-detected fish, though temporal patterns for the latter remain unknown. The location fish were tagged was a variable factor as well. The focus of our presentation will be results from more recent years.

Conditional occupancy modelling:
An alternative sampling design for rare and cryptic species

T. Arnold, M. Edwards, F. Iannarilli, M. Johnson, H. Specht, S. Stapleton, M. Weegman, and B. Yohannes

Department of Fisheries, Wildlife, and Conservation Biology
University of Minnesota
St. Paul, MN, USA
arnol065@umn.edu

Keywords: conditional sampling, detection probability, occupancy modelling, removal model.

Abstract: Although occupancy modelling is widely advocated for estimating occurrence of rare and cryptic species, simulation studies have indicated that occupancy models often perform poorly when true occurrence (ψ) and detection (p) are low ($\leq$ 0.4). We propose a conditional sampling design for rare and cryptic species, wherein investigators visit a larger sample of sites a single time and conduct k replicate surveys only at sites where organisms were detected during the first sampling round (i.e., repeat surveys are **conditional** upon detection in earlier rounds). This design assures that all replicate surveys provide data useful for estimating p, because there are no "wasted replicates" at unoccupied sites ($1 - \psi$). Better estimates of p allow for more precise estimates of occupancy at sites where organisms were undetected ($\hat{\psi}(1 - \hat{p})$); however, this approach reduces identification of sites known to be occupied (ψp^*). We contrasted our new proposed sampling design against traditional sampling (k replicates at all sites) and removal sampling (sampling until first detection or k replicates) using a range of true ψ (0.1, 0.2, 0.3, 0.4), true p (0.1, 0.2, 0.3, 0.4), k replicates (3, 5, 15), and cumulative sampling effort (sites × $\bar{k}$ = 120, 420, or 1500 total surveys). We assessed bias, precision, and mean square error (MSE) in $\hat{\psi}$ and $\hat{p}$, and measured overall performance based on A- and D-optimality of the variance-covariance matrix (Guillera-Arroita et al. 2010). For modelled values of ψ and p, removal sampling always underperformed alternative methods. Not surprisingly, conditional sampling provided the most reliable estimates of detection probability, but it also provided the most reliable estimates of occupancy when true ψ was $\leq$ 0.2. Traditional sampling designs provided the most reliable estimates of occupancy when true ψ was 0.3 or 0.4, but MSE for conditional sampling was only 3% larger and A- and D-optimality of ψ and p were 7% lower than traditional sampling when ψ = 0.3. Based on our simulations, we recommend conditional sampling designs (i.e., augmented sampling at sites where organisms were detected) whenever true occupancy is $\leq$ 0.3 and costs of sampling new sites approximate those of resampling previous sites.

References

Guillera-Arroita, G., Ridout, M.S., and Morgan, B.J.T. (2010). Design of occupancy studies with imperfect detection. *Methods in Ecology and Evolution*, 1:131-139.

A comparison of U.S. precipitation extremes under two climate change scenarios

M.J. Fix[a], **D. Cooley**[a], **S.R. Sain**[b], and **C. Tebaldi**[c]

[a]Department of Statistics
Colorado State University
Fort Collins, CO
miranda.fix@colostate.edu

[b]The Climate Corporation
San Francisco, CA

[c]Climate and Global Dynamics Laboratory
National Center for Atmospheric Research
Boulder, CO

Keywords: precipitation; climate change; generalized extreme value distribution; initial condition ensemble; pattern scaling

Abstract: Precipitation extremes are expected to increase in a warming climate, which may have serious socioeconomic and environmental impacts. This study leverages two initial condition ensembles conducted with the Community Earth System Model (CESM) to investigate potential changes in extreme precipitation under two climate change scenarios over the contiguous United States. We fit non-stationary generalized extreme value (GEV) models to annual maximum daily precipitation simulated from a 30-member ensemble under the RCP8.5 ("business-as-usual") scenario and a 15-member ensemble under the RCP4.5 (moderate mitigation) scenario. We then compare impacts using the 1% annual exceedance probability (AEP) level, which is the amount of daily rainfall with only a 1% chance of being exceeded in a given year. Under RCP8.5 between 2005 and 2080, the 1% AEP level is projected to increase by 17% on average across the U.S., and up to 36% for some grid cells. Compared to RCP8.5 in the year 2080, RCP4.5 is projected to reduce the 1% AEP level by 7% on average, with reductions as large as 18% for some grid cells. We also investigate a pattern scaling approach in which we produce predictive GEV distributions of annual precipitation maxima under RCP4.5 given only global mean temperatures for this scenario. We compare results from this less computationally intensive method to those obtained from our GEV model fitted directly to the CESM RCP4.5 output and find that pattern scaling produces reasonable projections.

Seasonal limitations and habitat functional responses in a spatially autocorrelated environment: Modelling winter distribution of white-tailed deer in Michigan, USA

S.T. O'Neil[a*], J.K. Bump[a], C.R. Webster[a], R. Doepker[b], S. Carson[b], and D.E. Beyer, Jr[b].

[a] Department of Forest Resources and Environmental Science
Michigan Technological University
1400 Townsend Ave, Houghton, MI 49931
*stoneil@mtu.edu

[b] Michigan Department of Natural Resources, Wildlife Division
1990 US-41 South, Marquette, MI, 49855

Keywords: generalized functional response, seasonal trade-off, species distribution

Abstract: Habitat selection and species distribution modelling rely on a comparison between occupancy (habitat use) and the overall availability of potential resources in a defined area. An existing challenge for habitat modelling involves the dynamic nature of the species-environment relationship. These dynamics could theoretically be modelled by observing temporal variation in resource availability and associated shifts in the species distribution, and can be approximated, known as a functional response in habitat selection (Mysterud & Ims 1998).

Habitat selection functional responses (HSFR) occur for a variety of reasons, including seasonal changes in the availability of food, cover, and conditions. Modelling the HSFR is especially important when vast differences exist in the distribution of resource availability between seasons; such is the case in the Upper Peninsula (UP) of Michigan, where heavy lake effect snows drive annual migrations of white-tailed deer (*Odocoileus virginianus*). Landscape drivers of deer winter range involve snow depth, elevation, and the presence of dense stands of conifers. However, these factors alone do not adequately explain the extent of deer wintering complexes (DWCs).

We fit a modified version of the generalized functional response (GFR) model (Aarts et al. 2013) to maps of deer winter occurrence across the UP. The model was based on a Bernoulli distributed response variable (presence/absence) and was fit to a suite of 13 environmental covariates, including 2nd order terms and interactions between 1st and 2nd order terms. We tested the influence of covariates at six spatial scales, using the best fitting scales of individual predictors in the final model. We updated the best fitting structure with a spatial random effect (conditional autoregressive, Besag et al. 1991). Models were fitted using Integrated Nested Laplace Approximations in R package INLA (Rue et al. 2009).

Latent Gaussian models accounting for spatial autocorrelation and functional responses in habitat selection resulted in excellent fit to deer winter occurrence data, indicating support for the modified GFR model and its specified interactions between food, cover, and conditions variables. Our results are useful for exploring the context of key drivers of seasonal habitats, and offer a framework for prediction under scenarios of climate and land cover change.

References

Aarts, G., J. Fieberg, S. Brasseur, and J. Matthiopoulos. 2013. Quantifying the effect of habitat availability on species distributions. Journal of Animal Ecology **82**:1135-1145.
Besag, J., J. York, and A. Mollie. 1991. Bayesian image-restoration, with 2 applications in spatial statistics. Annals of the Institute of Statistical Mathematics **43**:1-20
Mysterud, A., and R. A. Ims. 1998. Functional responses in habitat use: Availability influences relative use in trade-off situations. Ecology **79**:1435-1441.
Rue, H., S. Martino, and N. Chopin. 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. Journal of the Royal Statistical Society Series B-Statistical Methodology **71**:319-392.

Changing measurements of changing movements? Detecting movement patterns across multiple generations of tracking technology

L.R. Johnson[a], **P.H. Boersch-Supan**[a], **S.J. Ryan**[b], and **R.A. Phillips**[c]

[a]Department of Integrative Biology
University of South Florida
Tampa, FL, USA
lrjohnson@usf.edu, phb@usf.edu

[b]Department of Geography and Emerging Pathogens Institute
University of Florida
Gainesville, FL, USA
sjryan@ufl.edu

[b]British Antarctic Survey
Cambridge, UK
raphil@bas.ac.uk

Keywords: animal movement; tracking; albatross; maximum likelihood;

Abstract: Animal movement patterns can yield important information about foraging behavior and prey abundance and distribution. These are important components of understanding breeding behavior, juvenile and adult survival, and ultimately population dynamics. Over time, sensor technology for measuring movement patterns have changed and improved, particularly in terms of the temporal resolution of measurements. While these changes render old technologies obsolete, the older data themselves are still valuable, especially if new and old data can be analyzed together to ascertain if behaviors have changed over time. However, there are challenges in comparing animal movement and behavioral data across technologies that have aggregated measurements at different temporal scales, and it is unclear the extent to which this aggregation impacts conclusions that can be drawn from data.

To explore this conundrum, we combine a simulation study with an analysis of two suites of data on water landings collected a decade apart in 1992-1993 and 2002-2004, respectively, for two species of Antarctic albatrosses: Wandering (*Diomedea exulans*) and black-browed (*Thalassarche melanophrys*). We simulate foraging patterns under known step-length distributions, and aggregate these observations similarly to the sampling schemes that are utilized in the observational studies of the albatrosses. We then use a maximum likelihood approach to choose the best models under each sampling regime for the simulated data to evaluate our ability to correctly identify the distribution. We then use the likelihood-based approach to analyze the data for each species in both sets of data. Further we re-aggregate the more modern data to make a direct comparison with the historical set. This allows us to more closely examine the impact of the nature of the observational method on scientific conclusions and make suggestions for comparing and combining disparate data sets, to maximize their value.

Concentrations of heavy metals in estuarine water

A. Lim [a], C. Lueangthuwapranit [b] and *D. McNeil [c]

[a] Department of Mathematics and Computer Science
Faculty of Science and Technology, Prince of Songkla University
Pattani 94000, Thailand
apiradee.s@psu.ac.th

[b] Department of Technology and Industry
Faculty of Science and Technology, Prince of Songkla University
Pattani 94000, Thailand
chokchai.l@psu.ac.th

[c] Department of Statistics
Faculty of Science, Macquarie University
Sydney 2109, Australia
don.mcneil@mq.edu.au

Keywords: prevalence; estuarine; heavy metals; logistic and linear models; transformed data

Abstract: While some heavy metals including copper and zinc are essential for living organisms, others including lead and cadmium are toxic. We analysed nine heavy metals measured from monthly water samples during 2005-2015 taken from a tropical estuarine environment that serves as a major water source for irrigation, transportation, agriculture, aquaculture, fishing and recreation in a population exceeding 50,000 residents. One-litre samples were collected at ten sites between the river mouth and its source, and mercury, cadmium, copper, lead, zinc, chromium, nickel, manganese and iron concentrations were measured using atomic absorption. To estimate time-series trends, seasonal patterns and site differences for these concentrations, we separately fitted additive statistical models comprising these three factors to the data for each metal. To address statistical assumptions we first used logistic regression models to estimate prevalence and then fitted linear models to samples containing non-zero concentrations after transforming the data using power transformations to reduce skewness using square roots for all metals except iron (cube roots) and zinc (fifth roots). Results indicated that prevalence and non-zero concentration had similar seasonal and site patterns but different trends, thus justifying the decision to model them separately. Prevalence varied from 45-55% for lead and mercury to over 90% for iron and manganese, with respective maximum concentration levels in gm/L 0.0007 (Hg), 0.12 (Cd), 0.22 (Cu), 0.45 (Pb), 0.61 (Zn), 0.65 (Cr), 1.01 (Ni), 1.91 (Mn) and 8.4 (Fe).

Reference

Osman, A.G.M. and Kloas, W. (2010) Water quality and heavy metal monitoring in water, sediments, and tissues of the African catfish Clarias gariepinus from the River Nile, Egypt, *Journal of Environmental Protection*, 1(4):389-400.

*Corresponding presenter

Displaying water quality variables in tidal rivers

*N. Kongchouy [a], A. Lim[b] and D. McNeil[c]

[a]Department of Mathematics and Statistics
Faculty of Science,Prince of Songkla University
Hat Yai 90110, Thailand
nootchanath.k@psu.ac.th

[b] Department of Mathematics and Computer Science
Faculty of Science and Technology, Prince of Songkla University
Pattani 94000, Thailand
apiradee.s@psu.ac.th

[c]Department of Statistics
Faculty of Science, Macquarie University
Sydney 2109, Australia
don.mcneil@mq.edu.au

Keywords:water quality;graphical methods; factor analysis; correlation matrix

Abstract: Water quality plays acrucial role in the life, growth, and reproduction of living organisms. This study aimed tomodel water quality parameters and illustrate results obtained from samples usingeasily understandable graphs.Using standard methods, sampleswere collected monthly during 2005-2015 from tenupstream sites in the Na Thap tidal river that flowsinto the Gulf of Thailand. Thirty physicochemical variables of water quality were measured. A factor analysis model with promax rotation was fitted by maximum likelihood to the correlation matrix of these variables, after applying appropriate transformations to make distributions approximately normal.Five interpretable factors resulted and were illustrated using a bubble-plot correlation matrix. Linear models were fitted to selected variablesfrom each factor. Quantile-quantile probability plots indicated plausibility of normality assumptions, and to illustrate goodness-of-fit, observed data were compared with corresponding fitted values using scatter plots with fitted lines, using points of variable size ("polliwog" symbols) to account for transformation bias. Model results were displayed by graphing "democratic" confidence intervals for comparing levels of each factor with overall means (Sampantarak et al., 2011). Seasonal and location patterns and time-series trendsare highlighted using these graphs, and confounding and transformation bias are denoted by plotting additional points on the same axes.

Reference

Sampantarak, U., Kongchouy, N. and Kuning, M. (2011) Democratic confidence intervals for adjusted means and incidence rates. *American International Journal of Contemporary Research,* 1(2):38-43.

*Corresponding presenter

Bats in an 'Ecological Desert' -- Activity and Abundance of Bats in Commercial Coniferous Plantations

L. Kirkpatrick[1], Paul Lintott[3], D. Dent[1], S. Bailey[2], K. Park[1]

[1] Biological and Ecological Sciences Department,
University of Stirling,
Stirling, UK
Lucinda.kirkpatrick@stir.ac.uk

[2] Forestry Commission,
UK

[3]University of Exeter
Biosciences
Exeter
UK

Keywords: Chiroptera, Plantation management, multimodal, landscape ecology

Traditionally plantations have been managed for both timber and non timber resource extraction, and have rarely been investigated for biodiversity, particularly for mammalian species. Bats, which constitute one fifth of mammal species worldwide, have undergone large declines throughout Europe, most likely due to widespread habitat loss and degradation. While bat use of modified landscapes such as urban or agricultural environments has been well studied, intensively managed plantation woodlands have received less attention, particularly in Europe.

We assessed 286 stands across three of the largest, most intensively managed plantations in the UK for bat presence, activity and abundance in response to environmental characteristics at multiple spatial scales using an information theoretic approach (Burnham & Anderson, 2002). The influence of environmental characteristics on bat activity varied by species and genus, and appears to be constrained by structural complexity rather than prey availability. Occurrence and activity of clutter and edge adapted bats was associated with decreasing stand density and a more heterogeneous landscape whereas occurrence and activity of open adapted bats was highest at felled stands and decreased at detector points closer to mature conifer woodland. In addition the cryptic *Pipistrellus* species used different plantation areas based on foraging style despite morphological similarities, allowing both species to use plantation areas. This study demonstrates that with sympathetic management, previously under surveyed non native conifer plantations may have an important role in important for maintaining and improving bat populations, particularly for *Pipistrellus spp.*

Burnham, K., & Anderson, D. (2002). *Model selection and multimodel inference: a practical information theoretic approach*. Springer Science and Business Media.

Local and landscape scale influences on bat abundance, activity and behavior in native woodlands and exotic plantations

Lucinda Kirkpatrick[1], Kirsty Park[2]

[1]Presenting author:
Biological and Ecological Sciences,
University of Stirling
Scotland
Lucinda.kirkpatrick@stir.ac.uk

[2] Biological and Ecological Sciences,
University of Stirling
Scotland

Keywords: Chiroptera, plantation management, behaviour, landscape ecology

Abstract:

Non native, managed coniferous plantations have long been considered a poor or suboptimal habitat for bat species, with many studies demonstrating active avoidance in favour of other habitats such as native broadleaved woodlands. As yet however, no studies have explicitly compared coniferous plantations to broadleaved woodlands to determine the extent as to which bat use of plantations is a product of "spill over" from surrounding more appropriate habitat. Coniferous plantations comprise the majority of remaining forest land cover in Europe and understanding what influences diversity within them is key for effective management plans; as bat species represent one fifth of mammalian diversity in Europe and the majority rely on forests to at least some degree, understanding the processes behind bat use of different habitat types is vital for their effective conservation.

In this study fourteen paired coniferous plantation and native broadleaved sites were compared to determine whether the species composition and quality of woodland influences activity and abundance of bat species. All species whose geographical range encompassed the study area were caught or recorded in both plantation and native sites although there were gender and reproductive status differences between the two forest types, with native woodlands supporting a wider diversity of bat species. Activity was similar in coniferous plantations compared to native woodlands but differed between the two years, due to climatic conditions. Bat use of both plantation and native woodlands was influenced on both the local and the landscape scale and showed temporal differences between species. This demonstrates that plantations are not intrinsically poor habitats for all bat species, but with the correct management both native and plantation woodlands may have a key role to play in supporting bat populations, particularly in areas dominated by large, highly productive coniferous plantations.

Species traits, physical environment, and community interactions influence responses of wetland-breeding amphibians to environmental change

Courtney Davis[1], David Miller[1], Susan Walls[2], William Barichivich[2], Jeffrey Riley[3] and Mary Brown[4]

[1]Pennsylvania State University, [2]U.S. Geological Survey, Wetland and Aquatic Research Center, [3]U.S. Geological Survey, South Atlantic Water Science Center, [4]Cherokee Nation Technology Solutions contracted to the U.S. Geological Survey, Wetland and Aquatic Research Center

Of the myriad of issues facing amphibian populations worldwide, the direct and indirect effects of changing climate are among those most difficult to isolate and predict. Climate does not act on species in isolation. Instead the effects of climatic variability will be driven by interactions between climate, co-occurring species, and the physical environment in which interactions occur. Changes in regional precipitation patterns may be especially important for wetland breeding amphibians because of effects on wetland dynamics. We examined how 3 species of wetland breeding amphibians with varying life-history requirements responded to extreme weather events at St. Marks National Wildlife Refuge from 2009–2014, which was punctuated by periods of both severe drought and extensive flooding that facilitated colonization of fish to many previously unoccupied wetlands. We used dynamic multistate occupancy models that incorporate habitat dynamics and species interactions fit to observational data collected spring and fall of each year, to investigate how metapopulation dynamics (colonization and persistence) responded to fluctuating conditions during this period. Increased colonization rates of ornate chorus frogs (*Pseudacris ornata*), a species that preferentially selects ephemeral wetlands to breed, during drought events in 2010–2011 resulted in occupancy increases from 10% to 34% of short duration wetlands. Mole salamander (*Ambystoma talpoideum)* occupancy, however, decreased from 37% to 25% in long-duration wetlands because of reduced persistence probabilities that resulted from reduction of mean wetland hydroperiod during the drought. Occurrence of mole salamanders over the course of the study dropped from 39% to 19%. For both ornate chorus frogs and mole salamanders, colonization of wetlands by fish during flood events in later years reduced persistence probabilities and wetlands occupied by fishes were rarely colonized. In contrast, the pig frog (*Rana grylio*) showed high resilience to perturbations, with colonization and persistence being largely unaffected by changes in water availability or fish presence. Our results demonstrate that similar changes in water availability will have very different outcomes depending on the species of wetland breeding amphibian, and that these outcomes result from differences in habitats currently used, and responsiveness to changes in hydroperiod and to predation by fishes. Our simultaneous modeling of habitat and species annual dynamics provides a framework for disentangling the processes determining how species will respond to environmental change.

Shearwater monitoring in northwestern Mediterranean from opportunistic datasets: population trends, spatial patterns and global efficiency

N. Di Méglio [a], L. David [a] and P. Monestiez [bc]

[a] EcoOcéan Institute,
Montpellier, France
ecoocean@wanadoo.fr

[b] INRA, Biostatistics and Spatial Processes
Avignon, France
[c] CNRS, CEBC UMR 7372, Centre d'Etudes Biologiques
de Chizé, France
monestiez@avignon.inra.fr

Keywords: species distribution modeling, Distance Sampling, Poisson Kriging, Citizen sciences, Calonectris borealis, Puffinus yelkouan.

Abstract: Long term monitoring of marine species, with abundance indices, spatial distributions and feeding areas, gives essential information for conservation purpose or marine protected area design. We show, with the example of two seabird species - Cory's shearwater and Yelkouan shearwater - that opportunistic data from scientific surveys and citizen science programs can produce quite accurate maps and trend estimates. Data were heterogeneously distributed in time with four periods 1994-1996, 1997-2002, 2006-2008, 2010-2014 and based on a rigorous data collection protocol of DS (Distance Sampling) from opportunistic platforms and volunteer observers. We combined a bias correction from the fitted detection curve (DS) and a Poisson Kriging to account for the heterogeneity of the observation effort in space. We highlighted preferential feeding areas for both species with persistence over the four time periods and a population roughly stable over 20 years. The results are close to those from much costly scientific surveys (aerial survey, loggers-equipped birds) when comparable. Such opportunistic-data-based monitoring may represent a complement to large institutional programs with a better continuity in time. Added to other data sets, and with adapted statistical modeling, they may also provide information for conservation program or design of MPAs.

Uncertainty associated to the new French macroinvertebrate-based multimetric index (I_2M_2): towards a probabilistic ecological indicator

E. Billoir [a]*, R. Coulaud [a], F. Larras [a], V. Archaimbault [b], C. Mondy [c], P. Usseglio-Polatera[a]

[a] Interdisciplinary Laboratory for Continental Environments (LIEC)
University of Lorraine and CNRS (UMR 7360)
Campus Bridoux, Metz, France
elise.billoir@univ-lorraine.fr

[b] Irstea
Antony, France

[c] LEHNA
Université Lyon 1 and CNRS (UMR 5023)
Villeurbanne, France

Keywords: uncertainty assessment; ecological indicator.

Abstract: To meet the Water Framework Directive requirements, a new multimetric index (I_2M_2) has been designed for the invertebrate-based ecological assessment of French wadeable streams (Mondy et al., 2012). Integrating five taxonomy- and trait-based metrics selected for their high discrimination efficiency, low pressure specificity, high stability in least impaired conditions and low redundancy, this index is meant to identify impaired reaches for 17 anthropogenic pressure categories potentially impairing water quality or habitat. Based on I_2M_2, any river reach is so far assigned a unique ecological quality class among “Bad”, “Poor”, “Moderate”, “Good” and “High”. Here we aimed to assess the uncertainty that should be associated to the index, due to both (1) the *construction* of the I_2M_2 and (2) the successive normalized *field and lab* steps preceding its calculation (e.g. substrate mosaic description, field sampling, taxonomic identification, organism counting). *Construction* uncertainty was estimated by bootstrap sampling in the ‘development’ dataset. *Field and lab* uncertainty was estimated through an *ad hoc* field and lab experiments involving 4 operators who all performed the I_2M_2 protocol (AFNOR, 2009, 2010) on 24 sites for 2 years (1 campaign/year). Uncertainty of I_2M_2 was then assessed by Monte Carlo simulations accounting for c*onstruction* uncertainty, *field and lab* uncertainty, or both. This enabled to evaluate the relative importance of the uncertainty sources and to construct a model for the uncertainty that should be associated to I_2M_2 calculations. With uncertainty assessment any river reach can be assigned a probability distribution among the quality classes. Such a probabilistic indicator will better inform on the robustness of ecological assessment and will be much more helpful in identifying river management and restoration priorities.

References

Mondy, C., Villeneuve, B., Archaimbault, V. and Usseglio-Polatera, P. (2012) A new macroinvertebrate-based multimetric index(I2M2) to evaluate ecological quality of French wadeable streams fulfilling the WDF demands : a taxonomical and trait approach. *Ecological Indicators* 18: 452-467.

AFNOR. (2009) Qualité de l'eau - Prélèvement des macro-invertébrés aquatiques en rivières peu profondes. Norme homologuée XP T90-333, AFNOR, Paris.

AFNOR. (2010) Qualité de l'eau - Traitement au laboratoire d'échantillons contenant des macro-invertébrés de cours d'eau. Norme homologuée XP T90-388, AFNOR, Paris.

Brazilian humpback whale population size from design-based and model-based estimates

G.A. Bortolotto[a *], P.S. Hammond[b], A.N. Zerbini[c] and L. Thomas[d]

[a,b] Sea Mammal Research Unit
University of St Andrews
St Andrews, Fife, Scotland
&
[a,b,d] Centre for Research into Ecological and Environmental Modelling
University of St Andrews
St Andrews, Fife, Scotland
[a] gabdo@st-andrews.ac.uk
[b] psh2@st-andrews.ac.uk
[d] len.thomas@st-andrews.ac.uk
[c] Cascadia Research Collective & National Oceanic and Atmospheric Administration
Seattle, U.S.A.
alex.zerbini@noaa.gov

Keywords: distance sampling; density surface model; precision, abundance

Abstract: Animal abundance estimates are necessary to assess population status to support management or conservation actions. For cetacean species, a common way to do this is through Distance Sampling (DS; Buckland *et al.* 2001), which estimates depend on the survey design. Density Surface Models (DSM; Miller *et al.* 2013) may also be used to estimate animal abundance and it does not necessary need designed survey data to be performed, relying on model-based estimates. Abundance estimates for humpback whales in the coast of Brazil were performed with DS before, but DSM was never employed for that. Using data from a shipboard survey to estimate their abundance in 2008, estimates from both design-based and model-based methods were calculated in order to assess potential differences across them. Additionally, the realized effort did not fully meet the assumptions of design-based methods (e.g. uniform coverage probability) and the model-based estimates may assist in explain spatial variability. Covariate Beaufort sea scale was retained in the best detection function model according to the Akaike Information Criterion. The design-based approach resulted in an estimated abundance of 15,705 (CV = 0.232) whales. For the model-based estimate, transects were segmented in 290 units of 8 km, with a total effort of 2,340 km. A prediction grid with 1915 cells (8 km^2 each) was built over a 122,560 km^2 survey area. GAMs were performed with spatial covariates "easting", "northing" (computed from longitude and latitude, respectively), depth and survey stratum to model the abundance at each segment. The model selected by the smallest GCV score included geographic position and depth. Summed predicted abundance for the prediction grid resulted in 14,476 (CV = 0.073) whales. Point estimates were similar but the model-based abundance estimate was more precise than the design-based one. Features from the design of the survey must have contributed to increase the variation in the results of the latter method. In fact, variation in encounter rate was the main responsible for its higher variance. This study supports that the spatial covariates used in the model-based estimates explain a considerable portion of the variability in the data, and that the method may be used estimate humpback whales abundance in Brazil more precisely.

References

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L. and Thomas, L. (2001) *Introduction to distance sampling: Estimating abundance of biological populations*. Oxford University Press, Oxford. 488 p.

Miller, D.L., Burt, M.L., Rexstad, E.A. and Thomas, L. (2013) Spatial models for distance sampling data: Recent developments and future directions. *Methods in Ecology and Evolution*, 4: 1001-1010.

Temporal variation in willow ptarmigan detectability: research and management implications

G.G. Frye [a,] M.S. Lindberg [b], and R.A. Merizon [c]

[a] Institute of Arctic Biology, Department of Biology and Wildlife, and Department of Mathematics and Statistics
University of Alaska Fairbanks
Fairbanks, Alaska
ggfrye@alaska.edu

[b] Institute of Arctic Biology and Department of Biology and Wildlife
University of Alaska Fairbanks
Fairbanks, Alaska
mslindberg@alaska.edu

[c] Small Game Program
Alaska Department of Fish and Game
Palmer, Alaska
richard.merizon@alaska.gov

Keywords: willow ptarmigan; abundance; Poisson-binomial mixture model; indices; playback surveys

Ptarmigan (*Lagopus* spp.) are ecologically and culturally important species throughout their circumpolar ranges. As such, rigorous monitoring and management strategies are needed for ptarmigan populations. A fundamental component of monitoring and management efforts is the ability to obtain reliable estimates of population states and vital rates. A variety of survey techniques have been used to estimate and index ptarmigan abundance, but the accuracy of the data produced by many of these techniques is unknown. An unavoidable complication shared by most of these survey techniques is an imperfect detection process, which has the potential to yield inaccurate estimates, inferences, and predictions. We examined the role of imperfect detection in playback surveys for Willow Ptarmigan (*Lagopus lagopus*) in interior Alaska. We used temporally and spatially replicated counts to model the detection process. We fit Poisson-binomial mixture models in a Bayesian framework to estimate the probability of detecting willow ptarmigan as a function of (1) time of season and (2) time of day. Our results indicate negative effects of both of these temporal covariates. These results suggest that using raw counts from ptarmigan playback surveys is an unreliable approach to estimation. Similarly, simulations suggest that using playback indices to infer trends may be misleading, given the diel and seasonal variation in detection probability. We recommend that willow ptarmigan monitoring efforts account for heterogeneity in the detection process prior to drawing inferences or making predictions on the basis of playback survey data.

Ocean and Freshwater Influences on Columbia River Sockeye Salmon Adult Return Rates

S.G. Smith [a], J.G. Williams [a*], J.F. Fryer [b], M.D. Scheuerell [a], W.D. Muir [a*],
T.A. Flagg [a], R.W. Zabel [a], J.W. Ferguson [a*], and E. Casillas [a*]

[a] NOAA Fisheries
Northwest Fisheries Science Center
Seattle, Washington, USA
steven.g.smith@noaa.gov

[b] Columbia Inter-Tribal Fisheries Commission
Portland, Oregon, USA

* Retired

Keywords: sockeye salmon; smolt-to-adult survival; multi-model inference; Columbia River.

Abstract: Recently, adult sockeye salmon *Oncorhynchus nerka* returns to the Columbia River Basin have reached numbers not observed since the 1950s. To understand factors related to returns, we evaluated productivity changes by estimating smolt-to-adult return rates (SAR) for juvenile migration years 1985-2010. SAR varied between 0.2 and 23.5%. Highest values coincided with recent large adult returns. However, the largest return, in 2012, resulted not from increased survival, but from increased smolt production. We evaluated 19 different variables that could influence SARs, representing different freshwater and ocean conditions. We used small-sample corrected AIC to select among 2- and 3-variable models. The model with April ocean upwelling, Pacific Northwest Index (PNI) in the migration year, and PNI in the year before migration had 10 times the AIC_c-weight as the second-best-supported model, and R^2 of 82%. The variables April ocean upwelling and PNI in the migration year were by far the best of the candidate variables to explain variations in SAR. While our analyses were primarily correlative and limited by the type and amount of data available, changes in ocean conditions in the northern California current system, as captured by April upwelling and PNI, appeared to play a large role in the variability of SAR.

References

Williams, J.G., S.G. Smith, J.F. Fryer, M.D. Scheuerell, W.D. Muir, T.A. Flagg, R.W. Zabel, J.W. Ferguson, and E. Casillas. 2014. Influence of ocean and freshwater conditions on Columbia River sockeye salmon *Oncorhynchus nerka* adult return rates. *Fisheries Oceanography*, 23:210-224.

Impact assessment based on acoustic detections of marine mammals – a comparison of analysis techniques

B. Woelfing[a]

[a] Institute for Terrestrial and Aquatic Wildlife Research
University of Veterinary Medicine Hannover, Foundation
Germany
Benno.Woelfing@tiho-hannover.de

Keywords: impact assessment; acoustic survey; marine mammal; generalized additive model

Abstract:

The growth of the marine renewables industry has led to considerable effort to evaluate its effect on marine mammal species. Long-term time series of echolocating species presence can be obtained by recording their echolocation clicks. Using simulated scenarios as well as data from a German wind farm I compare different survey designs (placement of click detectors) and methods of analysis. I evaluate the use of different response variables: waiting time to the next click versus the number of minutes with detections per hour. Recommendations on the survey design and method of analysis will be given based on marine mammal density, the magnitude of natural fluctuations in marine mammal density and the presence of ecological gradients in the study area.

Sperm whale presence in the central and western North Pacific: assessing activity patterns in auto-correlated data using Generalized Estimating Equations and Generalized Additive Models

K. Merkens [a,b] and E. Oleson [a]

[a] NOAA NMFS Pacific Islands Fisheries Science Center
Honolulu, Hawaii
karlina.merkens@noaa.gov

[b] Lynker Technologies
Portland, Oregon

Keywords: sperm whale, auto-correlation, GEE, GAM

Abstract: Sperm whales (*Physeter macrocephalus*) are found throughout the world's oceans. Their ecological role as long-lived, top predators, and a history of intense harvest, makes them important species for monitoring. In the central and western North Pacific Ocean we deployed passive acoustic recording devices (High-frequency Acoustic Recording Packages, HARPs) at 13 locations to monitor presence of multiple cetacean species. These instruments produce a remarkable data set that is both long-term and highly detailed, and is therefore unique in the field of marine mammal biology. The data set was scanned by human analysts for detections of echolocation clicks of sperm whales, and the presence of these clicks was modelled using multiple parameters including sub-region and recording site, as well as daily, lunar and seasonal variables. Two different approaches were used to account for high levels of auto-correlation among the samples: Generalized Estimating Equations (GEEs) applied to finely binned data and Generalized Additive Models (GAMs) applied to more coarsely binned data. The results were consistent between the two models, indicating fairly robust patterns of presence over location and time. Sperm whales were present at all sites, but rates of presence ranged by site from 1-15% of time monitored. Presence across the whole region varied slightly on a seasonal cycle, but lunar cycles did not help explain any variability. These results expand our current understanding of the locations of sperm whales across the central and western North Pacific Ocean and demonstrate the usefulness of both GEE and GAM methods for analysing highly auto-correlated data.

Statistics for species-pair correlation with presence-absence records

K. Mainali [a], S. Bewick [a] and W. Fagan [a]

[a] Department of Biology
University of Maryland
College Park, MD, USA
kpmainali@gmail.com
sharon_bewick@hotmail.com
bfagan@umd.edu

Keywords: Pearson’s correlation, Jaccard’s index, independence, chi-square.

Abstract: Taxa pair correlation in spatial distribution has been historically estimated with Pearson’s correlation coefficient (*r*) for binary outcome (equivalent to a phi coefficient) and Jaccard’s index of similarity (*J*). These two metrics are different in variables, parameters and model structure. In this study, we examine important problems in statistical analyses of species co-occurrence data that have plagued correlative ecological and microbiome research; we use simulation and analyse human skin microbiome data. *J* and *r* have strong mismatch in determining the direction of relationship between taxa (positive or negative) and in identifying the significantly correlated taxa-pairs. We then explore the source of these discrepancies: specifically, we show how Pearson’s correlation is extremely sensitive to the relative frequency of co-absent sites and how co-absent sites can be non-informative and misleading for determining correlation, especially in microbiome analysis. Jaccard’s index is free of this problem. We determine statistical significance of *J* with binomial theorem and multinomial coefficient. We then show that Jaccard’s index can inflate false positives by its failure to correctly identify some situations of independence (of species pair) as determined by probability theory. We propose a remedy that Jaccard’s index should be filtered with another statistical test to determine independence such as chi-square test.

Partial Stratification in Two-Sample Capture-Recapture Experiments and Optimal Allocation of Sampling Effort

Lasantha Premarathna[a] and **Carl Schwarz**[b]

Department of Statistics and Actuarial Science
Simon Fraser University
8888 University Drive
Burnaby, BC, Canada V5A 1S6
[a] wpremara@sfu.ca [b] cschwarz@stat.sfu.ca

Keywords:: abundance; capture-recapture; Heterogeneity; MCMC; optimal Allocation; survey-design and analysis

Abstract: In closed populations, capture heterogeneity cause bias in estimate of abundance in two-sample capture-recapture (Lincoln-Petersen) experiments. Heterogeneity is related to observable fixed characteristic of the animal such as sex. If the each sampled animal can be stratified, then it is straightforward to obtain stratum-specific estimates. In many fishery experiments it is difficult to stratify (e.g. by sex) all the captured animals. In these cases a sub-sample of the captured animals at each sampling occasion is selected and further, more costly, measurements (e.g. sex determination) are made. Our data now consists of three parts-animals whose stratification variable is unknown, and sub-samples at each occasion where stratification variables are determined.

In this study we develop and apply new methods for these types of experiments using the method of MLE. In these types of experiments, various costs are associated such that the fixed cost for the study, cost to capture a sample of animals in each occasion and also cost to process the sub-samples. Furthermore, given the relative costs, the optimal allocation of sampling effort for a given cost is determined. Also we develop a method using Bayesian approach to account for additional information (e.g. prior information about the sex ratio). Moreover we develop methods for supplemental continuous covariates such as length using MLE and Bayesian approach. An illustrative example is presented applying these new methods to estimate the size of the walleye population in MilleLacs, MN.

Model averaging with interactions can be misleading

P. Dixon[a] and **J. Niemi**[a]

[a]Department of Statistics
Iowa State University
Ames IA, USA
pdixon@iastate.edu

Keywords: BACI design, impact assessment, posterior model probabilities, AIC

Abstract: We evaluate the consequences of model averaging when the quantity of interest is an interaction coefficient. Our focus is the analysis of data from a Before-After-Control-Impact (BACI) study design, which is often used to assess the impact of an ecological perturbation. In the simplest version, data are collected from two sites, one impacted by the perturbation; the other not, in two periods, one before the perturbation; the other after. Such data are commonly analyzed using a model with indicator variables for site, period, and their interaction. The effect of the perturbation is quantified by the magnitude of the interaction. Straightforward application of model averaging will consider 8 models, one for each combination of the three indicator variables.

Model averaging over 8 models is misleading because the posterior model probabilities depend on the coding of the indicator variables. For one example data set, the posterior probability of a non-zero interaction is 0.75 when the indicator variables are coded with 0/1 values and 0.36 when the indicator variables are coded with -1/+1 values. This problem arises because different sets of indicator variables are not equivalent (do not span the same subspace) for some of the 8 models being averaged.

Model averaging results are estimable, in the sense of being invariant to the choice of coding, when the set of models enforces the hierarchy principle. That is, inclusion of an interaction term in a model forces the inclusion of each component. For the two period x two site BACI design, there are 5 models in the hierarchical set: intercept only, period only, site only, period and site, and period, site, and their interaction.

We illustrate these ideas using both frequentist and Bayesian model averaging for studies of poison bait impacts on non-target organisms and oil spill impacts on sea-turtle nesting.

Modeling non-constant detection rates in removal sampled point-count surveys

A. Martin-Schwarze, **P. Dixon**, and **J. Niemi**

Department of Statistics
Iowa State University
Ames, Iowa, USA
adamms@iastate.edu

Keywords: abundance; N-mixture model; survival analysis; Bayesian

Abstract: Abundance estimates from animal point count surveys require accurate estimates of detection probabilities. The standard model for estimating detection from removal-sampled point count surveys assumes that organisms at a survey site are detected at a constant rate (Farnsworth et al. 2002; Amundson et al. 2014). However, the assumption of constant detection rates during surveys is often not justified. Detection rates can be influenced by organismal behaviors, including responses to observer presence and heterogeneity among population subgroups, and also by survey methods that affect observer effort. Failure to account for non-constant detection leads to biased estimates of detection and therefore of abundance.

Under the standard approach, detection probability is modeled by dividing the observation period into equal-duration intervals. We instead model the detection *rate* in continuous time via a time-to-detection distribution embedded within a hierarchical N-mixture framework using Bayesian methods. Our model is thus a combination of survival time-to-event analysis with unknown-N, unknown-p abundance estimation.

We apply this model to Ovenbird counts from the Chippewa National Forest and to datasets simulated under different time-to-detection patterns.

Models assuming constant detection rates produce biased estimates of detection when true detection rates vary with time, whereas models allowing for variable detection (assuming gamma, Weibull, or lognormal distributed times to detection) produce less biased estimates of the detection probability and nominal credible interval coverage. Models ignoring detection heterogeneity across subgroups yield biased estimates of detection when such heterogeneity exists, whereas models accounting for detection heterogeneity (modeled as a mixture) return reasonable coverage rates and can outperform heterogeneity-ignorant models even when there is no heterogeneity.

References

Amundson, C. L., Royle, J. A., and Handel, C. M. (2014). A hierarchical model combining distance sampling and time removal to estimate detection probability during avian point counts. *The Auk* 131, 476-494.
Farnsworth, G. L., Pollock, K. H., Nichols, J. D., Simons, T. R., Hines, J. E., Sauer, J. R., and Brawn, J. (2002). A removal model for estimating detection probabilities from point-count surveys. *The Auk* 119, 414-425.

Using Bayes' Theorem to Provide Evidence of Compliance with Incidental Take Permits

M. M. P. Huso [a] and D. H. Dalthorp [a]

[a] Forest and Rangeland Ecosystems Science Center
U. S. Geological Survey
Corvallis, Oregon USA 97331
mhuso@usgs.gov
ddalthorp@usgs.gov

Keywords: Incidental Take Permit, Endangered Species Act, Bayes Theorem, Evidence of Absence, Horvitz-Thompson

Abstract: The U.S. Fish and Wildlife Service has begun issuing 20- and 30-year permits to wind power developers for take under the Endangered Species Act and/or the Bald and Golden Eagle Protection Act, e.g., Indiana bat, Newell's shearwater, golden eagle, etc. Once a take limit is set and minimization and mitigation approaches agreed upon, additional actions may be necessary if the permitted take limit is exceeded. Accurately collecting and interpreting data to determine that take is within permitted limits can be quite difficult. When the target population is small, as might be expected for endangered species or species with low population densities, the likelihood of finding no carcasses may be high, yet observing no carcasses cannot necessarily be interpreted to mean zero or even low numbers of fatalities (Huso et al. 2015). Horvitz-Thompson type estimators can only return 0 with zero variance when no carcasses are observed. We present an approach based on Bayes' theorem that uses the observed carcass count (X) and information about the search process and estimated detection probabilities to provide posterior probabilities of the actual fatality being X, $X+1$, $X+2$, ... We review the basic functionality of the Evidence of Absence software (Dalthorp et al. 2014) developed to carry out its complex calculations. The ultimate purpose of this software is to give managers tools for designing monitoring programs to provide evidence of industry compliance with Incidental Take Permits. We describe modules currently under development that will help define decision points where evidence of compliance is insufficient or where adaptive management actions should be considered to reduce excessive take. These products will be useful for any species of concern in any region, from Golden eagles to black-capped vireos.

References

Dalthorp, D. H., M. M. P. Huso, D. Dail, and J. Kenyon. 2014. *Evidence of Absence Software*. Corvallis, OR: USGS. ISBN. Available online at http://pubs.usgs.gov/ds/0881/.

Huso, M. M. P., D. Dalthorp, D. Dail, and L. Madsen. 2015. Estimating turbine-caused bird and bat fatality when zero carcasses are observed. *Ecological Applications* 25:1213-1225.

Efficient, flexible estimation of time to decay of signs in indirect survey methods

P.T. Besbeas [a, b] and C.S. Nisiotis [a]
[a] Athens University of Economics and Business
Athens, Greece

[b] University of Kent
Canterbury, Kent, UK
ptb@kent.ac.uk

Abstract:

Surveys of rare or elusive species often rely on indirect signs such as dung or nests. Sign density is usually translated into animal density using sign production and decay rates. Traditional decay rate estimation methods entail following many signs from production to disappearance, or to a single follow-up visit establishing whether the sign has decayed or not. We focus on the latter sampling design, which is more time and cost effective in practice. For this design, Laing et al (2003) proposed a two stage estimation approach based on logistic regression, with mean time to decay estimated in the second stage. We propose a new framework for estimating the distribution of time to decay directly, in a single stage. The new framework identifies methodological issues that arise in the logistic regression approach. The new method is shown to be more efficient in general than Laing et al (2003) in simulation studies. In addition, the framework is more flexible, adapts to a parametric or non-parametric specification, and can accommodate data from a range of sampling designs. The method is illustrated by application to a population of roe deer *(Capreolus capreolus)* living in a mountainous Mediterranean area.

References

Laing et al (2003). Dung and nest surveys: estimating decay rates.
J. Appl. Ecol. 40, 1102-1111.

Tsaparis et al (2009). Estimating dung decay rates of roe deer (Capreolus capreolus) in different habitat types of a Mediterranean ecosystem: an information theory approach.
Europ. J. Wildl. Res. 55, 167-172.

New removal approaches for reptile and amphibian populations

Ming Zhou[a], Rachel S. McCrea[a], Eleni Matechou[a], Diana Cole[a], Richard A. Griffiths[b]

[a] School of Mathematics, Statistics and Actuarial Science
University of Kent, Canterbury, UK
mz214@kent.ac.uk; R.S.McCrea@kent.ac.uk; E.Matechou@kent.ac.uk; D.J.Cole@kent.ac.uk;

[b] Durrell Institute of Conservation and Ecology
University of Kent, Canterbury, UK

R.A.Griffiths@kent.ac.uk

Keywords: abundance; population dynamics; integrated population models; survey design and analysis.

Abstract: Classical removal experiments – the movement of protected animals or plants out of the path of development projects as a wildlife management tool – can be used to estimate the abundance of a population within a closed area. The classic removal model (Moran, 1951) assumes a constant capture probability and all animals are available for detection throughout the study, which results in a simple geometric decline of removed counts of individuals over time. However, the real data collected from some species exhibit unexpected fluctuations in the number of captured animals. The work is driven by real data on slow worms, *Anguis fragilis*, common lizards, *Zootoca vivipara* and great crested newts, *Triturus cristatus*, where existing approaches may give rise to misleading conclusions.

Reptiles and amphibians sometimes become undetectable as they may hide underground. This phenomenon can be modelled as a partial hidden process, where the underlying state process describes the movement pattern of animals between survey area and the area outside the study. We have developed an adaptive Hidden Markov Model with Robust Design which allows considerable flexibility in estimating demographic parameters, including estimates of transitions between underlying states and the size of populations. Moreover, the activity of reptiles and amphibians is highly sensitive to temperature and weather conditions. Therefore, the model is also extended by the incorporation of climatic covariates to account for time-varying detection probabilities. Comparisons are made with estimates obtained from fitting the traditional model, and it demonstrates that the performance of the new model is better under many ecological scenarios. We also consider the effect of sparse data and investigate the use of modelling different sources of data in conjunction with the removal data (Besbes et al, 2002).

As the success of these techniques is clearly dependent on the reasonable number of successive primary periods with multiple sampling occasions, we explore the effect of robust design on the precision of parameters using simulation. We are also able to show which combinations of parameters are estimable (Cole et al, 2010) when robust design reduces to a single secondary capture occasion within each primary sampling period.

References

Besbeas, P., Freeman, S.N., Morgan, B.J.T. & Catchpole, E.A. (2002) Integrating mark–recapture–recovery and census data to estimate animal abundance and demographic parameters. *Biometrics* **58**: 540–547.
Cole, D. J., Morgan, B.J.T. and Titterington, D. M. (2010) Determining the Parametric Structure of Non-Linear Models. *Mathematical Biosciences*, **228**, 16–30.
Moran, P.A.P. (1951) A mathematical theory of animal trapping, *Biometrika*, **38**, 307-311.

Specifying multi-species distance sampling models for estimating marsh bird abundance in the northern Gulf of Mexico

E.M. Adams [a] and B. Gardner [b]

[a] Biodiversity Research Institute
276 Canco Road
Portland, ME 04103
evan.adams@briloon.org

[b] School of Environment and Forest Sciences
University of Washington
Seattle, WA 98195

Keywords: marsh birds, abundance, multi-species, distance sampling

Abstract: Multi-species abundance or occurrence models are interesting to ecologists because they can quantify changes to communities in a given ecosystem and they can improve model fit for species when compared to individual species models. Often these approaches assume independence among abundance or occurrence of species by modeling each species as a single realization of a random distribution. However, there is often good reason to not assume independence among species. Many ecological relationships among species (predation, symbiosis, commensalism, etc.) strongly support the idea that species are not independently distributed relative to one another and that there is useful information in understanding these relationships. To test these ideas, we built a hierarchical community distance-sampling model that allows for correlation between species' abundances across sampling sites. We compared this model with individual models of the abundance of three marsh bird species (clapper rail, *Rallus crepitans*, seaside sparrow, *Ammodramus maritimus*, and least bittern, *Ixobrychus exilis*). Surveys were conducted at 1339 sites along the northern coast of the Gulf of Mexico, using a standardized protocol for marsh birds (Conway 2011). Estimates of detectability and availability were obtained using distance sampling and repeated visits per year, respectively. Using a multivariate normal distribution, we allowed estimates of the three species' abundance and their variances to be correlated. This model is useful for two main reasons: (1) species interactions can be disentangled using this method, allowing researchers to separate the effect of habitat and other environmental factors on the abundance of a given species versus the number of predators or competitors in the same location and (2) making efficient use of data and species correlations can improve predictions of species abundance over single species models, which is valuable for improving conservation and management decision-making.

Literature Cited:

Conway, C.J. 2011. Standardized North American marsh bird monitoring protocol. *Waterbirds* 34(3):319-346.

Analysing community data — how well do different statistical frameworks predict species- and community-level patterns?

Anna Norberg, Mathematical Biology Group, University of Helsinki
and Otso Ovaskainen, P.I., Professor at University of Helsinki, Finland
anna.norberg@ helsinki.fi

In community ecology, the interest lies by definition in the patterns of occurrence and abundance of multiple species. A large array of statistical frameworks has been developed for analyzing community ecological data sets describing the occurrences or abundances of species in a number of sampling units and the environmental attributes of those sampling units. Until recently, the emphasis with predictive community modelling methods has been on (single) species distribution models (SDMs), but today's joint species distribution models (JSDMs) have a particular advantage in that they can make use of correlations across species, which provide useful information for predicting occurrences of one species from others. In addition to computational differences, there are differences in the underlying biological hypothesis in single species versus community modelling — with single species modelling one assumes each species to responds independently to habitat variation, as opposed to joint responses of species assemblages.

The ability to make predictions is considered as the ultimate goal of ecology by many, and when doing predictive modelling, it is of importance to know how different techniques behave in various study settings, and with different species groups. It is essential to evaluate the model's predictive performance in both interpolation (model accuracy) and extrapolation (model generality). In our study, we compare various statistical frameworks (SDMs and JSDMs) in terms of their predictive performance. We parameterize the statistical frameworks with training data, and then use them to predict independent validation data representing both similar and different environmental conditions under which the training data were acquired. The simulated data are generated with dynamical models mimicking communities with either competitive, mutualistic or trophic interactions. The real data sets involve diatoms, trees and butterflies, and they are contrasting in terms of the numbers of species, the fraction of rare species, the number of sampling units, and the spatial context in which the sampling units are acquired.

The overreaching aim of this work is to examine whether JSDMs succeed better in prediction than SDMs for multispecies occurrence data. In particular, we ask how well models within these two categories perform in predicting species occurrence patterns under similar conditions as in the data used for the model parameterization (accuracy), as well as how well they perform in predicting communities outside the original training data (generality). We measure the model's prediction success by calculating multiple measures at both species and community levels, the latter including measures of species richness, community dissimilarity, and co-occurrence.

Habitat Suitability Criteria Development for Multiple Fish Species and Lifestages in the Susitna River, Alaska

A. Shelly, D. Reiser, and M. Gagner

[a] R2 Resource Consultants, Inc.
Redmond, WA
ashelly@r2usa.com

Keywords: habitat suitability, species distribution model, in-stream flow modeling.

Abstract: The State of Alaska is currently investigating the development of a large hydroelectric project on the Susitna River located in Southcentral Alaska. The Susitna is a large glacial feed river with a drainage area of over 19,000 square miles, flowing more than 300 miles from its glacial origin deep in the Alaskan Range to Cook Inlet near Anchorage. The river is home to all five pacific salmon species and is heavily utilized for recreational and subsistence fishing. Large inter- and intra-annual variation in fish habitat is common as the river completely ices over in winter, then experiences flooding and potential channel changes as the ice breaks up each spring. Substantial longitudinal variation in geomorphology and multiple inputs from fish-bearing tributaries results in a diverse and patchy fish assemblage. Habitat suitability criteria (HSC) models are a form of species distribution modelling used to evaluate the impacts of riverine flow modifications to fish habitat. These models have taken many forms since their original development in the 1980s, from univariate hand-drawn "best professional judgment" curves to more complex multivariate techniques (Ahmadi-Nedushan, 2006). We selected and applied multivariate mixed-effects logistic regression to build HSC curves based on fish utilization and habitat availability data collected during intensive and extensive field surveys in the middle and lower reaches of the Susitna River in 2013 and 2014. These models provide a relative measure of local habitat preference, while allowing for longitudinal differences in abundance for individual species. The models will be used to compare relative habitat availability as a function of flow and various hydroelectric operational scenarios.

References

Ahmadi-Nedushan, B., A. St-Hilaire, M. Berube, E. Robichaud, N. Thiemonge, and B. Bobee. (2006) A review of statistical methods for the evaluation of aquatic habitat suitability for instream flow assessment. *River Research and Applications,* 22(5):503-523.

Joint dynamic species distribution models: a tool for community ordination and spatiotemporal monitoring

James T. Thorson[a], James N. Ianelli[b], Elise A. Larsen[c], Mark D. Scheuerell[d], Cody Szuwalski[e], Elise F. Zipkin[f]

[a] Northwest Fisheries Science Center,
Seattle, WA, USA
James.Thorson@noaa.gov

[b] Alaska Fisheries Science Center
Seattle, WA
jim.ianelli@noaa.gov

[c] National Socio-environmental Synthesis Center
Annapolis, MD
ealarsen@umd.edu

[d] Northwest Fisheries Science Center
Seattle, WA
Mark.Scheuerell@noaa.gov

[e] Bren School of Environmental and Resource Management
University of California
Santa Barbara, CA
cszuwalski@bren.ucsb.edu

[f] Department of Integrative Biology,
Michigan State University
East Lansing, MI 48824
ezipkin@msu.edu

Keywords: spatiotemporal model; species ordination; species distribution model

Abstract: Species distribution models (SDM) are increasingly used to analyze count, presence/absence, and presence-only data sets. There is a growing literature regarding dynamic SDM (incorporating temporal variation in species distribution), joint SDM (analyzing the correlated distribution of multiple species), and geostatistical models (accounting for similarity between nearby sites caused by unobserved covariates). However, no previous study has combined all three attributes. We therefore develop spatial dynamic factor analysis (SDFA), which accounts for spatial similarity when estimating one or more "factors." Each factor evolves over time following a density-dependent (Gompertz) process, and the log-density of each species is approximated as a linear combination of different factors. We demonstrate SDFA using two multispecies case studies (an annual survey of bottom-associated species in the Bering Sea, and a seasonal survey of butterfly density in the continental USA). Case study applications show that that SDFA can be used for species ordination, i.e., showing that dynamics for butterfly species within the same genus are significantly more correlated than species from different genera. We also demonstrate how SDFA can rapidly identify dominant patterns in community dynamics, including the decline and recovery of several Bering Sea fishes following 2008, and the "flight curves" typical of early or late-emerging butterflies. We conclude by suggesting future research that could incorporate phylogenetic relatedness or functional similarity.

Spatio-temporal models for stream networks

J.M. Ver Hoef

National Marine Mammal Laboratory
NOAA-NMFS Alaska Fisheries Science Center
Seattle, Washington, USA
jay.verhoef@noaa.gov

Keywords: spatial statistics; time series; geostatistics; kriging

Abstract: Spatial statistical and time series models are combined in the emerging field of spatio-temporal statistical models, although there are few applications for stream networks. I review recent literature, and then discuss three main classes of models: 1) additive models, 2) separable models, and 3) nonseparable models. I present exploratory graphics to help choose among these models. The additive and separable models allow easy application of existing spatial stream-network models with time series models, and I show how to take computational advantage of these constructions. I then develop additive and separable models for stream temperature data from Oregon. This example illustrates estimation of regression coefficients and predictions in space and time at unsampled locations for stream networks, along with uncertainty bounds for estimates and predictions.

Species distribution modelling with combined data sources

Ian W. Renner[a] and Olivier Gimenez[b]

[a] School of Mathematical and Physical Sciences
The University of Newcastle
Callaghan, NSW, Australia
Ian.Renner@newcastle.edu.au

[b] Biodiversity and Conservation department
Centre d'Ecologie Fonctionnelle et Evolutive/CNRS
Montpellier, France
olivier.gimenez@cefe.cnrs.fr

Keywords: species distribution modelling; point process modelling; occupancy modelling; combined likelihood.

Abstract: A number of methods have been developed to model the distribution of a species for different types of data. For example, point process models provide a useful framework for presence-only data, logistic regression for presence-absence data obtained through single-visit systematic surveys, and occupancy modelling for detected/undetected data obtained through repeat-visit surveys. All of these methods can be used to obtain an estimate of the desired response, the intensity of the species. In situations for which multiple sources of data are available to model a species, these sources may be combined into a joint likelihood expression (Dorazio 2014; Fithian et al, 2015). In this presentation, I will explore questions of interpretation of the output from these combined approaches, as well as propose extensions to current practice through the introduction of a LASSO penalty. This approach will be demonstrated by modelling the distribution of the Eurasian lynx in eastern France.

References

Dorazio, R.M. (2014) Accounting for imperfect detection and survey bias in statistical analysis of presence-only data. Global Ecology and Biogeography, 23:1472-1484.
Fithian, W., Elith, J., Hastie, T., and Keith, D.A. (2015) Bias correction in species distribution models: pooling survey and collection data for multiple species. *Methods in Ecology & Evolution*, 6:424-438.

Spatiotemporal models, scale, and the problem of population distribution and abundance

L. Polansky[a,b], K. Newman[a,c], M. Nobriga[a,d], and L. Mitchell[a,e]

[a] U.S. Fish and Wildlife Service
850 Guild Ave., Suite105, Lodi, CA 95240, USA
[b] leo_polansky@fws.gov
[c] ken_newman@fws.gov
[d] matt_nobriga@fws.gov
[e] lara_mitchell@fws.gov

Keywords: data aggregation, delta smelt, fisheries, generalized additive models, geostatistical models, population ecology

Abstract: Usually the mechanisms, patterns, and questions about the distribution and abundance of populations are scale dependent (Levin 1992). These rang from small-scale events that elicit individual responses to large-scale ecosystem drivers. Analysis of disaggregated, survey specific data, and robustly aggregating these results for biological inference at larger scales, offers an approach to integrate scale dependent understanding. We used a generalized additive model (GAM) framework that included soap film smoothers (Wood et al. 2008) to describe catch density changes of spawning stage delta smelt (*Hypomesus transpacificus*), a rare annual estuarine fish species. The most supported model included time, space, and several salient environmental factors for making predictions. Results show that environmental conditions salient for determining large-scale distributional patterns in fact explain very little model deviance. While suitable environmental conditions may be necessary for increased densities, they are far from sufficient. In contrast, accounting for tidal conditions, which are believed to influence local distributions through impacts on local movement behavior, can have a substantial impact on estimates of large-scale aggregated abundance indices. Model-based abundance indices had substantially reduced estimate variance compared with design-based ones, and offer a way to reduce sampling bias under additional assumptions. Despite the increased difficulties concomitant to analyzing disaggregated survey data of rare and elusive species, it is necessary for understanding the import of factors determining the distribution and abundance at different scales, identifying where information is lacking, and improving estimates of population indices. We discuss some of the modeling difficulties encountered, what worked and what did not, and what models might be most ideal in future work.

Levin, S.A. (1992) The problem of pattern and scale in ecology. *Ecology*, 73, 1943–1967.
Wood, S.N. (2008) Soap film smoothing. *Journal of the Royal Statistical Society B*, 70, 931–955.

A note to the scientific program committee for help in evaluation of this abstract: This work was shaped very much by statistical pragmatism and conservation needs when trying to make inference from a rich yet complex dataset. As such there is little novel statistical theory or simulation. This would be an empirical data analysis heavy talk, but would provide some hopefully stimulating ideas for future statistical work as well as a sense of the difficulties encountered when more classic time series models of population are insufficient for data and questions.

Joint estimation of the dynamics of structured populations: accounting for correlation between vital rates

Edgar J. González[a] and Benjamin M. Bolker[b]

[a] Departmento de Ecología y Recursos Naturales
Universidad Nacional Autónoma de México
Mexico City, DF, Mexico
edgarjgonzalez@ymail.com

[b] Department of Mathematics & Statistics
McMaster University
Hamilton, ON, Canada
bolker@mcmaster.ca

Keywords: integral projection models, population dynamics, temporal correlation

Abstract: Population dynamics of many species vary according to continuous variables such as size or age, i.e. the vital rates (survival, growth and reproduction) of its members differ depending on their particular characteristics. To account for this, integral projection models (IPMs; Easterling et al. 2000) model each vital rate as a function of these variables and then integrate them into a single model. This separate modelling does not allow for the inclusion of correlations between vital rates, correlations that would be expected to exist due to trade-offs in resource allocation both within a single year for annual species or between years for perennial ones. To exemplify how such correlations can be estimated, we simulated data of populations displaying trade-offs in their vital rates. Under a likelihood framework, we used a hybrid optimization algorithm to obtain both estimates for the vital rates and their correlations. We compared results obtained under the traditional independence assumption and our approach. As expected, assuming independence between vital rates when they are actually correlated can bias results, and therefore the conservation/management recommendations derived from these models.

References

Easterling, M.R., Ellner, S.P. & Dixon, P.M. (2000). Size-specific sensitivity: applying a new structured population model. *Ecology*, 81: 694–708.

Analysing population trends with Generalized Additive Mixed Models

Jonas Knape

Department of Ecology
Swedish University of Agricultural Sciences
Uppsala, Sweden
jonas.knape@slu.se

Keywords: population trend; monitoring.

Abstract: Generalized Additive Models are commonly applied to estimate trends in the abundance of populations from large-scale, count based, monitoring surveys. I explore a conceptually simple tweak to the structure of GAM models typically used for such data, usually consisting of a smooth temporal trend, site effects and covariates. By adding a temporal random effect, in addition to the smooth temporal trend, I argue that noise in the population trajectory is better filtered out enabling more reliable inference about the long-term tendencies of the population. The approach is evaluated using simulations, as well as in applications to data from the Swedish Breeding Bird survey.

Estimating the contributions of anthropogenic and environmental factors to Florida manatee mortality

J.A. Hostetler[a], C. Fonnesbeck[b], C.J. Deutsch[c], and M.C. Runge[d]

[a] Marine Mammal Program, Fish and Wildlife Research Institute
Florida Fish and Wildlife Conservation Commission
100 8th Avenue SE
St. Petersburg, FL 33701
Jeffrey.Hostetler@MyFWC.com

[b] Department of Biostatistics
Vanderbilt University
Nashville, TN 37203, USA
Chris.Fonnesbeck@Vanderbilt.edu

[c] Wildlife Research Laboratory, Fish and Wildlife Research Institute
Florida Fish and Wildlife Conservation Commission
1105 SW Williston Rd., Gainesville, FL 32601
Chip.Deutsch@MyFWC.com

[d] USGS Patuxent Wildlife Research Center
12100 Beech Forest Road
Laurel, MD 20708
MRunge@USGS.gov

Keywords: Florida manatee, Bayesian analysis, mortality, cause of death, hierarchical modeling, proportions, red tide, cold.

Abstract: We estimated the relative contributions of cold, red tide, and three anthropogenic causes to mortality of the endangered Florida manatee (*Trichechus manatus latirostris*). We formulated a Bayesian hierarchical model to account for the uncertainty associated with the large number of carcasses for which cause of death was unknown (32.3%), and thereby estimated the proportions of death due to each cause, without having to condition on carcasses of known cause. We also estimated cause-specific mortality rates above the baseline for unusual mortality events using estimates of baseline mortality rates and an additive mortality assumption. In all age classes and regions of Florida, watercraft collision was the largest single cause of death, with water-control structures and marine debris accounting for fewer fatalities. In the Southwest region of Florida, red tide was a significant contribution to baseline mortality, with much higher mortality rates during unusual mortality events. The level of additional cold mortality varied with age class, winter severity, and quality of warm-water habitat; these patterns were mostly in the expected directions (additional mortality greater for younger manatees, for manatees in poor quality habitat, and during more severe winters). Bayesian credible intervals were generally wide, particularly for age classes and regions where data were sparse and where the fraction of undetermined deaths was large. We discuss implications of possible violations of model assumptions. These results may provide valuable information for analyzing threats to manatees and guiding management priorities. In general, this approach is relevant for a variety of applications where allocating unknown observations to categories is hampered by possible bias or high system uncertainty.

How should we define a plankton population?

L. Lindroos [a] and A. Lindén [b]

[a] Department of Environmental and Marine Biology
Åbo Akademi University
Turku, Finland
louise.lindroos@abo.fi

[b] Research- and development institute Aronia
Novia University of Applied Science
Ekenäs, Finland
andreas.linden@novia.fi

Keywords: Long-term monitoring data, phytoplankton, time-series analysis, population dynamics

Abstract: Long-term monitoring data of phytoplankton are collected annually in the Baltic Sea from multiple stations to cover vast areas, typically without replicates from the same stations. While data from stations close to each other are obviously correlated, it is unclear to which degree some or all stations can be thought to represent the same large population in species specific dynamical analyses. To investigate this question we studied the population dynamics of four phytoplankton taxa present at twelve offshore stations in the Baltic Sea using multivariate autoregressive state-space modelling. The data have been collected by the Finnish Environment Institute since 1979. We compared models, where samples from the twelve stations were considered the same population, where the stations were divided into five areas based on hydrographic- and environmental conditions, and models considering all stations as separate populations. We further compared models with equal or unequal density dependence and/or observation error variances, and with different process error correlation structures between the populations. We evaluated model support in terms of parsimony according to their AICc scores. Preliminary results indicate that the most parsimonious models concern one large population, or grouping into five areas with distinct populations, depending on species of concern. The approach of treating all twelve stations as different populations gained low support. Among the models with five populations, the most parsimonious variants always showed correlated process error and common estimates of density dependence. Our results supports the idea of treating samples from different stations as replicates from larger populations, and that populations close to each other show similar dynamics. However, such grouping should be well justified and may be relevant to do on a species specific basis.

Statistical estimation of hierarchical age-and-sex-structured production models for the Kuril harbor seals

T. Kitakado

Department of Marine Bioscience
Tokyo University of Marine Science and Technology
Tokyo, Japan
kitakado@kaiyodai.ac.jp

Keywords: age-structured models; Bayesian estimation; harbor seals; hierarchical models; integrated likelihood.

Abstract: The population of Kuril harbor seals off Cape Erimo in northern Japan had dramatically declined by the 1970s due to overhunting, and it had once faced with a risk of extinction. Since then, owing to protection measures, the population size has shown a steady recovery while the damage to fishery by the seals has also increased. So, it is necessary to conduct a population assessment and develop a resource management strategy to achieve a balanced objective between the conservation of population and mitigation of the damage to fishery. Here, hierarchical age-and-sex-structured population models were developed by taking account of a stochastic and density-dependent stock-recruitment relationship and applied to counts data for pups and non-pups from shore-based surveys continuously conducted during the breeding and molting seasons. Parameters such as carrying capacity, resilience and mortality rate were estimated using integrated likelihood. In addition, Bayesian estimation was used to examine the impact of assumed prior information on parameters like mortality rates derived from other studies/experiments. The estimation results showed that the models used in the analyses fitted well to the data. The population level has exceeded at least 60% of the carrying capacity though the extent of recovery depends on the model assumption. Stochastic simulations showed that the population size will still increase at the rate of current level of bycatch and be viable. It was also suggested that future management procedures involving culling adults while avoiding unintentional bycatch of yearling animals would be effective to achieve the balance objective for the conservation and management of the Kuril harbor seals off Cape Erimo. More objective harvest controls are also discussed the estimated population models under a framework of the so-called management strategy evaluation.

Stability, resilience, and the strength of density dependence in stochastic populations

Brian Dennis[a], Mark L Taper[b], and José Miguel Ponciano[c]

[a]Department of Fish and Wildlife Sciences
University of Idaho
Moscow, ID, USA
brian@uidaho.edu

[b]Department of Ecology
Montana State University
Bozeman, MT, USA
markltaper@gmail.com

[c]Department of Biology
University of Florida
Gainesville, FL, USA
josemi@ufl.edu

Keywords: density dependence; ecological resilience; ecological stability; diffusion process, stochastic population growth; demographic stochasticity; environmental stochasticity.

Abstract: Many concepts in population ecology have traditionally been defined in the language of deterministic modeling, among them density dependence, stability, and resilience. However, ecological populations are invariably buffeted by a variety of stochastic forces, as attested by the thousands of population time series now available in large online databases. Here we recast stability, resilience, and the strength of density dependence as properties of stochastic systems, using diffusion processes as modeling tools. Various statistical properties of diffusion models of population growth provide stochastic replacements for deterministic ideas. In diffusion processes, the type and nature of stochasticity affecting a population becomes tightly entangled in the forces causing population change, and emergent dynamical behaviors are found. The inclusion of stochastic components in models has the added benefit of allowing stability-related quantities to be estimated from data. We conclude that a complete understanding of stability, resilience, and density dependence in ecology must include stochasticity.

A Generalized Estimating Equation approach to Multivariate Adaptive Regression Splines

J. Stoklosa[a] and **D.I. Warton**[a]

[a]School of Mathematics and Statistics and Evolution & Ecology Research Centre
The University of New South Wales
New South Wales, Australia
j.stoklosa@unsw.edu.au

Keywords: GEE; GLM; MARS; Model selection; Negative binomial regression; Non-parametric smoothing

Abstract: Multivariate adaptive regression splines (MARS) is a popular nonparametric regression tool often used for prediction and uncovering important data patterns between the response and predictor variables. The standard MARS algorithm assumes normality and independence between continuous response variables. In this paper we extend MARS to generalized estimating equations, and we refer to this MARS-for-GEEs algorithm as “MARGE”. Through simulation we show that the proposed algorithm has improved predictive performance compared with the original MARS algorithm when using correlated and/or non-normal response data and is competitive with alternatives in the literature, especially for problems with multiple interacting predictors. The proposed algorithm is applied to various ecological data types.

A Simple Subsampling Approach to Modelling Big Mark-Recapture Data Sets

S.J. Bonner[a], **W. Burchett**[b] and **C. Brown**[c]

[a]Department of Statistical and Actuarial Sciences
University of Western Ontario
London, Ontario Canada
sbonner6@uwo.ca

[b]Department of Statistics
University of Kentucky
Lexington, Kentucky USA
woodrow.burchett@uky.edu

[c]Biological Science
University of Tulsa
Tulsa, Oklahoma USA
charles-brown@utulsa.edu

Keywords: Bayesian Inference; Big Data; Cliff Swallows; Mark-Recapture; Maximum Likelihood Inference; Multi-State Models

Abstract:

Modern mark-recapture models are complex beasts that may include envirnomental or individual covariates; partially observed covariates modelling the effects of fitness, spatial variation, or other factors that vary both between individuals and over time; random effects that account for other unexplained sources of variation; and other hierarchical components. The likelihood functions for these models are often computationally intensive to evaluate and this makes it difficult to fit these models with the available tools, both frequentist and Bayesian. This is particularly problematic when working with big data sets recording a large number of individuals captured over many occasions and may lead researchers to select models that are too simple for their data or to limit the number of models they consider in their selection and testing processes. We propose a simple approach to reduce the computational burden of fitting Cormak-Jolly-Seber type models that condition on the first release of each marked individual. Our method is based on modelling a smaller, systematically subsampled data set considering only the events within some small number of occasions following each release of a marked individual. The reduced data set retains maximal information about the parameters of interest and models can often be fit to the reduced data in a fraction of the time it would take to fit the full data with little loss in precision. Moreover, systematically subsampling in this way does not introduce bias so that standard models can be fit to the submsampled data without correction or further model development. The approach can be applied in either the frequentist or Bayesian framework to fit a broad range of models including multi-state models, models with time-varying individual covariates, spatially explicit models, and complex hierarchical models; and also leads to a natural method for model checking. We illustrate the new approach by comparing the run time and precision of parameter estimates for a variety of models fit to data from more than 235,000 marked cliff swallows captured and marked in Nebraska over a 35 year period.

Data cloning can guide study design to ensure parameter estimability in complex ecological models

Stephanie J. Peacock[a,b,*], Martin Krkosek[b], Mark A. Lewis[a,c] and Subhash Lele[c]

[a] Department of Biological Sciences
University of Alberta
Edmonton, Alberta
stephanie.j.peacock@gmail.com

[b] Department of Ecology and Evolutionary Biology
University of Toronto
Toronto, Ontario

[c] Department of Mathematics and Statistics
University of Alberta
Edmonton, Alberta

[*]Presenting author

Keywords: Bayesian, data cloning, mechanistic models, estimability, parameter estimation.

Abstract: The statistical tools available to ecologists are becoming increasingly sophisticated, allowing more complex, mechanistic models to be fit to ecological data. Such models have the potential to provide new insights into the processes underlying ecological patterns, but the inferences made are limited by the information in the data. Statistical inestimability of model parameters due to insufficient information in the data is a problem too-often ignored by ecologists employing complex models. Here, we show how a new statistical computing method called data cloning can be used in simulation studies to assess the estimability of model parameters and inform study design before data are collected. A case study of parasite transmission from farmed to wild salmon highlights that assessing the estimability of ecologically relevant parameters should be a key step when designing studies in which fitting complex mechanistic models is the end goal.

Adventures in video – processing massive video data for animal abundance estimation

C. R. Donovan, N. B. Erichson, A. Webb, D. L. Borchers & S. Menze
Centre for Research in Ecological and Environmental Modelling (CREEM)
School of Mathematics and Statistics
University of St Andrews
crd2@st-andrews.ac.uk

Keywords: big data, image processing.

Abstract: High-definition aerial surveying is becoming an increasingly popular means to survey animal populations, particularly in the context of Environmental Impact Assessments (EIAs). The data gathered naturally avoids some of the biases found in other surveying methods, but still presents substantive modelling challenges, such as extremely large data volumes, moving sensors and dynamic backgrounds. The data in question is effectively 500 megapixel video, producing 1TB per hour. Treated directly, the spatio-temporal matrices are huge e.g approximately 500 million × 80,000 for one hour of data. We present here the approaches to processing such data, both in terms of logistics (hardware/software) and initial algorithms for automated object extraction - a necessary step before object classification i.e. animals must be located and identified to be counted.

References

Erichson, N.B.E. & Donovan, C.R. (2016) Randomized Low-Rank Dynamic Mode Decomposition for Motion Detection. *Submitted to Journal of Computer Vision and Image Understanding*. arXiv:1512.03526 [cs.CV]

A Best Practices Approach for Detecting Change at Marine Renewable Energy Sites

Hannah Linder [a] and John Horne [b]

[a] School of Aquatic and Fishery Sciences
University of Washington
Seattle, Washington
lindeh@uw.edu

[b] School of Aquatic and Fishery Sciences
University of Washington
Seattle, Washington
jhorne@uw.edu

Keywords: marine renewable energy; regression models; time-series; simulation scenarios; intervention analysis.

Abstract: In compliance with the National Environmental Policy Act (NEPA), an evaluation of environmental effects is mandatory for obtaining permits for any MRE project. Evaluation procedures include an initial assessment of baseline conditions, and on-going monitoring during operation to determine if biological conditions change relative to the baseline. Currently, there are no best practices established for the analysis of MRE monitoring data. An approach was established to evaluate and recommend analytic models used to characterize and detect change in biological monitoring data. The approach is composed of six steps: review current MRE monitoring practices, identify candidate models for evaluation, fit models to a baseline dataset, develop simulated scenarios of change post-construction of an MRE project, evaluate the fit of models to simulated data, and produce recommendations on the choice of analytic model for monitoring data. The evaluation process was performed on two sets of empirical data that are representative of common biological monitoring data: normal and non-normal continuous time-series. The evaluated models include: linear regression models, time series models, and nonparametric machine learning models. Many of these models have yet to be used in MRE monitoring studies, and their evaluation will advance the techniques for effective environmental monitoring. Model fit diagnostics Root-Mean-Square-Error and Mean-Absolute-Scaled-Error were used to measure accuracy of model predicted values. A power analysis was used to evaluate the ability of each model to measure and detect change from baseline conditions in each scenario dataset. Results of this evaluation will be used to generate guidelines on choice of models applicable for detecting change in environmental monitoring data. The creation of standardize guidelines for selecting the model used to analyze MRE environmental monitoring data will ensure the most accurate analysis of data and enable effective cross-comparisons among MRE sites. Consequently, the creation of a best practices approach for selecting a model will advance the growth of MRE development.

Evidence, errors, and information criteria

J.M. Ponciano[a] and **B. Dennis**[b] and **M. L. Taper**[c]

[a]Department of Biology
University of Florida
Gainesville, FL, USA
josemi@ufl.edu

[b]Department of Fish and Wildlife Sciences
University of Idaho
Moscow, ID, USA
brian@uidaho.edu

[c]Department of Ecology
Montana State University
Bozeman, MT, USA
markltaper@gmail.com

Keywords: Evidence functions; model misspecification; probability of misleading evidence; information criteria

Abstract: The shortcomings of Neyman-Pearson hypothesis testing approaches for model selection are well known and documented in statistics. However, the reliability of information criteria as model selection tools has not been adequately studied. Richard Royall (1997) formalized the use of likelihood ratios to interpret observations as evidence for a given model versus another. Lele (2004) extended this framework into more general evidence functions including certain information criteria. The evidence function framework also allows for conditional error probabilities reminiscent but superior to Neyman-Pearson's Type I and Type II error rates in that the evidence error probabilities all go to zero with increasing sample size. These are the probabilities of strong misleading evidence and the probabilities of weak evidence. However, Royall's original work is limited in that the comparison is between two fully specified models and one of the models is assumed to have generated the data (i.e. no model misspecification). In this work, we develop an analytical and computational characterization of Royall's errors incurred in model selection via AIC, SIC, CAIC, AICc, HQIC, EIC1, and a new adjusted version of the EIC1 under a variety of configurations of model misspecification. SIC, CAIC, HQIC, and the new adjusted EIC1 behave consistently as evidence functions in that their error probabilities approach zero as sample size increases. However, AIC, AICc, and EIC1 cannot be considered as evidence functions because their error probabilities do not go to zero with increasing sample size but instead approach positive constant values. Thus, the behavior of AIC, AICc, and EIC1 in model selection is similar in error structure to Neyman-Pearson hypothesis testing.

References

Royall, R. (1997) Statistical Evidence: A likelihood Paradigm, London: Chapman and Hall. Lele, S.R. (2004) Evidence functions and the optimality of the law of likelihood. In: Taper ML, Lele, SR (eds). *The nature of scientific evidence: statistical, philosophical and empirical considerations.* The University of Chicago Pressm Chicago, pp 191-216.

Model projections in model space: Multimodel inference beyond model averaging.

Mark L. Taper[a] and José M. Ponciano[b]

[a] Department of Ecology
Montana State University
Bozeman, MT, USA
MarkLTaper@gmail.com

[b] Department of Biology
University of Florida
Gainesville, FL, USA
jm.ponciano@gmail.com

Keywords: model space; model averaging; model selection; AIC

Abstract: Information criteria have had a profound impact on modern science. They allow researchers to estimate which of a set of model is closest to the generating process. Unfortunately, information criterion comparison does not tell how good the best model is. The standard analysis ignores that there is an estimable divergence relationship amongst all of the models as well as divergences from each model to the generating process. Using both sets of divergences, a model space can be constructed including an estimated location for the generating process. Thus, not only can the analyst determine which model is closest to the generating process, she can also determine how close to the generating process the best model is. Properties of the generating process estimated from these projections are more accurate than those estimated by model averaging.

R coding for a mixed crowd: Statistical analysis in environmental consultation

P. Molloy[a] and M. O'Neill[b]

[a] Stantec Consulting Ltd.
Saanichton, BC, Canada
philip.molloy@stantec.com

[b] Stantec Consulting Ltd.
Sidney, BC, Canada
meghan.oneill@stantec.com

Keywords: environmental consultation; interdisciplinary programming; Statistical programming; R.

Abstract: With the growth of R in the statistical community, ecology graduates are increasingly choosing R as their preferred statistical programming language. Thereafter, many leave academia to join industry or consulting companies and work with colleagues from diverse backgrounds, with variability in statistical training and familiarity with R. Further, project teams within a consulting environment are dynamic, meaning that people often pick up work previously developed by others. This situation calls for thoughtful coding practices that consider a mixed, dynamic team and foster collaboration and perpetuate clear, defensible analysis without loss of efficiency. Several coding guides are available to promote professional practices but they typically assume coders are working with other programmers who are familiar with and fluent in R (or programming generally). Here, we adapt these best-practice guides and adapt them to provide R-coding advice for statistical ecologists working in interdisciplinary, cross-sector teams. Our recommendations are designed to promote clarity of code to facilitate collaboration with R novices and hand-over of scripts within a dynamic team. More generally, we suggest that our recommendations are of value to all R coders to foster clear, clean code that facilitates peer review and the incorporation of junior researchers.

Large Scale Density estimation of blue and fin whales: combined whale distribution and density estimates using bearing data.

Danielle Harris [a], Jennifer Miksis-Olds[b], Len Thomas[a] and Julia A. Vernon[b]

[a]Centre for Research into Ecological and Environmental Modelling,
University of St Andrews,
St Andrews, Scotland, UK
dh17@st-andrews.ac.uk

[b]Applied Research Laboratory,
The Pennsylvania State University,
State College, USA

Keywords: density estimation, animal distribution, acoustic monitoring.

Abstract: Passive acoustic monitoring of marine mammals is common and, with properly applied statistical techniques, it is possible to estimate absolute animal density from acoustic recordings. The most appropriate density estimation method to use depends on how much detail about an animal's location can be derived from the recordings, which is often determined by the number and configuration of deployed instruments. At best, three-dimensional locations of calling animals can be estimated from acoustic data, but recordings can yield little or no information about animals' positions. A suite of density estimation methods exist that can be applied to different types of acoustic survey data, but the variety of survey scenarios is such that novel methods are still required.

Here, we present a new method for estimating cetacean density using acoustic data, where only horizontal bearings to calling animals are estimable. Measured call signal-to-noise ratios (SNR) are also required, as well as auxiliary information about call source levels, sound propagation, and call production rates. Firstly, acoustically masked areas in the study region (where transmission loss is predicted by sound propagation models to be too high for calls ever to be detected) are identified and excluded from the analysis. The bearing and SNR measurements, source level data, and transmission loss predictions are then used to create a probability density function for each call location, as a function of range from the instrument. Next, characterizing the automatic detection process gives the probability of detecting a call as a function of SNR, enabling detection probabilities to be estimated for each received call. Each detected call is then scaled by its associated probability of detection to account for undetected calls, resulting in a call abundance estimate. The resulting estimates are then smoothed using a spatial Generalized Estimating Equation to give an estimated density surface, accounting for spatial autocorrelation. If an appropriate call production rate is available, the call density surface can be rescaled to reflect animal density.

Results from simulations and a pilot study are presented. Simulations were conducted using recordings of "Sri Lankan" blue whale calls from Comprehensive Nuclear-Test-Ban Treaty Organisation instruments in the Indian Ocean as a motivating example. From 100 simulation runs, the median bias in estimated call density was -12%. Fin whale recordings from another CTBTO array in the Pacific Ocean were then analysed for the pilot study.

Are the fishes imperiled? Evaluating uncertainty from long-term fish monitoring in the San Francisco Delta

K.A. Millers [a] and J.T. Peterson [a, b]

[a] Department of Fisheries and Wildlife
Oregon State University
Corvallis, OR, USA
kimberley.millers@oregonstate.edu

[b] US Geological Survey, Oregon Cooperative Fish and Wildlife Research Unit
Oregon State University
Corvallis, OR, USA
jt.peterson@oregonstate.edu

Keywords: long-term monitoring; fish; San Francisco Delta; abundance; occupancy.

Abstract: Estuaries are at risk from anthropogenic disturbances and climate change through changes to water quality, habitat, hydrologic regimes and other factors that threaten the health, abundance and diversity of fishes inhabiting such ecosystems. Our ability to manage and conserve populations is often determined from understanding the current status of populations and their long-term trends in association with environmental conditions. However, sampling bias and variance are the primary factors that influence data quality and inferences of catchability, abundance and occupancy. This study evaluates the effectiveness of current monitoring efforts of fishes in the San Francisco Delta and estimates the bias in long-term population trends. The San Francisco Delta comprises the Sacramento River, San Joaquin River and their surrounding estuaries. This Delta has undergone heavy modification over the past 150 years and consequently impacted resident fish populations. Extensive monitoring of fish populations has been conducted throughout the estuary since the late 1960's using a variety of active sampling gear, such as benthic trawls, otter trawls, and beach seines. Using simulation and long-term monitoring data, we estimate the bias and variability of abundance and occupancy for selected fish species in the San Francisco Delta. We compare the performance of alternative statistical methods that included N-mixture models, state-space models, and Bayesian methods to evaluate population dynamics. We compare population estimates and uncertainty from the different model types and examine the relationship between catchability and environmental conditions to assess drivers of bias in the estimates. The findings suggest that both spatial and temporal factors have varying influences on our ability to estimate population abundance and occupancy. Additionally, the use of different models to estimate population dynamics (e.g. various N-mixture model) of fish species demonstrate how inferences are markedly influenced by statistical method and sample design. This study provides an informative approach to considering the status and long-tern trends of resident fishes in the San Francisco Delta. By evaluating the bias and variability of monitoring data managers can find improved population estimates which can allow more appropriate management actions to be taken.

Count data collected using a robust design: models, results and recommendations

E. Matechou[a], **K. Winner**[b], **R. Hutchinson**[c], **J.S. Pippen**[d], **L. Ries**[e]

[a]School of Mathematics,
Statistics & Actuarial Science
University of Kent
Canterbury, Kent
e.matechou@kent.ac.uk

[b]College of Information
and Computer Sciences
University of Massachusetts, Amherst
Amherst, Massachusetts
kwinner@cics.umass.edu

[c]School of Electrical Engineering
and Computer Science
Department of Fisheries and Wildlife
Oregon State University
Corvallis, Oregon
rah@eecs.oregonstate.edu

[d]MPG Ranch
Florence, Montana
jpippen@mpgranch.com

[e]Department of Biology
Georgetown University
Washington D.C.
lries@sesync.org

Keywords: robust design; count data; phenology; imperfect detection; flight curves

Abstract: We propose novel models for count data collected using a robust-design framework in an open population. We present results on real butterfly count data from the MPG ranch in Montana, US (`http://mpgranch.com/`) as well as simulated data. Our experiments highlight the increase in precision of parameter estimates when the robust design is employed as opposed to the more standard line-transect single-count sampling approach. Based on these results, we make recommendations under a fixed effort assumption to answer the question: "Assuming equal cost, is it better to conduct additional primary or secondary sampling occasions?" with a particular focus on the impacts for citizen science. We also explore the effect of informative covariates for detection and emergence on this problem. Note: while our experiments focus on count data collected via transects, our analysis is extensible to any count-collecting sampling protocol.

Zooplankton prevalence and abundance in the Na Thap tidal river of Thailand

*A. Lim[d], C. Lueangthuwapranit[a], S. Chesoh[b] & D. McNeil[c]

[a] Department of Technology and Industry
Faculty of Science and Technology, Prince of Songkla University
Pattani 94000, Thailand
chokchai.l@psu.ac.th

[b] Department of Science
Faculty of Science and Technology, Prince of Songkla University
Pattani 94000, Thailand
chesoh.s@hotmail.com

[c] Department of Statistics
Faculty of Science, Macquarie University
Sydney 2109, Australia
don.mcneil@mq.edu.au

[d]Department of Mathematics and Computer Science
Faculty of Science and Technology, Prince of Songkla University
Pattani 94000, Thailand

Keywords: prevalence; abundance; estuarine zooplankton; factor analysis; regression model

Abstract: In aquatic food-webdynamics, phytoplankton and zooplankton are primary and secondarycomponents, respectively.The objective of this study wasinvestigate theprevalence and abundance ofzooplankton in a tropical tidal river estuary, using phytoplankton abundance at the same location and time as the predictor in a linear model and seasonal, annual and upstream location factors as covariates.Samples were collected from ten sites at bi-monthly intervals from 2005-2015 using standard plankton nets. Individuals were counted and categorized into 139 identifiable taxa. A factor analysis model with promax rotation was fitted by maximum likelihood to the correlation matrix of these plankton counts after adding 1 and log-transforming to satisfy the normality assumption.After reducing the number of factors by aggregating 31 rare taxa into a single group and amalgamating six fitted factors into three interpretable groups plus a further group that did not fit the factor model, the first group was found to comprise oligohaline zooplankton, with the second and third groupscomprising polyhalineandmesohaline zooplankton, respectively, and a remaining group comprising euryhaline zooplankton. We then fitted models with four additive predictors, namely bi-month (six levels), year (11 levels), site(up to 10 levels) and phytoplankton abundance (four quartile-defined levels) to estimate prevalence of zooplankton in each group (using logistic regression) and corresponding abundance in populated samples (using linear regression on logarithms of counts). Q-q plots indicatedplausibility of normality assumptions, and results were displayed by graphing confidence intervals for comparing levels of each factor with overall means.Oligohaline, mesohaline and euryhaline zooplankton were positively associated with the higher phytoplankton abundance. This study confirmed that prevalence and abundance of zooplankton in estuarine environment depend on season, trend, upstream location, as well as phytoplankton availability (Lueangthuwapranit et al., 2011), with minimal confounding among these predictors.

Reference

Lueangthuwapranit, C., Sampantarak, U. and Wongsai, S. (2011) Distribution and abundance of phytoplankton: influence of salinity and turbidity gradients in the Na Thap River, Songkhla Province, Thailand. *Journal of Coastal Research,* 27(3):585-594.

*presenter

The Influence of the Turning Angle?

V.L. Goodall [a], L.P. Fatti [b] and N. Owen-Smith [c]

[a] Department of Statistics
Nelson Mandela Metropolitan University
Port Elizabeth, South Africa
victoriagoodall@gmail.com

[b] School of Statistics and Actuarial Science
University of the Witwatersrand
Johannesburg, South Africa
paulfatti@gmail.com

[c] School of Animal, Plant and Environmental Sciences
University of the Witwatersrand
Johannesburg, South Africa
Norman.Owen-Smith@wits.ac.za

Keywords: Hidden Markov models; turning angle; ungulate movement; activity states; Kruger National Park.

Abstract: Hidden Markov models have become a popular technique for the analysis of animal movement data. Common input metrics are step-length and the turning angle between successive locations. We used a simulation to investigate the effect of including the turning angle in the modelling process. Our case study uses hourly movement data for three ungulate species in the Kruger National Park, South Africa. Simulated movement data similar to these species, indicated that the accuracy of the state allocation did not improve much by including the turning angle in the analysis, particularly for models with more than two latent states. The inclusion of the turning angle increases the complexity of the model, while circular data presents additional challenges for the convergence of the likelihood function due to the circular nature of the data. For cohesive ungulate herds, the turning angle between successive displacements is less ecologically relevant [1]. A comparison of the state allocation accuracy using the models fitted with and without the turning angle, showed minimal differences.

References

[1] Owen-Smith N, Goodall VL, Fatti LP. Applying mixture models to derive activity states of large herbivores from movement rates obtained using GPS telemetry. Wildlife Research. 2012;39(Journal Article):452-62.

Hidden Markov models for acceleration and three-dimensional movement data in a soaring raptor

T. Photopoulou[a], **M. Murgatroyd**[b], **R. Langrock**[c], **V. Leos-Barajas**[d],
L. Underhill[b], **A. Jenkins**[e], **A. Amar**[f], **W. Bouten**[g] and **R. Altwegg**[a]

[a]Statistics in Ecology, Environment and Conservation
Department of Statistical Sciences, University of Cape Town, Rondebosch 7701, South Africa
TP: theoni.photopoulou@uct.ac.za, RA: res.altwegg@gmail.com

[b]Animal Demography Unit, University of Cape Town, Rondebosch 7701, South Africa
MM: verreauxs@gmail.com, LU: les.underhill@uct.ac.za

[c]Department of Business Administration and Economics
Bielefeld University, Bielefeld, Germany
roland.langrock@uni-bielefeld.de

[d]Department of Statistics, Iowa State University, Ames, IA, USA
vianey@iastate.edu

[f]Percy Fitzpatrick Institute for Ornithology, University of Cape Town
Rondebosch 7701, South Africa
arjun.amar@uct.ac.za

[e]AVISENSE Consulting, 10 Harrier Circle, Imhoff's Gift
Kommetjie 7975, South Africa
andrew@avisense.co.za

[g]Computational Geo-Ecology, Institute for Biodiversity and Ecosystem Dynamics
University of Amsterdam, Amsterdam, The Netherlands
w.bouten@uva.nl

Keywords: movement modelling; acceleration; tracking data; hidden Markov model; raptor

Abstract: Raptors are becoming increasingly vulnerable to anthropogenic activities but for many species, we have a patchy understanding of their movements and daily activities, or how these relate to environmental effects and topographic features. Understanding these interactions is key to identifying current and future threats, and mitigating them. We apply hidden Markov models (HMM), in a maximum likelihood framework, to tracking data from black eagles (*Aquilla verreauxii*) to investigate flight behaviour and activity patterns in the Western Cape, South Africa. Fine scale GPS data were used to investigate detailed flight behaviour using periods of consecutive high resolution locations. Samples of three-axis acceleration data, though information-rich, were collected intermittently and at lower temporal resolution between sampling occasions. We investigate the effect of weather and topographic variables (1) on activity levels, using a model for acceleration, and (2) on flight behaviour using a multivariate model for horizontal displacement, turning angle and change in altitude. We show that even infrequently collected acceleration data can be used to obtain a crude estimate of activity without the requirement for behavioural context; and we implement a three-dimensional model, including change in altitude, for animal movement in a volant species.

Geolocation of Atlantic cod off New England using archival tagging data and hidden Markov models

C. Liu[a,1], **G.W. Cowles**[a], **D.R. Zemeckis**[a], **S.X. Cadrin**[a], and **M.J. Dean**[b]

[a]Department of Fisheries Oceanography
School for Marine Science and Technology (SMAST)
University of Massachusetts Dartmouth
New Bedford, MA, USA
[1]cliu3@umassd.edu

[b]Annisquam River Marine Fisheries Station
Massachusetts Division of Marine Fisheries
Gloucester, MA, USA

Keywords: hidden Markov model; geolocation; movement tracking; Atlantic cod

Abstract: Understanding fish movement is critical for describing spatial processes and population dynamics. Archival electronic tags present the opportunity to acquire high resolution data on fish movements. Geolocation methods using archival tags have been commonly used to estimate daily positions of pelagic species using tidal features, temperature, depth, and other environmental data. However, the development and validation of alternative methods is required for geolocation of demersal species, because of considerable error in estimated positions. In this effort we developed geolocation methods for Atlantic cod off New England using hidden Markov models (HMMs). The geolocation approach is based on a modification of the observation likelihood and behavior models of an existing HMM framework (Pedersen et al., 2008) and addresses both region- and species-specific challenges. The HMM emission probabilities are described by the likelihood model which compares environmental data recorded on the tag with those derived from an oceanographic model. Likelihood distributions are depth- and temperature-based with tidal-based exclusion. The transition probabilities are described by the behavior model which constrains the horizontal movement of the fish. The Viterbi algorithm was applied to the HMM-constructed posterior probability distribution to generate the most probable track. Validation experiments were performed using a comprehensive quantitative skill assessment process relying on stationary tags moored on the seafloor, double-tagged fish (archival tag and acoustic transmitter), and simulated tracks. Known data, including fish locations and activity level metrics, showed good agreement with those estimated by the HMM geolocation model. Processing the recovered archival tagging data using the developed geolocation methods is expected to improve our understanding of cod movements and population dynamics, which will be helpful for informing future stock assessments and fishery management plans.

References

Pedersen, M.W., Righton, D., Thygensen, U.H., Andersen, K.H., and Madsen, H. (2008) Geolocation of North Sea cod (*Gadus morhua*) using hidden Markov models and behavioural switching. *Canadian Journal of Fisheries and Aquatic Sciences*, **65**(11):2367–2377.

Dynamic social networks based on movement

H. R. Scharf[a], **M. B. Hooten**[b,a], **B. K. Fosdick**[a],
D. S. Johnson[c], **J. M. London**[c], and **J. W. Durban**[d]

[a]Department of Statistics
Colorado State University
Fort Collins, CO, USA
henry.scharf@colostate.edu

[b]U.S. Geological Survey
Colorado Cooperative Fish and Wildlife Research Unit
Department of Fish, Wildlife, and Conservation Biology
Colorado State University
Fort Collins, CO

[c] NOAA Alaska Fisheries Science Center
National Marine Fisheries Service
Seattle, WA, USA

[d] NOAA Southwest Fisheries Science Center
National Marine Fisheries Service
La Jolla, CA, USA

Keywords: dynamic social network; animal movement; *Orcinus orca*; hidden Markov model; Gaussian Markov random field

Abstract: Network modeling techniques provide a means for quantifying social structure in populations of individuals. Data used to define social connectivity are often expensive to collect and based on case-specific, *ad hoc* criteria. Moreover, in applications involving animal social networks, collection of these data is often opportunistic and can be invasive. Frequently, the social network of interest for a given population is closely related to the way individuals move. Thus telemetry data, which are minimally-invasive and relatively inexpensive to collect, present an alternative source of information. We develop a framework for using telemetry data to infer social relationships among animals. To achieve this, we propose a Bayesian hierarchical model with an underlying dynamic social network controlling movement of individuals via two mechanisms: an attractive effect, and an aligning effect. We demonstrate the model and its ability to accurately identify complex social behavior in simulation, and apply our model to telemetry data arising from killer whales. Using auxiliary information about the study population, we investigate model validity and find the inferred dynamic social network is consistent with killer whale ecology and expert knowledge.

References

Scharf, Henry R., et al. "Dynamic social networks based on movement." arXiv preprint arXiv:1512.07607 (2015).

A Switching Hidden Markov Movement Model for rapidly identifying behavioral states from animal tracks.

K. Whoriskey [a] M. Auger-Méthé [a], C. M. Albertsen [b], F. G. Whoriskey [c],
and J. Mills Flemming [a]

[a] Department of Mathematics and Statistics
Dalhousie University
Halifax, N.S., Canada
kwhoriskey@dal.ca; auger-methe@dal.ca; joanna.flemming@dal.ca

[b] National Institute of Aquatic Resources
Technical University of Denmark
Charlottenlund, Denmark
cmoe@aqua.dtu.dk

[c] Ocean Tracking Network
Dalhousie University
Halifax, N.S., Canada
fwhoriskey@dal.ca

Keywords: Ocean Tracking Network; SHMMM; HMM; SSM; animal movement; behavioral states

Abstract: Animals move in order to maximize their probability of survival and reproduction. Identifying underlying drivers of animal movement (behavioral states) is required for understanding how and why animals use available space, and thus informs proper management of both species and ecosystems. In the aquatic realm, researchers often employ satellite telemetry technology for documenting animal paths or tracks through time. Hidden Markov models (HMMs) are a common method for estimating behavioral states from animal telemetry data with minimal error, and can be rapidly fitted via Maximum Likelihood methods. For those tracks measured with error, state-space models (SSMs) provide an alternative method for estimating behavioral states, but tend to be fitted with comparatively slow Bayesian methods. One particularly popular Bayesian SSM is the Discrete-time Correlated Random Walk with Switching (DCRWS) of Jonsen et al. (2005). We utilized the process equation of the DCRWS to develop a new HMM for estimating behavioral states from animal tracks assumed to have minimal error. We used the R-package Template Model Builder (TMB) to fit the model within a Maximum Likelihood framework that drastically reduced computational time compared to the Bayesian implementation. This model, which we call the Switching Hidden Markov Movement Model (the SHMMM), will be of great utility to movement ecologists who are interested in implementing the DCRWS but have highly accurate data. Furthermore, while many HMMs cannot account for measurement error, the SHMMM should be directly extendible. We report on our current efforts to extend the SHMMM into a SSM framework using TMB, which has previously shown great promise for rapidly analyzing error-prone animal tracking data (Albertsen et al. 2016).

References

Albertsen, C. M., K. Whoriskey, D. Yurkowski, A. Nielsen, and J. Mills Flemming. 2015. Fast fitting of non-Gaussian state-space models to animal movement data via Template Model Builder. *Ecology*, 96:2598-2604.

Jonsen, I. D., J. Mills Flemming, and R. A. Myers. 2005. Robust state-space modeling of animal movement data. *Ecology,* 86:2874-2880.

Modeling diurnal variation in multiphasic movement

B. M. Bolker[a] and **M. Li**[b]

[a]Departments of Mathematics & Statistics and Biology
McMaster University
Hamilton, Ontario
bolker@mcmaster.ca

[b]Department of Biology
McMaster University
Hamilton, Ontario
author2@institut.com

Keywords: correlated random walk; information-theoretic approaches; temporal variability

Abstract: We extend the now-standard multiphasic framework for animal movement, which uses a hidden Markov model (HMM) to combine sequences of correlated random walks with different characteristics, to allow for diurnal variation in behaviour. This straightforward addition improves the fit of a model for Florida panther movements. Relative to temporally homogeneous HMMs, predictions based on the diurnal model more accurately capture the autocorrelation and diurnal structure of the data. More generally, we use this problem to explore the issue of model complexity; as shown by van de Kerk *et al.* (2015), the BIC-optimal number of hidden states in a HMM can be surprisingly and inconveniently large when using large, automatically collected data sets. This is a probably a result of structural simplicity in the models; adding appropriate dimensions of complexity to the model reduces the BIC-optimal number of states, while adding computational challenges.

References

van de Kerk, M., D. P. Onorato, M. A. Criffield, B. M. Bolker, B. C. Augustine, S. A. McKinley, and M. K. Oli. Hidden Semi-Markov Models Reveal Multiphasic Movement of the Endangered Florida Panther. *Journal of Animal Ecology* 84, no. 2 (March 1, 2015): 576-85. doi:10.1111/1365-2656.12290.

Detecting change in benthic ecosystems with digital imagery: monitoring design in a challenging environment

Nicholas Perkins[a], Nicole Hill[b], Scott D. Foster[c], Geoff Hosack[d], Neville Barrett[e]

[a,b,e] Institute of Marine and Antarctic Studies, University of Tasmania, Australia
Nicholas.Perkins@utas.edu.au
Nicole.Hill@utas.edu.au
Neville.Barrett@utas.edu.au
[c,d] Commonwealth Scientific and Industrial Research Organisation, Hobart, Tasmania, Australia
scott.foster@csiro.au
geoff.hosack@csiro.au

Keywords: marine benthos, AUV, large-scale monitoring, INLA, geostatistical modelling, sampling design, repeat transects

Abstract: As benthic ecosystems face increasing pressure from sources such as climate change and overfishing, monitoring is essential to quantify any impacts and inform management decisions. Recent advances in sampling platforms, such as Autonomous Underwater Vehicles (AUVs), equipped with digital cameras have allowed a greater range of habitats to be surveyed. This imagery, and the ability to make repeated samples that are geolocated, gives researchers a way to monitor changes in the cover of target biota. However, there are a number of survey and sampling issues that need to be explored in order to optimise the information gained from AUV deployments. These include: (i) Efficiently subsampling and scoring the vast quantities of AUV imagery produced in an individual deployment; (ii) The spatial repeatability of transects and the consequences for detecting temporal trends; and (iii) The design of surveys over large spatial domains that are able to detect change. To address these aspects of monitoring with AUVs, we use a variety of approaches including simulated sampling scenarios and geostatistical modelling using the Integrated Nested Laplace Approximation (INLA) method. We show that for biota that are likely to be the target of monitoring programs, the inclusion of more images at the expense of points per image is likely to give increased precision in transect cover estimates. We use model estimates of the range and spatial variance of indicator species to examine the effect on cover estimates of repeat transects at various displacements and the implications of decreased image sampling. By using simulations, we provide generalizable results that encompass a range of potential distributions. We highlight the importance of considering the characteristics of the target biota (e.g. size, abundance and the distributional properties) when designing surveys and sampling schemes. We also show that for the majority of target species, a relatively high (< 5 m) precision in the placement of repeat transects is required to have a reasonable chance of detecting temporal change. Further temporal modelling of potential range shifts and changes in abundance of indicator species highlight how AUVs may be capable of providing a monitoring platform over large scales. By taking into account a wide range of sampling issues, distributional properties of biota and technical limitations, we are able to provide practical advice for the use of benthic imagery as a long-term monitoring tool.

Evaluating and deploying sniffer dogs to detect invasive plants in the Australian alps

C.E. Hauser[a], I. Veltheim[a], B. Crase[a] and G. Guillera-Arroita[a]

[a] School of BioSciences
University of Melbourne
Melbourne, Australia
chauser@unimelb.edu.au

Keywords: detection rate, survey design

Abstract: Since the 1999 discovery of Hawkweeds (*Hieracium* spp) in Australia's Alpine National Park, managers have relied upon human search teams to detect and treat infestations. Controlled experiments have shown that searchers' detection probabilities are substantially less than one under typical survey conditions. Detection rate is severely hampered in the presence of other yellow-flowering species or dense vegetation, and when Hawkweed is not flowering (Hauser *et al.* 2012).

Trained detection dogs rely upon scent rather than sight for detection and may thus offer a complementary approach to human searchers. In this study, we designed and conducted a set of Hawkweed search trials to assess the detection capabilities of a newly trained detection dog in the Alpine National Park. We placed potted Hawkweeds and control plants in replicated plots, recorded the path of the detection dog and her handler, and measured the time taken for the dog to detect the Hawkweeds. We use these data in a statistical model of time-to-detection and test the effect of Hawkweed species, Hawkweed abundance, background vegetation and weather variables on search performance. This model can be compared directly to the time-to-detection model for human searchers. We will discuss the circumstances in which a detection dog may be a cost-effective alternative to current survey methods, and give recommendations for deploying human and canine searchers in future surveys.

References

Hauser, C.E., Giljohann, K.M., Moore, J.L., Garrard, G., Kendal, D. & McCarthy, M.A. (2012) Observing the observer: experiments that estimate weed detectability. International Statistical Ecology Conference, July 3-6 in Sundvollen, Norway.

ID: 7

A Road Map for Designing and Implementing a Biological Monitoring Program

Biological monitoring programs, such as status and trends monitoring, are an omnipresent feature of many natural resource agencies. Statisticians working with these agencies typically provide assistance in two ways, designing data collection procedures and analyzing the collected data. Yet the impact of such assistance on actually improving the resulting management practices depends crucially on factors beyond data collection and analysis – clarity of management objectives, construction of conceptual and quantitative models linking actions to outcomes, the indicators selected for measurement, and effective communication of results to managers. In order to increase the impact of our statistical contributions, improve the scientific quality of monitoring programs, and, ultimately, to help our collaborators achieve their objectives, we've developed a Road Map that provides a practical, sequential planning structure for ensuring monitoring programs yield results that lead to improved management practices. The Road Map facilitates the important role of statisticians can play by including clarifying management objectives and the other factors listed above. The utility of the Road Map will be illustrated in application to the development of a monitoring program focused on entrainment of the delta smelt, an endangered fish in the San Francisco Bay Delta.

submitted by Ken Newman
Mathematical Statistician
US Fish & Wildlife Service

ken_newman@fws.gov

Optimal adaptive sampling design for occupancy monitoring

Kristin M. Broms [a], Mevin B. Hooten [c, b], and Ryan M. Fitzpatrick [d]

[a] Department of Fish, Wildlife, and Conservation Biology
Colorado State University
Fort Collins, CO
Kristin.broms@rams.colostate.edu

[b] Department of Fish, Wildlife, and Conservation Biology
Department of Statistics
Colorado State University
Fort Collins, CO
hooten@rams.colostate.edu

[c] U.S. Geological Survey
Colorado Cooperative Fish and Wildlife Unit
Fort Collins, CO

[d] Aquatic Wildlife Research Group
Colorado Parks and Wildlife
Fort Collins, CO
ryan.fitzpatrick@state.co.us

Keywords: design criterion; multi-species occupancy model; optimization; plains fish

Abstract: Static design-based approaches to sampling, such as stratified random sampling, generalized random tessellation sampling (GRTS, Stevens and Olsen 2004), or space-filling designs, are common methods to select sampling locations when little previous data exists. In contrast, adaptive sampling designs are an underutilized model-based approach to monitoring of ecological systems. Popularized through environmental monitoring schemes, they use the information from data that have already been collected for more efficient sampling going forward. Spatially explicit optimal adaptive monitoring designs have only recently been used in ecological studies (Hooten et al., 2009; Hooten et al., 2012). We demonstrate the components of an optimal adaptive sampling design using a Colorado Park and Wildlife monitoring program for plains fishes in the South Platte River basin in northeastern Colorado, USA. We develop a design criterion for the multi-species occupancy modelling framework that does not depend on the response variable directly. New optimally selected sampling locations are identified by minimizing this design criterion annually. We compare our optimal designs to random designs to illustrate the reduction in uncertainty and survey cost savings.

References

Hooten, M.B., C.K. Wikle, S. Sheriff, and J. Rushin (2009). Optimal spatio-temporal hybrid sampling designs for ecological monitoring. *Journal of Vegetation Science*, 20: 639-649.

Hooten, M.B., B.E. Ross, and C.K. Wikle. (2012). Optimal spatio-temporal monitoring designs for characterizing population trends. Gitzen, R.A., J.J. Millspaugh, A.B. Cooper, and

D.S. Licht (eds). In: *Design and Analysis of Long-Term Ecological Monitoring Studies*. Cambridge University Press.

Royle, J.A. (2002) Exchange algorithms for constructing large spatial designs. *Journal of Statistical Planning and Inference*, 100: 121-134.

Stevens Jr, D.L., and Olsen, A.R. (2004) Spatially balanced sampling of natural resources. *Journal of the American Statistical Association,* 99: 262-278.

Approximating survey standard errors in simulations evaluating optimal survey designs for estimating population change

J. H. Reynolds[a]

W. L. Thompson[b]

[a] Western Alaska Landscape Conservation Cooperative
Anchorage, Alaska, USA
joel_reynolds@fws.gov

[b] U.S. Fish and Wildlife Service
National Wildlife Refuge System
Hadley, MA, USA
bill_thompson@fws.gov

Keywords: monitoring design, survey design, power analyses, per-unit-variance, population process, observation process, simulation.

Abstract: Selecting a monitoring design to detect change through time in an ecological resource requires balancing the speed with which a given level of change can be detected against the cost of monitoring. Planning studies allow one to assess these tradeoffs and identify optimal design choices for a specific scenario of change. However, such studies are not often conducted. Even worse, they seem least likely to be conducted when they offer the most insight – when survey methods and monitoring designs are complex and not well captured by simple statistical models. In part, this is likely due to the time and effort required to conduct the extensive simulations usually required to explore the performance of the suite of observation process designs of interest over the relevant range of potential population process changes. We present an approach where a limited number of detailed simulations of the population and observation processes were conducted then the results used to empirically model the sampling distributions of the survey quantities of interest and their standard errors; these empirical models were then used to simulate, much more quickly, the performance of the full suite of observation process designs of interest over the relevant range of potential population process changes. A key feature of this approach was the use of the per-unit-sampling variance from Kish (1965) in developing the empirical model of the sampling distribution of the standard errors of interest for any sample effort level and population process state. The approach is illustrated in an investigation of competing designs for monitoring declining brown bear (*Ursus arctos)* densities in southwestern Alaska, the results of which were reported in Reynolds et al. (2011).

References

Kish, L. (1965) *Survey Sampling*. Wiley and Sons, New York.
Reynolds, J.H., Thompson, W.L, and Russell, B. 2011. Planning for success: identifying effective and efficient survey designs for monitoring. *Biological Conservation* 144, 1278-1284.

A Sampling Strategy Designed to Maximize the Efficiency of Data Collection of Food Web Relationships

K.E. Sauby [a] and M.C. Christman [b]

[a] Department of Biology
University of Florida
Gainesville, Florida, USA
ksauby@ufl.edu

[b] MCC Statistical Consulting LLC
Courtesy Associate Professor
Department of Biology and Department of Statistics
University of Florida
Gainesville, Florida, USA
marycchristman@gmail.com

Keywords: adaptive cluster sampling, *Opuntia*, cacti, cactus moth, *Cactoblastis cactorum*

Abstract: For studies involving populations of rare and/or clustered individuals, adaptive cluster sampling (ACS) can increase efficiency and reduce variance in comparison to designs such as simple random sampling. This sampling design involves first the collection of a primary set of units from the population according to a probability sampling scheme (such as simple random sampling without replacement). For any primary unit whose variable of interest satisfies a given criterion (for example, Species X is present or the abundance of Species X meets or exceeds a certain threshold number), neighboring units are selected until no additional neighboring units satisfy the criterion. A major disadvantage of ACS, however, is that the final sample size is unknown at the initiation of sampling. In this study we propose a restricted ACS design, where the number of sampled units neighboring any primary unit is limited to 12. We explore the properties of this sampling design and propose bias correction techniques. To illustrate this design and its properties, we present an example dataset collected according to restricted ACS to estimate occupancy and abundance of two species of native prickly pear cacti in Florida (*Opuntia* species; one of conservation concern) and two specialist cactus moths (the native moth *Melitara prodenialis* and the invasive South American *Cactoblastis cactorum*) that feed only on prickly pear cacti. We find that the Horvitz-Thompson mean estimator of prickly pear cactus occupancy (on which the sampling criterion is based) ranges from negative to positive bias as the primary sample size increases. In contrast, Horvitz-Thompson estimates of mean cactus insect occupancy are less biased.

Penalized Likelihood Methods for Occupancy Models

Rebecca A. Hutchinson [a,b], Jonathon J. Valente [c], Sarah C. Emerson [d],
Matthew G. Betts [c], and Thomas G. Dietterich [a]

[a] School of Electrical Engineering and Computer Science
[b] Department of Fisheries and Wildlife
[c] Department of Forest Ecosystems and Society
[d] Department of Statistics

Oregon State University
Corvallis, Oregon, USA

Corresponding author: rebecca.hutchinson@oregonstate.edu
jonathon.j.valente@gmail.com
emersosa@stat.oregonstate.edu
matthew.betts@oregonstate.edu
tgd@oregonstate.edu

Keywords: occupancy models, imperfect detection, penalized likelihood.

Abstract: Occupancy models are employed in species distribution modelling to account for imperfect detection during field surveys (Bailey et al 2014). While this approach is popular in the literature, problems can occur when estimating the model parameters. In particular, the maximum likelihood estimates can exhibit bias and large variance for data sets with small sample sizes, which can result in estimated occupancy probabilities near 0 and 1 ('boundary estimates'). In our recent work (Hutchinson et al 2015), we explored strategies for estimating parameters based on maximizing a penalized likelihood. Penalized likelihood methods augment the usual likelihood with a penalty function that encodes information about what parameter values are undesirable. We introduced penalties for occupancy models that have analogues in ridge regression and Bayesian approaches, and we compared them to a penalty developed for occupancy models in prior work (Moreno and Lele 2010) on both simulated and empirical data. We concluded that penalized methods may be of practical utility for fitting occupancy models with small sample sizes, and we released R code that implements these methods through the `unmarked` package (Fiske and Chandler 2011). In this presentation, we will review the findings from our paper and discuss considerations for choosing among these methods when modelling occupancy.

References

Bailey, L.L., MacKenzie, D.I., Nichols, J.D. (2014) Advances and applications of occupancy models. *Methods in Ecology and Evolution*, 5:1269–1279.

Fiske, I. & Chandler, R. (2011) unmarked: An R package for fitting hierarchical models of wildlife occurrence and abundance. Journal of Statistical Software, 43,1–23.

Hutchinson, R. A., Valente, J. J., Emerson, S. C., Betts, M. G., Dietterich, T. G. (2015) Penalized likelihood methods improve parameter estimates in occupancy models. *Methods in Ecology and Evolution*, 6:949–959.

Moreno, M. and Lele, S.R. (2010) Improved estimation of site occupancy using penalized likelihood. *Ecology*. 2:341–6.

Parameter Estimation Issues in a Simple Occupancy Model

A. Bowe [a] and B. Dennis [b]

[a] MS Candidate: Statistical Science
University of Idaho
Moscow, ID
abowe@uidaho.edu

[b] Department of Statistical Science
University of Idaho
Moscow, ID
brian@uidaho.edu

Keywords: estimability, occupancy modelling, data cloning.

Abstract: Occupancy modeling is becoming increasingly popular in wildlife management as a method of monitoring trends in wildlife populations. One of the primary motivations for using occupancy modeling is the ability to make inferences about large landscape patches with a reduced number of surveys. However, this increased versatility comes at the risk of degrading the estimability of the model parameters. Previous research (Hubbard 2014) has explored the presence of inherent identifiability issues in occupancy models, but little work has been done on the estimability the key parameters of these models: detection probability (p) and occupancy probability (psi).

We used a combination of bootstrapped profile likelihoods, data simulation and the data cloning techniques of Lele et al. (2010) to study estimability across a range of values for p and psi and for the number of sites and surveys. Preliminary results suggest estimability issues are present at smaller numbers of sites, or smaller numbers of surveys, and/or as either p or psi approaches the boundaries (0 or 1).

References

Hubbard, B. (2014). *Parameter Redundancy with Applications in Statistical Ecology* (Doctoral dissertation, University of Kent).

Lele, S. R., Nadeem, K., & Schmuland, B. (2012). Estimability and likelihood inference for generalized linear mixed models using data cloning. *Journal of the American Statistical Association.*

Application of a multi-level occupancy model with covariates to urban birds in South-Eastern Australia

W. A. Robinson [a] and G.W. Luck [a]

[a] Institute for Land Water and Society
Charles Sturt University
Thurgoona, New South Wales, Australia
wrobinson@csu.edu.au
galuck@csu.edu.au

Keywords: multi level model; occupancy; covariates, urban birds.

Abstract: In this applied research, we use multi-level occupancy models (Nicholls et al. 2008; Pavlacky et al. 2012) to estimate occupancy of common urban birds and their relationship with several environmental predictor variables at small and large scales. Large-scale occupancy (ψ) was estimated for 20 bird species in 18 towns spread across several thousand square kilometres in regional Victoria and New South Wales, Australia. The sampling was conducted over 4 time periods, in four independent neighbourhoods within each town and five accompanying neighbourhood-specific environmental covariates were measured in the field or calculated from GIS layers. We also modelled Small-scale occupancy (θ) (the occupancy of the species within a neighbourhood, given that it is present in the town), and the detection probability (p) within a neighbourhood when present in the town. We used AICc to compare the base model, $\{\hat{\psi}, \hat{\theta}, \hat{p}\}$ to models containing $\{\hat{\psi}, \ \hat{\theta} \times Cov_{1\ to\ 5}, \ \hat{p}\}$ and $\{\hat{\psi}, \hat{\theta} \times Cov_x, \hat{p}_{1\ to\ 5}\}$ to identify environmental covariates that improved the base model. Overall, few species had improved models from fitting the covariates, but those that did confirmed intuitive expectations. For example, canopy feeding species like striated pardolotes were more likely to occur in neighbourhoods with higher density of native vegetation, and more difficult to detect in neighbourhoods that had higher levels of impervious surfaces (more urban development).

References:

Nichols, J. D. L. L. Bailey, A. F. O'Connell, N. W. Talancy, E. H. C. Grant, A. T. Gilbert, E. M. Annand, T. P. Husband, and J. E. Hines. 2008. Multi-scale occupancy estimation and modelling using multiple detection methods. Journal of Applied Ecology 45:1321-1329

Pavlacky, D. C., J. A. Blakesley, G. C. White, D. J. Hanni, and P. M. Lukacs. 2012. Hierarchical multi-scale occupancy estimation for monitoring wildlife populations. Journal of Wildlife Management 76:154-162.

Uncertainty in biological monitoring: a framework for data collection and analysis to account for sampling bias

V. Ruiz-Gutierrez[a], **M.B. Hooten**[b] and **E.H. Campbell Grant**[c]

[a]Conservation Science Program
Cornell Lab of Ornithology, Cornell University
Ithaca, NY
vr45@cornell.edu

[b]U.S. Geological Survey,
Colorado Cooperative Fish and Wildlife Research Unit
Department of Statistics,Colorado State University
Fort Collins, CO
mevin.hooten@colostate.edu

[c]U.S. Geological Survey,
Patuxent Wildlife Research Center
S.O. Conte Anadromous Fish Laboratory
Greenfield, MA
ehgrant@usgs.gov

Keywords: detection, false negative, false positive, monitoring, occupancy, probit link

Abstract: Biological monitoring programs are increasingly relying upon large volumes of citizen science data to improve the scope and spatial coverage of information, challenging the scientific community to develop model-based approaches to improve inference. Recent statistical models in ecology have been developed to accommodate false-negative errors, although recent work points to false positive errors as equally important sources of bias. This is of particular concern for the success of any monitoring program given rates as small as 3% could lead to the overestimation of the occurrence of rare events by as much as 50%, and even small false positive rates can severely bias estimates of occurrence dynamics. We present an integrated, computationally efficient Bayesian hierarchical model to correct for false positive and negative errors in detection/non-detection data. Our model combines independent, auxiliary data sources with field observations to improve the estimation of false positive rates, when a subset of field observations cannot be validated *a posteriori* or assumed as perfect. We evaluated the performance of the model across a range of occurrence rates, false positive and negative errors, and quantity of auxiliary data. The model performed well under all simulated scenarios, and we were able to identify thresholds for biological scenarios and quantity of auxiliary data required for improved inference. We applied our model to a large-scale, citizen-science monitoring program for anurans in the Northeastern U.S., using auxiliary data from an experiment designed to estimate false positive error rates. Not correcting for false positive rates resulted in biased estimates of occupancy in 4 of the 10 anuran species we analyzed, leading to an overestimation of the average number of occupied survey routes by as much as 70%. The framework we present for data collection and analysis is able to provide reliable inference for occurrence patterns using data from a citizen-science monitoring program. However, our approach is applicable to data generated by any type of research and monitoring program, independent of skill level or scale, when effort is placed on obtaining independent information on false positive rates.

Teasing apart climatic and habitat effects on the distribution dynamics in the Swiss breeding bird community

Jérôme Guélat [a] and Marc Kéry [b]

Swiss Ornithological Institute
Sempach, Switzerland
[a] jerome.guelat@vogelwarte.ch
[b] marc.kery@vogelwarte.ch

Keywords: occupancy, birds, colonization, extinction, distribution dynamics.

Abstract: A good knowledge of species distribution dynamics is fundamental to understand many ecological processes and also to evaluate the conservation status of bird species. We combined three disparate datasets (two breeding bird atlases and one common breeding bird survey scheme) in a single modeling framework to understand the drivers of the national distribution dynamics of the community of Swiss breeding birds during the last 22 years. We used dynamic site occupancy models (MacKenzie *et al.* 2003) which allow us to explicitly express occurrence change as a function of colonization and persistence processes, while accounting for presence/absence measurement error (i.e., imperfect detection). We used climatic covariates to model the colonization and extinction probabilities to investigate possible effects of climate change. In particular we quantified the relative importance of elevation and temperature on the dynamics and related it to different species traits in the whole avian community. We also studied the sensitivity of distribution dynamics to these variables. Site occupancy models also provide an explicit way of modeling the observation process. This allows us to account for the differences in the sampling protocols and in the observation intensity during the whole study period.

References

MacKenzie, D.I., Nichols, J.D., Hines, J.E., Knutson, M.G. & Franklin, A.B. (2003). Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. *Ecology*, **84**, 2200–2207.

Stretching applications of occupancy models: from local knowledge to monitoring crime

Koustubh Sharma[a], Belinda Wright[b] and Julia Taubmann[c]
[a] Snow Leopard Trust
Seattle, USA
koustubhsharma@gmail.com

[b] Wildlife Protection Society of India
New Delhi, India
belinda@wpsi-india.org

[c]University of Freiburg
Germany
j.taubmann@googlemail.com

Keywords: snow leopard; tiger; distribution; poaching; large datasets; monitoring.

Abstract:

Large amount of data are being constantly generated by various government and non-government agencies across the world. Most of these datasets are at best used to provide trends and patterns without any attention to the uncertainties caused by imperfect detectability. In absence of empirical estimation of detection probabilities, the inferences are often unreliable and nonreplicable. Occupancy based models provide the flexibility to relax certain assumptions and redefine interpretations. We use two different datasets to demonstrate the power of occupancy analysis in making meaningful inferences – one using community knowledge in Alay Mountain range over a period of 20 years, and other using reports on tiger crime from India across 40 years.

In the Alay Mountains spanning 14,000 km^2, information on historic and current sightings of snow leopard and other sympatric species were collected by interviewing local community members within 49 400-km^2 grids. Analysis of data revealed high probabilities of local extinction in 39% sites for snow leopard, and 84% and 47% respectively for brown bear and Marco Polo sheep. Lynx was the only species that seemingly colonized about 41% of the areas. The assumption of closure of sites to changes was relaxed, hence necessitating interpretation of occupancy probability as probability of site use. Probabilities of colonization and extinction were reinterpreted as probabilities of moving in/abandoning of areas by the species of interest. Factors such as profession, experience and lapsed time affected probability of detection and retention of information by individual respondents, used as survey points in the modelling process.

Wildlife crime such as poaching and illegal trafficking poses a significant threat to wildlife. An increase in the number of reports of wildlife crime does not essentially translate into an increase in the crime rate. It can be a factor of better policing and detection, and/or a true increase in the instances of wildlife crime. We used reports of tiger crime collected over the past 40 years to estimate probabilities of occurrence and detection of tiger crimes across the 605 districts of India. Analysis not only identifies 73 tiger crime hotspots across the country, but also helps establish certain hypotheses of preferred modus operandi of poachers and traffickers using information theoretic approach. Distinct estimation of probability of detecting tiger crime during a year as opposed to probability of its occurrence within a district in a stipulated period (3-7 years) provides empirical basis of monitoring a threat and the efficiency of addressing it.

Stretching the application of occupancy models opens up avenues for utilizing large datasets in making robust and reliable inferences to monitor species distributions and changes across large spatial scales and crime across spatial and temporal scales.

Extensions to the Hybrid Symbolic-Numerical Method Investigating Identifiability

D. J. Cole[a] and R. Choquet[b]

[a] School of Mathematics, Statistics and Actuarial Science,
University of Kent, Canterbury, England
d.j.cole@kent.ac.uk

[b] Centre d'Ecologie Fonctionnelle et Evolutive, Le Centre national de la recherche scientifique, Montpellier, France
Remi.CHOQUET@cefe.cnrs.fr

Keywords: identifiability; mark-recovery; parameter redundancy; profile.

Abstract:

If two parameters are confounded so that they only ever appear in a model as a product, then it is not possible to estimate the two parameters individually – only their product. Such a model is termed parameter redundant and will be non-identifiable. A non-identifiable model does not have a unique set of maximum likelihood estimates and its standard errors are undefined. In practice, most ecology models are fitted using numerical methods so this can be missed, leading to biased and inaccurate standard errors as well as incorrect biological interpretations. So it is therefore essential to know whether or not a model is non-identifiable.

In complex models it is not always as obvious whether a model is non-identifiable. In such cases, symbolic methods can be used to detect parameter redundancy and also give combinations of parameters that can be estimated, as well as provide general results for specific classes of models (see for example Cole *et al,* 2010). Symbolic methods are not straightforward to use, especially for someone with a non-mathematical background, as they require the use of a symbolic algebra package such as Maple. Symbolic methods are not the only way to detect parameter redundancy; numeric methods exist as well, though these can be inaccurate. A hybrid symbolic-numeric method has been developed that is easier to apply and more accurate (Choquet and Cole, 2012), but cannot be used to find combinations of parameters that can be estimated, and derive general results.

Here we present extensions to the hybrid symbolic-numeric method that can be used to find estimable combinations of parameters, and derive general results. This creates a method that has all the functionality of the symbolic method but is much simpler to use. It can be executed using R rather than a symbolic algebra package, creating an automatic method that can be added to existing software. We illustrate these methods using mark-recovery models, though the methods are generally applicable.

References

Choquet, R. and Cole, D.J. (2012) A Hybrid Symbolic-Numerical Method for Determining Model Structure. *Mathematical Biosciences*, 236, 117-125.

Cole, D. J., Morgan, B. J. T and Titterington, D. M. (2010) Determining the Parametric Structure of Non-Linear Models. *Mathematical Biosciences*, 228, 16-30.

A state-space autoregressive model to quantify patterns of ontogenetic growth

M. E. Brooks [a], C. Clements [a], and A. Ozgul [a]

[a] Department of Evolutionary Biology and Environmental Studies
University of Zurich
Zurich, Switzerland
Mollie.Brooks@ieu.uzh.ch
Chris.Clements@ieu.uzh.ch
Arpat.Ozgul@ieu.uzh.ch

Keywords: linear mixed models; maximum marginal likelihood; empirical Bayes; TMB; cost of reproduction.

Abstract:

Demographic rates of individuals vary due to many factors (e.g. age, environment, genetics, disease, energy allocation), some of which are more easily observable than others. Known covariates are commonly included in models, but recently, interest has been increasing for ways to quantify heterogeneity that may be due to unobserved factors.

Here, we present a state-space (S-S) autoregressive approach for quantifying patterns of ontogenetic growth. The work is motivated by a long-term study of Soay sheep in St. Kilda archipelago where recapture rates are around 50%. Unlike other growth models, the S-S approach simultaneously allows for the use of non-consecutive observations due to imperfect recapture and allows for the inclusion of carryover effects from temporally varying covariates (e.g. population density, weather, and reproductive effort). We compare this approach to an autoregressive growth model fit in a linear mixed model (LMM) framework, which cannot use non-consecutive observations. We compare the S-S and LMM approaches for the sheep data and for simulated datasets with varying recapture rates and observation error.

We fit the S-S models using maximum marginal likelihood (i.e. empirical Bayes) in the new R package TMB. This package is typically 100 to 500 times faster than Gibbs sampling approaches that are more traditionally used for models with latent variables. Each model contained more than 10,000 latent variables and took only 20 seconds to converge. The speed allowed us to do extensive model selection on covariates for the sheep study.

In the sheep study, the S-S model showed that female growth rates decreased with age, population density, and while lactating. After accounting for the effects of covariates, individual heterogeneity was greater than temporal heterogeneity, residual process error, and the effect size of weather. The LMM, on the other hand, failed to capture the reduction of growth in females that allocate energy to reproduction. The LMM had lower estimated individual heterogeneity, but higher temporal heterogeneity.

For the simulated datasets, S-S models were less biased than LMMs. With low recapture rates, LMMs tended to underestimate the intercept, overestimate the autoregressive term, and underestimate the random effect of individual. Observation error caused LMMs to be biased in the opposite direction to imperfect recapture rates.

Our S-S approach is more flexible than size-at-age models and more robust than LMMs. Improved estimates of individual heterogeneity may have made it possible to detect the cost of reproduction in sheep, unlike in previous analyses of the same population. Using TMB allowed us to achieve the flexibility and accuracy of S-S models at a speed more comparable to LMMs.

Quantifying nonbreeding from mark-recapture data and nest records

A.M. Lee[a,b], **F. Daunt**[c], **S. Wanless**[c] and **J. Reid**[a]

[a]Institute of Biological and Environmental Sciences
University of Aberdeen
Aberdeen, UK
lee@alumni.ntnu.no

[b]Department of Environmental Science, Policy & Management
University of California, Berkeley
Berkeley, USA

[c]Centre for Ecology & Hydrology, Bush Estate, Penicuik, Midlothian
Penicuik, UK

Keywords: population ecology, mark-recapture, floaters, nonbreeding,

Abstract: Many populations, particularly of birds, contain adult nonbreeding individuals that do not initiate breeding in a given year. These nonbreeders can be of several types and belong to different life stages. Understanding the composition of the nonbreeding segment of populations is important to accurately predict future population dynamics and responses to environmental change. However, because population censuses and studies have traditionally focused on breeders, and because nonbreeders tend to be more difficult to detect, data on nonbreeders are often lacking. A long-term study of a population of European shags (*Phalacrocorax aristotelis*) on the Isle of May, Scotland has uncovered evidence of lowered breeding activity in years with adverse environmental conditions. Separating nonbreeders from breeders by direct observation is often impossible in this population, making it difficult to establish whether the lowered breeding activity is caused by delayed recruitment of young individuals, skipped breeding years by experienced breeders, or late life individuals that cease breeding prior to death. Using mark-recapture models (Cormack-Jolly-Seber and multi-event) and nest records, we have combined a variety of methods to quantify the frequency of different types of nonbreeding in this population.

Using continuous-time spatial explicit capture-recapture (SECR) models to learn about animal activity.

G. Distiller[a] and **D. Borchers**[b]

[a]Department of Statistical Sciences
Statistics in Ecology, Environment and Conservation (SEEC)
University of Cape Town
Greg.Distiller@uct.ac.za

[b]School of Mathematics and Statistics
Centre for Research into Ecological and Environmental Modelling (CREEM)
University of St Andrews
dlb@st-andrews.ac.uk

Keywords: continuous-time spatial capture-recapture, modelling animal activity

Abstract: While spatially explicit capture-recapture (SECR) models were primarily developed to estimate density they also provide a framework to investigate animals' activity and space use patterns (Royle et al. (2013), Sutherland et al. (2015)). Continuous-time SECR models (Borchers et al. 2014) provide a tool for investigating activity in the temporal domain. We develop SECR models with flexible hazard functions that allow us to draw inference about animal activity patterns through time. We illustrate the methods using data from a camera trap study of jaguars and a baited single-catch trap study of possums, and model the detection hazard using regression splines to allow flexible modelling of the daily detection hazards for these two species. With the incorporation of spatial or habitat variables in the temporal detection hazard, this approach has potential to provide a tool for investigating how habitat use changes continuously in time, over any time scale.

References

Borchers, D.L., Distiller, G., Foster, R., Harmsen, B., and Milazzo, L. (2014) Continuous-time spatially explicit capture-recapture models, with an application to a jaguar camera-trap survey. *Methods in Ecology and Evolution*, 5(7), 656-665.

Royle, J. A., Chandler, R. B., Gazenski, K. D., and Graves, T. A. (2013). Spatial capture-recapture models for jointly estimating population density and landscape connectivity. *Ecology*, 94, 287-294.

Sutherland, C., Fuller, A. K., and Royle, J. A. (2015). Modelling non-euclidean movement and landscape connectivity in highly structured ecological networks. *Methods in Ecology and Evolution*, 6(2), 169-177..

A Spatially-Explicit Robust-Design Capture-Recapture Model Applied to Dolphin Photo ID Data Collected Following the *Deepwater Horizon* Spill

Trent L. McDonald[a], Fawn Hornsby[a], Todd Speakman[b], Eric Zolman[b], Keith D. Mullin[c], Carrie Sinclair[c], Patricia Rosel[d], Len Thomas[e], Lori Schwacke[b]

[a]Western EcoSystems Technology, Inc.
Laramie, Wyoming, USA
tmcdonald@west-inc.com

[b]National Centers for Coastal Ocean Science
National Oceanic and Atmospheric Administration
Hollings Marine Laboratory, Charleston, South Carolina, USA

[c]Southeast Fisheries Science Center
National Marine Fisheries Service
Pascagoula, Mississippi, USA

[d]Southeast Fisheries Science Center
National Marine Fisheries Service
Lafayette, Louisiana, USA

[e]Centre for Research into Ecological and Environmental Modelling
University of St. Andrews,
St. Andrews, Scotland

Keywords: mark-recapture; dolphins; *Deepwater Horizon* oil spill

Abstract: In recent years, spatially-explicit capture-recapture models have become popular, and represent a significant advance in analysis when capture locations are known. Most spatial capture models treat activity centers as latent variables, estimate them, and adjust capture probabilities accordingly. In closed models, this allows density to be estimated directly without hard study area boundaries. In open models, spatial capture models account for occasions when animals are off the study area and nearly un-catchable, which adjusts survival rates for estimated immigration and emigration. Here, we combine open and closed models for robust-design (Kendall et al. 1997) data by replacing the closed portion with a spatially-explicit model while leaving the open portion intact. We applied this model to dolphin photo ID data collected in two estuaries following the *Deepwater Horizon* oil spill in 2010, and present survival, density and abundance estimates.

References

Kendall, William L., Nichols, James D., and Hines, James E. 1997. Estimating temporary emigration using capture-recapture data with Pollock's robust design. Ecology, 78(2), 563-578.

Capture-recapture estimation based upon zero-truncated and one-inflated count data

P. Kaskasamkul[a] and **D. Böhning**[b]

Mathematical Sciences
University of Southampton
Southampton, England
[a]pk3e13@soton.ac.uk
[b]d.a.bohning@soton.ac.uk

Keywords: capture-recapture; ratio plot; one-inflated count data, one-truncated geometric model; zero-truncation one-inflation geometric model

Abstract: The purpose of capture-recapture studies is to estimate the size of an elusive target population. This is frequently accomplished by means of a zero-truncated count distribution where the count represents the number of times the respective member of the target population has been identified. In practice, however, many researchers working in the field of capture-recapture experience data with an excess of counts of ones. This might happen by the nature of data or by the fact that the data are classified into the wrong group which shows a form of one-inflation. Consequently, some classical estimators overestimate, particularly, we see severe overestimation bias in Chao's lower bound estimator which is widely used, and even worse, used in the belief that it represents a valid lower bound of the true population size. To cope with this problem, three estimators are developed from two models; the first two estimators are modified versions of Turing's and the maximum likelihood estimator, both in the context of a one-truncated geometric model. The third estimator implements a full maximum likelihood approach for the zero-truncated one-inflated geometric model by using a nested EM algorithm. The new estimators are compared with existing conventional estimators in both, simulations and applications. The results reveal that the suggested estimators might provide considerable benefits if compared to commonly used estimators.

Confounding in dynamic ecological processes

T.J. Hefley [a], M.B. Hooten [b,a], E.M. Hanks [c], D.P. Walsh [d], and R.E. Russell [d]

[a] Department of Fish, Wildlife, and Conservation Biology
and
Department of Statistics
Colorado State University
Fort Collins, Colorado, USA
trevor.hefley@colostate.edu

[b] U.S. Geological Survey, Colorado Cooperative Fish and Wildlife Research Unit
Colorado State University
Fort Collins, Colorado, USA
mevin.hooten@colostate.edu

[c] Department of Statistics
Pennsylvania State University
State College, Pennsylvania, USA
hanks@psu.edu

[d] U.S. Geological Survey, National Wildlife Health Center
Madison, Wisconsin, USA
rerussell@usgs.gov
dwalsh@usgs.gov

Keywords: ecological diffusion; generalized linear mixed model; spatial confounding.

Abstract: Ecological processes are dynamic in space and time. The dynamics generate dependence among observations of the system. Analyzing ecological data often requires modeling dependencies created by these processes. In many applications, the generalized linear mixed model (GLMM) is used with a random effect to account for such dependence. Covariates are often included in the fixed effects of a GLMM and may be collinear with the random effects, which can complicate the implementation and affect inference (e.g., Hodges and Reich, 2010; Hanks et al. 2015). We demonstrate the potential for confounding in traditional covariance-based spatial GLMMs and propose a constructive approach in which a dynamic spatio-temporal model (e.g., statistical partial differential equation model) that explicitly incorporates spatial covariates serves as the random effect. We illustrate our approach using spatially varying ecological diffusion to model the spatial dynamics of chronic wasting disease in white-tailed deer in Wisconsin.

References

Hanks, E. M., Schliep, E. M., Hooten, M. B., and Hoeting, J. A. (2015). Restricted spatial regression in practice: geostatistical models, confounding, and robustness under model misspecification. *Environmetrics* 26, 243-254.

Hodges, J. S. and Reich, B. J. (2010). Adding spatially-correlated errors can mess up the fixed effect you love. *The American Statistician* 64, 325-334.

Network shape influences the strength of growth and phenology responses to climate change in juvenile salmon

A.H. Fullerton [a], B.J. Burke [a], J.J. Lawler [b], and C.E. Torgersen [c]

[a] Fish Ecology Division
NOAA Northwest Fisheries Science Center
Seattle, Washington
aimee.fullerton@noaa.gov
brian.burke@noaa.gov

[b] School of Environmental and Forest Sciences
University of Washington
Seattle, Washington
jlawler@u.washington.edu

[c] USGS Forest and Rangeland Ecosystem Science Center
Cascadia Field Station, University of Washington
Seattle, Washington
cet6@u.washington.edu

Keywords: individual based model; thermal heterogeneity; spatial structure

Abstract: The diversity of freshwater habitats to which Pacific salmon are adapted may afford some protection against the potentially deleterious impacts expected as a result of climate change (e.g., reduced growth and survival, altered phenology). However, it is unclear what role the spatial properties of river networks might play in determining risk. Specifically, the spatial structure of a stream network dictates, to some extent, its thermal heterogeneity patterns. Patchiness in water temperature may provide opportunities for salmon to behaviorally optimize their growing conditions. We hypothesized that (a) climate change will alter growth (such that, depending on initial conditions, change could be positive or negative) and cause earlier fry emergence and smolt outmigration; but that (b) these changes will be smaller in networks having high spatial structure (e.g., dense tributary junctions). We developed a spatially structured individual based model (spatial scale = 100 1-km reaches; temporal scale = 730 12-h time steps) in which growth and movement decisions were governed by water temperature and conspecific density. We evaluated growth and phenology of Chinook salmon under 6 climate scenarios in each of 3 networks of increasing spatial complexity – long, typical, and compact. Despite similar thermal conditions and egg placement across networks, differences in growth accumulated throughout the growing season. Contrary to our expectation, salmon grew best and more fish reached a threshold size needed to outmigrate in the long network (i.e., the shape having the lowest spatial complexity). Moreover, ontogenetic shifts occurred earlier in the long network. Across the climate scenarios, results were similar for the long and typical networks but differed in the compact network. This model provides a powerful tool for developing hypotheses about how fish may respond to altered thermal regimes given different spatial contexts. Results from carefully designed simulations can be used to convert existing conservation plans into climate adaptation plans, and to prioritize locations for restoration activities intended to maintain or create thermal diversity.

A spatio-temporal marked point process model

Charlotte M Jones-Todd
cmjt@st-andrews.ac.uk

Abstract

I consider a novel marked spatio-temporal point process model which deals with complex dependency structures inherent in many ecological data, and focus on how such an approach can be used to infer about dependency between the latent spatial processes which drive the spatial distribution of marks conditional on their location over time. This modelling technique is applied to *Azteca sericeasurs* nest data from a 300 hectare coffee plantation situated in the Soconusco region of Chiapas, Mexico.

Using the Stochastic Partial Differential Equation (SPDE) approach for estimating the latent spatio-temporal processes, and `R-INLA` we fit the proposed model in a Bayesian framework. The modelling procedure incorporates a new approach to fitting a classic point process model to the data, additionally enabling us to infer the relationship between *A. sericeasurs* nests and *Coccus viridis* infested coffee bushes accounting for the spatial and temporal correlation conditional on their location.

A spatio-temporal exposure-hazard model for assessing the impact of GM plants' pollen on non-target organisms

E. Walker [a,b*], S. Soubeyrand [a], M. Leclerc [c], J-F. Rey [a] and A. Messéan [b]

[a] INRA, BioSP
UR546 Biostatistics and Spatial Processes,
84914 Avignon, France
Emily.walker@paca.inra.fr

[b] INRA, Eco-Innov
UAR1240 Eco-Innov (Impacts écologiques des innovations en production végétale),
78850 Thiverval- Grignon, France

[c] INRA, IGEPP
UMR 1349 IGEPP,
35653 Le Rheu, France

Keywords: Exposure-hazard model, landscape simulation, particle dispersal, point processes, pollutant source, risk assessment, stochastic geometry, xenobiotics, spatially and temporally-explicit model, spatial ecology.

Abstract: The cultivation of Genetically Modified (GM) crops within agricultural landscapes may represent a potential hazard that might have substantial impacts on the populations of non-target organisms (NTOs) living among agroecosystems. The exposure-hazard model we present is based on (i) tools of stochastic geometry (marked polygon and point processes) for structuring the landscape and describing the location of exposed individuals, (ii) a method based on a dispersal kernel describing the dissemination of contaminant particles from polygon sources, and (iii) an (eco)toxicological equation describing the toxicokinetics and dynamics of contaminants in affected individuals of the exposed population. The design of the model was motivated by concerns raised about the impact of Genetically Modified (GM) maize pollen on non-target species, namely non-target Lepidoptera. Simulations were run to analyse results in terms of exposed individuals' mortality in several realistic situations. However, the model can cope with a wide range of hazards whose spatial patterns need to be considered, such as toxic emissions from industrial areas having health effects to surrounding populations or pathogen spread from wild infectious populations to livestock or crops.
The model is implemented in an R package named *SEHmodel* (Spatial Exposure-Hazard model) that contains functions and methods for quantifying spatio-temporal variation in contamination risk around known polygon sources of contaminants, and quantifies the impact of the contaminants on the surrounding population of individuals which are located in habitat areas and are susceptible to the contaminants.

How does the landscape mediate the effects of thinning on productivity in young growth forests: An investigation and implications for management

K.L Oken [a], D.V. D’Amore [b], E.A. Steel[c]

[a] Quantitative Ecology & Resource Management
University of Washington
Seattle, Washington, USA
okenk@uw.edu

[b] Pacific Northwest Research Station
USDA Forest Service
Juneau, Alaska, USA
ddamore@fs.fed.us

[c] Pacific Northwest Research Station
USDA Forest Service
Seattle, Washington, USA
asteel@fs.fed.us

Keywords: forests, growth, heterogeneity, thinning, landscape, ecosystem services

Abstract: Young growth forests represent a critical nexus of a wide variety of ecosystem services including natural resources in the way of wood products, as well as conservation services such as carbon sequestration. However, our ability to optimize the realized value of these forests is compromised by their extreme variability in growth rates and inconsistency of responses to management. Therefore, in order to optimize forest value, a rigorous analytical approach is needed to provide a more thorough understanding of the patterns and drivers of productivity of young growth forests and how such stands respond to different management strategies. Many variables impact overall stand productivity; managers can control some of these variables, such as intensity of pre-commercial thinning, whereas others are innate characteristics of the landscape, such as the underlying hydrology and soil types. Data gathered through long-term forest plot measurements can provide a rich foundation for investigating trends in growth and site response within forested areas. We analysed a unique long-term data set measuring growth of trees in young growth stands in 284 plots throughout the extensive perhumid coastal temperate rainforest of Southeast Alaska. Stands were measured approximately annually for 30-100 years, and were set up in a randomized block design to measure the effect of four different thinning intensities. Using mixed-effects models, we show that the most intense thinning treatment decreases the carbon content within a stand by nearly 40% even 40 years after treatment because the faster growth rates of individual trees cannot compensate for the fewer trees in thinned stands. However, the extreme variability among stands presents an impediment to applying the results to specific management objectives. Therefore, we explored how landscape factors such as soil water content mediate the effect of thinning on stand productivity, and also consider the consequences of this spatial heterogeneity for management.

Hierarchical Bayesian models to quantify forest dynamics at the scale of individual trees from remote sensing data

T. Trevor Caughlin [a]

[a]School of Forest Resources & Conservation
University of Florida
Gainesville, FL
trevorcaughlin@ufl.edu

Keywords: Landsat, state-space model, forest dynamics, reforestation, hierarchical Bayes.

Abstract: Forest dynamics drive large-scale ecological processes, including theterrestrial carbon cycle, forest succession, and biodiversity maintenance at landscape, regional and continental scales. Individual tree vital rates (growth, survival and reproduction) underlie these dynamics, and are primarily estimated from repeated measurements of tree stems in forest inventory plots. However, logistical constraints limit the spatial coverage of these field measurements and extrapolation to larger scales can propagate spatial errors. Remote sensing data could provide a solution to the spatial mismatch between field data and large scale forest dynamics.

I demonstrate a hierarchical Bayesian approach to link remote sensing and forest inventory data to predict forest dynamics across a heterogeneous landscape covering a 3809 km^2 area in western Panama. I use generalized linear models to predict LiDAR-derived metrics of forest structure, including tree canopy cover and tree height, from 30 x 30 m resolution data from the Landsat archive. I then develop a state-space model with tree canopy cover and tree height as output variables and tree vital rates as input variables. This modelling approach can incorporate data on individual trees from forest inventory plots to inform demographic parameters over large spatial extents. Bayesian methods enable a clear route to propagating uncertainty in parameter estimation throughout the model, including measurement error from remote sensing classification. I apply the model to quantify whether forest succession over a ten year period was limited more by seed arrival or by the growth and survival of established trees. Meeting the demand for large-scale predictions of forest dynamics will require synthesizing site-specific data across a range of empirical studies. Statistical models that integrate multiple data sources to make spatially-explicit forecasts over large areas could support the UN's target of restoring hundreds of millions of hectares of degraded land within the next 15 years, and help mitigate global climate change.

A novel method to improve estimates of predator diet compositions

P.E. Moriarty[a], T.E. Essington,[b] and E.J. Ward[c]

[a] School of Aquatic and Fishery Sciences
University of Washington
Seattle, WA
pmoriart@uw.edu

[b] School of Aquatic and Fishery Sciences
University of Washington
Seattle, WA
essing@uw.edu

[c] Northwest Fisheries Science Center
National Oceanic and Atmospheric Administration
Seattle, WA
eric.ward@noaa.gov

Keywords: food webs; diet estimation; mixture model; maximum likelihood.

Abstract: Stomach content data is frequently used to study predator diets. Unfortunately, these data pose several statistical challenges for analysis and estimation that have hindered our ability to create quantitative diet estimates from stomach content data. For instance, extreme events in predator consumption are common, and there may be covariance between amount of food consumed and the type of food consumed. Existing methods do not effectively address these challenges, nor do they allow for the ability to create an appropriate likelihood function for diet estimates, which would allow for formal model selection and likelihood-based parameter estimation. We developed a novel model that addresses these challenges to quantitatively estimate the proportion of a predators' diet that consists of a given prey type. Simulation testing of our model indicates that our predator diet proportions have comparable accuracy and increased precision as compared to a (stomach-mass) weighted mean, and are also an improvement over averaging individual sample diet proportions. Additionally, we apply our model to stomach content data for multiple predators to compare the accuracy and precision of prey contribution estimates to traditional diet estimation methods. When applied to common prey items, the accuracy and precision of our model is similar to the weighted mean. However, in the presence of complex cases of extreme events and covariance in the real datasets, these estimation methods sometimes differ in accuracy and precision. Our model represents an important step in developing statistical methods to address the challenges of stomach content data within a likelihood framework, so that the resulting diet proportion estimates can be more widely used.

References

Pirotta, E., Matthiopoulos, J., Mackenzie, M.L., Scott-Hayward, L. & Rendell, L. (2011) Modelling sperm whale habitat preference : a novel approach combining transect and follow data. *Marine Ecology Progress Series*, 436:257–272.

Wang, Z., Bovik, A.C., Sheikh, H.R., Member, S., Simoncelli, E.P. & Member, S. (2004) Image Quality Assessment : From Error Visibility to Structural Similarity. *IEEE transactions on image processing*, 13:600–612.

Testing and recommending methods for fitting size spectra to data

Andrew M. Edwards[a,b], **James P. W. Robinson**[b], **Michael J. Plank**[c],
Julia K. Baum[b] **and Julia L. Blanchard**[d]

[a]Pacific Biological Station, Fisheries and Oceans Canada,
Nanaimo, British Columbia, Canada.
andrew.edwards@dfo-mpo.gc.ca

[b]Department of Biology, University of Victoria,
Victoria, British Columbia, Canada.
jamespwr@uvic.ca, baum@uvic.ca

[c]School of Mathematics and Statistics, University of Canterbury, Christchurch, New Zealand, and
Te Pūnaha Matatini, a New Zealand Centre of Research Excellence.
michael.plank@canterbury.ac.nz

[d]Institute for Marine and Antarctic Studies, University of Tasmania,
Hobart, Tasmania, Australia.
julia.blanchard@utas.edu.au

Keywords: individual size distribution; ecosystem indicators; ecosystem approach to fisheries; biomass size spectrum; abundance size spectrum; bounded power-law distribution; truncated Pareto distribution.

Abstract: A size spectrum characterises the distribution of a property, such as body size, across an ecological community. Size spectra are often used as ecosystem indicators of marine systems. They have been fitted to data from various sources, including groundfish trawl surveys, visual surveys of fish in kelp forests and coral reefs, sediment samples of benthic invertebrates and satellite remote-sensing of chlorophyll. Over the past decade several methods have been used to fit size spectra to data. We document eight such methods and test their accuracy on simulated data. We find that maximum likelihood estimation is the only method that is consistently accurate and that yields reliable confidence intervals for the exponent of the individual size distribution (which is related to the slope of the size spectrum). We apply the eight methods to a twenty-year time series of bottom-trawl fish survey data. The choice of method directly influences the ecological conclusion of whether or not fishing has impacted the community structure. Furthermore, maximum likelihood estimation is the only method that can properly account for the binned nature of survey data, and we therefore recommend its use for fitting size spectra and providing advice to fisheries managers.

A multistate dynamic occupancy model for spatially aggregated sessile communities

K. Fukaya[a], **J. A. Royle**[b], **T. Okuda**[c], **M. Nakaoka**[d], and **T. Noda**[e]

[a]The Institute of Statistical Mathematics
Tokyo, Japan
kfukaya@ism.ac.jp

[b]USGS Patuxent Wildlife Research Center
Laurel, MD, USA
aroyle@usgs.gov

[c]National Research Institute of Far Seas Fisheries, Fisheries Research Agency
Kanagawa, Japan
okudy@affrc.go.jp

[d]Hokkaido University
Hokkaido, Japan
nakaoka@fsc.hokudai.ac.jp

[e]Hokkaido University
Hokkaido, Japan
noda@ees.hokudai.ac.jp

Keywords: Classification error; Community dynamics; Multistate dynamic occupancy model; Spatial heterogeneity

Abstract: Markov models have been applied to sessile assemblage data to infer transition probabilities. Estimates may be biased, however, when the classification error of the occupancy state exists, which lead to developments of new methods that account for such an observation error (Conway-Cranos and Doak 2011, Fukaya and Royle 2013). In these methods, however, any spatial structure of the community is not considered explicitly. Since sessile organisms typically represent a spatially aggregated pattern, it seems likely that the local structure of the community composition affects observation, state transitions, or both. Hence, accounting for such spatial dependence may be a key to capture underlying community dynamics and observation processes accurately. In this study, we propose a hierarchical model to estimate Markov community dynamics of sessile species which explicitly considers the spatial variation in species frequency and its effect on observation processes. The relative frequency of each species (state) 'near' an observation point is assumed to be related to the probability distribution of observed data at that location, while a nonparametric kernel regression estimator is used to predict it. We apply this model to simulated data as well as real intertidal sessile community data, showing that the new model provides corrected estimates of transition probabilities, even when the spatial dependence of community structure is anisotropic.

References

Conway-Cranos, L. L. and Doak, D. F. (2011) Sampling errors create bias in Markov models for community dynamics: the problem and a method for its solution. *Oecologia*, 167:199–207.
Fukaya, K. and Royle, J. A. (2013) Markov models for community dynamics allowing for observation error. *Ecology*, 94:2670–2677.

Development of a mixed hierarchical multinomial regression model to analyze community composition

E. Billoir [a], J. Arce Funck [a], C. Crenier [a], P. Usseglio-Polatera [a], J. Bouquerel [a], V. Felten [a]
and M. Danger [a]

[a] Interdisciplinary Laboratory for Continental Environments (LIEC)
University of Lorraine and CNRS
Campus Bridoux, Metz, France
elise.billoir@univ-lorraine.fr

Keywords: multinomial regression; random factors; nested factors; ecosystem functioning.

Abstract: The ecological objective of our study was to study the composition of benthic macro-invertebrates community in different streams in France, according to their level of nutrients and of metallic contamination. Four stream categories were to be compared, so-called *reference*, *metals*, *nutrients*, *metals+nutrients*. The 250 species retrieved in the field samples were split into seven functional groups, and we were interested in the potential effect of stream category on the proportions of each functional group within sampled field communities. Based on abundance data (counts) of 250 species in many river reaches of each stream category, we aimed at testing whether a category effect can be statistically highlighted, taking into account inter-reaches variability within stream categories. To this aim, we developed a model being (i) mixed, including one fixed factor “stream category” and one random factor “river reach”, (ii) hierarchical, since factor “river reach” is subordinate to factor “stream category” and (iii) with a multinomial error model (for the probability, for each sampled macro-invertebrate, to be in each functional groups). For the estimation of model parameters, we developed a Bayesian implementation using BUGS/JAGS. This provided us with credible intervals of the model coefficients, hence for each functional group we could assess the effect size and statistical significance of the two factors under study. To our knowledge, such a mixed hierarchical multinomial regression model is a new approach to analyze community data. It could be useful in many situations where a partition in categories is to be explained by nested factors.

Estimating temporal change in mean and variance of community composition via location scale additive models

Gavin L. Simpson

Institute of Environmental Change & Society
and Department of Biology
University of Regina
Regina, Canada
gavin.simpson@uregina.ca

Keywords: Generalised additive model; location and scale; variance; irregular time series; community dynamics; critical transitions; environmental change

Abstract: Many ecological and environmental time series are irregularly sampled in time. This is especially so for long data series (e.g. due to varying funding levels over time) and palaeoecological time series from sedimentary archives. In the case of palaeoecological data, the sampling is often truly irregular owing to unknown variation in sediment accumulation rates at the time of collection. Additive models (Hastie & Tibshirani, 1990) have been used successfully to model temporal trends in the mean of ecological and environmental time series (e.g. Ferguson *et al*, 2008; Curtis & Simpson, 2014). However, often other parameters related to the variance or the shape of the response distribution are of particular interest (e.g. increasing variance as an early warning indicator of critical transitions [Scheffer *et al*, 2015]).

Here, I describe the use of location-scale additive models (Rigby & Stasinopoulos, 2005; Wood *et al* 2015) to model temporal change in the mean and variance of lake sediment time series records of systems experiencing environmental change; specifically lakes affected by eutrophication. Using annually resolved sediment time series I compare results from the location-scale models with dynamic linear model and stochastic volatility model fits, confirming previous results showing increased ecosystem variance. I then proceed to fit location-scale additive models to irregularly sampled lake sediment records to investigate hypotheses of altered community dynamics following nutrient enrichment.

References

Curtis, C.J., & Simpson, G.L. (2014). Trends in bulk deposition of acidity in the UK, 1988–2007, assessed using additive models. *Ecological Indicators*, 37, Part B, 274–286.
Ferguson, C.A., Carvalho, L., Scott, E.M., Bowman, A.W., & Kirika, A. (2008). Assessing ecological responses to environmental change using statistical models. *Journal of Applied Ecology*, 45(1), 193–203.
Hastie, T.J. and Tibshirani, R.J. (1990) *Generalized Additive Models.* Chapman & Hall/CRC.
Rigby, R.A., & Stasinopoulos, D.M. (2005). Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society. Series C, Applied Statistics*, 54(3), 507–554.
Scheffer, M., Carpenter, S.R., Dakos, V., & van Nes, E.H. (2015). Generic Indicators of Ecological Resilience: Inferring the Chance of a Critical Transition. *Annual Review of Ecology, Evolution, and Systematics*, 46(1), 145–167.
Wood, S.N., Pya, N., & Säfken, B. (2015). Smoothing parameter and model selection for general smooth models. *arXiv [stat.ME]* Retrieved from `http://arxiv.org/abs/1511.03864`.

Joint Modeling of Species Interaction dynamics for Multi-state Processes with Imperfect Detection

J.E. Jones and A.J. Kroll

Timberlands Technology Department
Weyerhaeuser Company
PO Box 9777
Federal Way, WA 98063-9777

Keywords: co-occurrence, multi-state, ecological interactions, community dynamics

Abstract: Modeling dynamics of species interactions is a critical component of community ecology. Although species co-occurrence models have long been available, recent methodological advances support extensions to incorporate temporal population dynamics and variation in the detection process. However, current dynamic co-occurrence models have focused on marginally two-state systems, e.g., presence or absence for a given species, despite scientific evidence that the influence of species interactions may extend beyond species distributions to behaviors such as mating success or reproduction. We propose a joint model of species interaction dynamics that incorporates multi-state processes. In our model, one or more states of a subordinate species are assumed to depend on the state of a dominant species, but a reciprocal dependency does not exist. The basic framework we outline can be extended to multiple species, can include interaction lag effects, and deals efficiently with cases where data are missing for one species but not the other. We apply the model to data from a long-term Spotted owl monitoring program in a region with increasing Barred owl densities.

Hidden Semi-Markov Model for Female California Sea Lion Attendance

Jeff Laake, Devin Johnson*, and Sharon Melin

Alaska Fisheries Science Center,
National Marine Fisheries Service, NOAA
7600 Sand Point Way NE, Seattle, WA 98115 USA
jeff.laake@noaa.gov

Keywords: Attendance, Behavior data, California sea lion, Detection, Marked animals, Semi-Hidden Markov Model

Abstract: Hidden Markov models can be used to describe multi-state animal encounter history data with incomplete observations. However, the time spent in each state is geometrically distributed which is not appropriate for some types of data including attendance data for female California sea lions. Female sea lions come to the rookery prior to giving birth, give birth to a pup and remain at the rookery for an extended period of time nursing and caring for the pup. Afterwards, the females alternate feeding trips at sea and periods on land nursing the pup. The attendance cycle (at-sea, on-land), in particular the time at sea, is an important determinant of the health and growth of the pup that may be affected by prey availability. Attendance can be monitored with radio or satellite transmitters but this entails a cost and may limit the sample size. Alternatively, the rookery can be observed daily and the presence of individually identifiable females can be recorded. Observational attendance data is incomplete because some females may be present at the rookery but not seen. Using a sample of marked females with pups observed at San Miguel Island, we demonstrate the use of a hidden semi-Markov model with 4 states (pre-birth, birth, at-sea and on land) to derive estimated distributions for the time at-sea and on-land. These types of models have considerable potential to help describe pinniped behavior from observational data.

Linking northern fur seal dive behavior to environmental variables in the eastern Bering Sea

Ruth Joy*, Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, British Columbia V5A1S6 Canada
rutherfordjoy@gmail.com

Mike G. Dowd, Department of Mathematics and Statistics, Dalhousie University, Halifax, Nova Scotia B3H4R2 Canada

Brian C. Battaile, Marine Mammal Research Unit, University of British Columbia, Vancouver, British Columbia V6T1Z4 Canada

Pamela M. Lestenkof, Marine Mammal Research Unit, University of British Columbia, Vancouver, British Columbia V6T1Z4 Canada

Jeremy T. Sterling, National Marine Mammal Laboratory, Alaska Fisheries Science Center, Seattle, Washington 98115–6349 USA

Andrew W. Trites, Marine Mammal Research Unit, University of British Columbia, Vancouver, British Columbia V6T1Z4 Canada

Richard D. Routledge, Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, British Columbia V5A1S6 Canada

Abstract:

The population of northern fur seals *(Callorhinus ursinus)* in the Pribilof Islands, Alaska has declined dramatically during the past 35 years. Understanding fur seal's foraging behaviour at sea may provide important insights into what factors are associated with foraging success, particularly if these factors affect adult females whose success is also linked to pup survival. We propose an augmented state space methodology for studying behavioural patterns using high-resolution movement time series. We show how non-stationary time series models that describe systems for which parameters evolve slowly over time relative to the state dynamics can be estimated at relevant time scales for behavioural inference. This framework allows us to relate the time-varying parameter estimates of an auto-regressive system model to the seal's at-sea behaviour. The at-sea behaviour states of eleven lactating female northern fur seals were then matched, spatially and temporally, to a set of environmental variables, some of which were averages that represented the oceanic conditions over a large spatial area. The mismatch of scale between seal behaviour and the spatial variables was accounted for by applying an error-in-covariate Bayesian hierarchical model. Using this approach, we were able to link together northern fur seals that went to disparate regions of the eastern Bering Sea, with widely variable information about their underlying environmental fields into a single model. This application of a hierarchical model relates changes in identifiable behavioural states of the northern fur seal to changes in the Alaska commercial groundfish industry over a diurnal foraging cycle.

Statistical modeling and testing of the usage of oceanographic features by a top marine predator species

N. A. Pelland [a,b,*] and J. T. Sterling [a]

[a] National Marine Mammal Laboratory
NOAA Alaska Fisheries Science Center
Seattle, WA, USA
*nap2@uw.edu

[b] Joint Institute for the Study of the Atmosphere and Ocean
University of Washington
Seattle, WA, USA

Keywords: fronts, mesoscale eddies, physical oceanography, top marine predator, northern fur seal, *Callorhinus ursinus*

Abstract: Research on the ecology of top marine predator species often contains anecdotal examples of association between predators and discrete, small-scale oceanographic and bathymetric features such as surface fronts, mesoscale eddies, and canyons. In recent decades, advances in satellite tagging – along with the observations of the oceanic environment – present an opportunity for rigorous statistical testing of these associations and thus an increased understanding of how small-scale physical features influence the foraging environment for top predators. The process of doing so is a challenge that lies at the intersection of marine ecology, physical oceanography, and applied statistics. Here, we describe, from a physical oceanographic and ecological perspective, recent findings related to this research problem for the northern fur seal (NFS; *Callorhinus ursinus*), a wide-ranging pinniped species inhabiting the subarctic North Pacific Ocean. Large-scale satellite tagging programs initiated by NOAA's Marine Mammal Laboratory confirm historical interpretations that during their winter migration and summer foraging, NFS congregate in regions where eddy generation and surface fronts are common. During winter, statistical evidence supports the idea that NFS utilize the edges of mesoscale eddies during area-restricted horizontal movement that may be indicative of foraging, though there is substantial variability within and between individual animals and eddies. Individuals appear to obviously utilize some oceanographic features, but ignore others with few discernible differences – implying covariates or processes that are poorly understood or hidden from the current observing technologies. The method used for identification of surface features from oceanographic data is an important choice that may require tuning for specific situations. Identification of fronts through the use of Lagrangian Coherent Structures is one common approach, though has some theoretical and practical caveats, in particular when applied to near-coastal data. We discuss potential mechanisms that may account for predator usage of fronts and mesoscale eddies, and how these could inform future hypothesis formulation and testing.

Additional References

Pelland NA, Sterling JT, Lea M-A, Bond NA, Ream RR, Lee CM, et al. (2014) Fortuitous Encounters between Seagliders and Adult Female Northern Fur Seals (*Callorhinus ursinus*) off the Washington (USA) Coast: Upper Ocean Variability and Links to Top Predator Behavior. PLoS ONE 9(8): e101268. doi:10.1371/journal.pone.0101268

Nordstrom CA, Battaile BC, Cotte C, Trites AW (2013) Foraging habitats of lactating northern fur seals are structured by thermocline depths and submesoscale fronts in the eastern Bering Sea. Deep-Sea Research II 88-89: 78-96. dx.doi.org/10.1016/j.dsr2.2012.07.010

Using agent-based models to predict behavioral and physiological responses of top predators to environmental change: a case study with Weddell seals (*Leptonychotes weddellii*)

R.S. Beltran[a,b], J.W. Testa[c], J.M. Burns[b]

[a] Department of Biology and Wildlife
University of Alaska Fairbanks
Fairbanks, Alaska, USA
roxanne.beltran@gmail.com

[b] Department of Biological Sciences
University of Alaska Anchorage
Anchorage, Alaska, USA

[c] National Marine Mammal Laboratory
NOAA National Marine Fisheries Service
Seattle, Washington, USA

Keywords: Bioenergetics, pinnipeds, agent-based model, climate change, fishing pressure

Abstract: One of the crucial scientific challenges of this century is characterizing the vulnerability of ecosystems to global change. Bioenergetic models can be used to estimate total energy requirements by extrapolating simple physiological calculations to population-level metrics; however, these models often fail to link energy deficiencies with reproductive consequences or consider behavioral plasticity, and thus cannot be used to predict population-level consequences. An alternative approach is to use agent-based models which permit unique individuals to interact with variable extrinsic conditions (e.g. weather, prey), link energy deficiencies to reproductive and survival consequences, and allow individuals to adapt their behaviors. Here, we present an agent-based, ecophysiological model that simulates the energy balance of adult, female Weddell seals (*Leptonychotes weddellii*). The inputs include physiological parameters and population-wide ranges for the duration and phenology of life history events. Energy intake depends on foraging effort and stochastic prey availability at each timestep, whereas energy expenditure is calculated from time- and behavior-specific demands. The simulated animals select their activities (forage, nurse pup, molt, rest) based on body condition and life history constraints (i.e. dependent pup). Following model development and validation with empirical data, we ran simulations and compared the responses of individuals under baseline conditions to scenarios with reduced prey availability. A 10% reduction in prey availability resulted in seals foraging more and resting less (from 52.2%±6.2% resting to 40.3±8.4% resting). At the end of the year-long simulations, animals in the baseline simulation were in significantly better condition than animals with reduced prey availability (T-test, $t_{28}=5.6$, $p<.0001$). Our model successfully explored decision-based energy allocation strategies that occur under energetic stressors and elucidated how extrinsic conditions may impact individual fitness. Predicting the behavioral and physiological responses of predators is valuable for the study of global change biology and can be used to inform management decisions in polar regions.

Estimating Seasonal Behavior States from Bio-logging Sensor Data

J.M. London [a], D.S. Johnson [a], P.B. Conn [a], B.T. McClintock [a], M.F. Cameron[a], and P.L. Boveng[a]

[a] NOAA Alaska Fisheries Science Center
Seattle, Washington, USA
josh.london@noaa.gov

Keywords: hidden semi-Markov model; bio-loggers; animal behaviour; bearded seal.

Abstract: The seasonal timing of key, annual life history events is an important component of many species' ecology. Seasonal periods important to marine mammals often do not align well with typical labels (i.e., spring, summer, winter, fall). The timing of key life history events is well documented only for species found in accessible rookeries or breeding areas. Our knowledge of seasonal timing for species widely dispersed in inaccessible or remote habitats is poor. Here, we employed data from bio-logging sensors and new statistical modeling to identify and estimate timing of seasonal states for 7 adult bearded seals in the Bering Sea. These seals are reliant on the seasonal sea ice for pupping, nursing, breeding and molting and these seasons can be characterized by more time spent hauled out on ice and by changes in dive behavior. We are especially interested in the pupping-breeding-molting season, but also used this approach to identify seasonal structure in the non-breeding period. Seasonal periods were treated as separate behavior states that correspond to a hidden Markov process. Hidden Markov models (HMM) are commonly used to estimate behavior states (e.g., foraging, resting, transit) from telemetry data. Typical HMMs, however, have no temporal memory of state assignments and would likely not capture seasonal level states. To address this, we relied on a hidden semi-Markov model and explored specification of the transition matrix for the states to mimic the sequential timing of seasons. Dive and haul-out behavior as well as movement characteristics from bio-loggers were used to estimate these states. The timing and extent of sea ice in the Bering Sea is predicted to change dramatically over the next 50 years and we anticipate bearded seals might adjust the timing of these life history events in response to those changes.

Leveraging constraints and biotelemetry data to estimate the location of harbor seal haul-out sites

Brian M. Brost
Department of Fish, Wildlife, and Conservation Biology, Colorado State University

Mevin B. Hooten
U.S. Geological Survey, Colorado Cooperative Fish and Wildlife Research Unit
Department of Fish, Wildlife, and Conservation Biology, Colorado State University
Department of Statistics, Colorado State University

Robert J. Small
Division of Wildlife Conservation, Alaska Department of Fish and Game

Abstract. The location of central places like dens, nests, or resting sites can be determined by sighting individuals during aerial or ground-based surveys; however, direct observation may only provide a snapshot of the animal's behavior if surveys are infrequent, and could be altogether impractical when surveys are encumbered by remote locations, rugged terrain, or otherwise difficult conditions. We address these issues using a model-based approach for locating central places from satellite telemetry data. Our framework consists of an observation model that accounts for large telemetry location error and a highly flexible mixture model specified using a Dirichlet process to identify the location of central places. We apply our framework to Argos satellite telemetry data on harbor seals (*Phoca vitulina*) near Kodiak Island, Alaska. Harbor seals exhibit fidelity to terrestrial haul-out sites, locations that may warrant special management or conservation status because they exhibit unique characteristics and are associated with sensitive life history events. We also quantify temporal patterns in haul-out use by incorporating ancillary behavioral data into the model, and synthesize results across individuals with a formal meta-analysis to obtain population-level inference concerning spatio-temporal haul-out use.

Modeling innovations for fisheries assessments and management: are there any?

James Ianelli

Resource Ecology and Fisheries Management
Alaska Fisheries Science Center
National Oceanic and Atmospheric Administration
7600 Sand Point Way NE, Seattle WA
Jim.Ianelli@noaa.gov

Keywords: Fisheries management; Statistical modeling, computational methods

Abstract: Modern fisheries modeling practices for management purposes require extensive specifications of statistical uncertainty, both structural and in estimation. Evaluating risk is an important part of providing management advice (on catch limits etc) and use of ensemble approaches including complex ecosystem models is growing. Such methods have drawn parallels to hurricane track predictions but in our resource setting we never know precisely the actual impact nor its intensity. However, observations on how managers react to uncertain outcomes and the relative impacts are rarely evaluated. We argue that one area where applied resource management scientists can benefit is through meta-analysis of the interaction of available data, how they are analyzed and ultimately principles that affect decisions. Domestic examples from Alaskan fisheries are contrasted with experience in international resource management arenas.

Estimating changes in species interactions and community stability in a kelp forest ecosystem with a multivariate time series model

M.D. Scheuerell [a], E.J. Ward [b], and S.L. Katz [c]

[a] Fish Ecology Division
Northwest Fisheries Science Center
National Marine Fisheries Service
Seattle, Washington USA
mark.scheuerell@noaa.gov

[b] Conservation Biology Division
Northwest Fisheries Science Center
National Marine Fisheries Service
Seattle, Washington USA
eric.ward@noaa.gov

[c] School of the Environment
Washington State University
Pullman, Washington USA
steve.katz@wsu.edu

Keywords: time series; hierarchical Bayesian; food web; species interactions; stability.

Abstract: Global-scale losses of apex predators have led to profound changes in community structure and ecosystem processes, but the effects of predator reintroductions on recipient ecosystems remain poorly understood. We examined a long-term data set from a kelp forest in the California Channel Islands to characterize how species interactions and stability change over time following the reintroduction of sea otters. To do so, we developed a hierarchical Bayesian version of a time-varying, vector autoregressive, state-space model (TVVARSS). The general approach is that the log-density of an organism at time t is a linear combination of 1) its log-density at the previous time step t-1; 2) the sum of all pairwise interactions with predators, prey, and competitors at the previous time step t-1; 3) the effect of the environment; and 4) random environmental variation not explained by the covariates or model. This random variation may include catastrophic events, effects of unsampled species, changes in population age structure, or changes in habitat or prey availability. Our analysis identified multiple interactions within the food web that changed coincident with the establishment of otters, including interactions consistent with the classic trophic cascade (effects of predators on grazers, and grazers on primary producers subsequent to otter establishment). By allowing species interactions to vary over time, we also found that stability decreased over time, which may have been initiated by fishing effects, but was ultimately sustained via increased otter numbers. Therefore, explicit consideration of the potential changes in the interactions among species following predator reintroduction, and their integrated effects on stability, may be more important than simply quantifying species abundance or diversity.

References

Ives, A.R. and Dakos, V. (2012) Detecting dynamical changes in nonlinear time series using locally linear state-space models. *Ecosphere* 3: art58.

Ives, A.R., Dennis, B., Cottingham, K.L. and Carpenter, S.R. (2003) Estimating community stability and ecological interactions from time-series data. *Ecological Monographs*. 73: 301-330.

Size structured multispecies models to evaluate indirect effects of fishing

Nis S Jacobsen[a*] and Philipp Neubauer[b]

[a] School of Aquatic and Fisheries Science,
University of Washington
Seattle, WA, USA
nisjac@uw.edu

[b] Dragonfly Data Science
Wellington, New Zealand
neubauer.phil@gmail.com

Keywords: Size spectrum, Ecosystem based fisheries management, model calibration.

Abstract: Size structured multispecies models to evaluate indirect effects of fishing. Models based on physiological structure of interacting fish populations, often called size-spectrum models are becoming increasingly popular for evaluating the impact of fisheries on a strategic time scale. Size spectrum models have evolved from models representing a community to trait-based models and into models of specific ecosystems. One immediate challenge is the application of statistical techniques to properly estimate parameters and account for uncertainty in predictions. Here I present how a general trait-based multispecies model can be calibrated to large marine ecosystems using few parameters, and how these models can be used to predict indirect effects of fishing by accounting for trophic cascades and changes in food availability. I present calibrations of five large marine ecosystems, i.e. The North Sea, the Baltic Sea, The North East US Continental Shelf and the Benguela Current, all representing data rich ecosystems. I outline future venues for how the models can be used in ecosystems ranging from data-poor to data rich systems incorporating uncertainty and parameter estimation, and the possible role of size spectrum models in evaluating the impact of fisheries on ecosystems.

Visualizing Population Genetics

L. F. McMillan and **R. M. Fewster**

Department of Statistics
University of Auckland
Auckland, New Zealand
louise.mcmillan@gmail.com

Keywords: genetic assignment; population structure; multilocus genotype data; saddlepoint approximation

Abstract: Genetic assignment is the process of assigning individuals to source populations based on their genetic profiles. We propose a method for visualizing genetic assignment data by characterizing the genetic distribution of each candidate source population. This method improves upon the assignment method of Rannala and Mountain (1997) by calculating appropriate graph positions for individuals for which some genetic data are missing. Individuals with missing data are plotted at their population quantiles which we obtain using a saddlepoint approximation. The saddlepoint method also provides a way to visualize results from leave-one-out procedures.
We demonstrate our method using simulated data and microsatellite data from ship rats (*Rattus rattus*) captured on Great Barrier Island, New Zealand, and smaller surrounding islands. The visualization method makes it much easier to detect features of population structure and predict the accuracy of assignment results. It improves upon assignment software such as GeneClass, which has no visualization, and STRUCTURE, for which interpretation can be difficult when multiple population memberships are displayed for single individuals.

References

Rannala, B. and Mountain, J.L. (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences*, 94:9197-9201.

Density Approximation in Population Genetics

W. Zhang[a], **J. Liu**[b], **J. Goodman**[a] and **R.M. Fewster**[a]

[a]Department of Statistics
The University of Auckland
Auckland, New Zealand
wzha217@aucklanduni.ac.nz
jesse.goodman@auckland.ac.nz
r.fewster@auckland.ac.nz

[b]University of Michigan - Shanghai Jiao Tong University Joint Institute
Shanghai, China
stephen.liu@sjtu.edu.cn

Keywords: stationary distribution; linkage disequilibrium; maximum entropy principle; diffusion approximation

Abstract: The linkage disequilibrium coefficient r^2 is a measure of statistical dependence of the alleles possessed by an individual at two genetic loci. It is an important part of the multilocus structure of natural populations and has also been widely used in association studies to search for the locations of disease-causing genes on chromosomes. However, so far most emphasis has been put on seeking the quantitative properties of r^2 or other measures of linkage disequilibrium. In this work, we propose a method to approximate the population-wide probability distribution of r^2 at stationarity by the maximum entropy principle, given a finite sequence of the distribution's moments. For the model considered in this work, we generalize an elegant analytic method proposed by Song and Song (2007) to compute the stationary moments of r^2. Our method is based on the diffusion approximation, under which we demonstrate that for large numbers of models in population genetics, moments of the stationary distribution can be obtained without knowing the probability distribution itself. We illustrate our approach by recovering the stationary distribution of r^2 for a two-locus model incorporating mutation and recombination. Further, to show the robustness of the maximum entropy method, we apply it to simulated data to investigate the sampling distribution of r^2, which has been studied in the literature; however, compared with existing methods, our approach performs much better, especially when dealing with large samples.

References

Song, Y. S. and Song, J. S. (2007) Analytic computation of the expectation of the linkage disequilibrium coefficient r^2. *Theoretical Population Biology*, 71:49-60.

Using high-resolution mosquito surveillance data and a mechanistic metapopulation model to predict within-city spread of dengue fever

C.B. Leach[a,b], **C.T. Webb**[b], **A. Eiras**[c], **K. Pepin**[d], and **J.A. Hoeting**[a]

[a]Department of Statistics
Colorado State University
Fort Collins, Colorado
clint.leach@gmail.com

[b]Department of Biology
Colorado State University
Fort Collins, Colorado

[c]Department of Parasitology
Federal University of Minas Gerais
Belo Horizonte, Brazil

[d]National Wildlife Research Center
United States Department of Agriculture
Fort Collins, Colorado

Keywords: mechanistic modeling; metapopulation; epidemiology; stan;

Abstract: Dengue fever is a major public health risk in Brazil, with over a million human cases reported in 2015. The dengue virus is spread by the mosquito vector *Aedes aegypti*, and as such, prevention of dengue fever is accomplished primarily through municipal mosquito control programs. However, at the municipal scale, connecting patterns of mosquito abundance to human disease risk is complicated by human movement and heterogeneous mixing between human and vector populations in space. As a result, mosquito control policies would benefit greatly from the development of mechanistic models that can explicitly incorporate human-vector transmission processes and human movement. In this work, we adapt a single population model of dengue spread into a metapopulation context, where neighborhoods within a city are coupled by human movement. We combine this model with high-resolution mosquito surveillance data and a Bayesian hierarchical framework to estimate the patterns of human movement needed to predict observed weekly case reports from the city of Vitoria, Brazil. Implemented and fit using `stan` (Stan development team, 2015), the metapopulation model is able to successfully capture seasonal dengue dynamics at the city scale and performs better than models that assume city-wide homogeneous mixing between humans and mosquitoes and models that treat each neighborhood independently. Capturing these more detailed mixing patterns will help to better predict human disease risk from mosquito surveillance data and thereby better inform the allocation of mosquito control efforts within the city.

References

Stan Development Team. 2015. RStan: the R interface to Stan, Version 2.8.0 http://mc-stan.org/rstan.html.

Integrating spatial-capture recapture methods into ecological disease transmission models

Robin E. Russell[1], Tonie Rocke[1], Katherine L. D. Richgels[1,2], and Dan Walsh[1]
(1)National Wildlife Health Center, US Geological Survey, Madison, WI, (2)School of Veterinary Medicine, University of Wisconsin, Madison, WI

Keywords: disease ecology, spatial-capture recapture, plague, prairie dogs

Abstract Text:

Parameterizing models of wildlife disease transmission is a difficult process due to a variety of factors, including a lack of information on spatial structure of host organisms. Spatial capture recapture (SCR) techniques provide an estimate of average movement distances for animals from basic mark-recapture information. We apply SCR methodology to data collected from several prairie dog species, including Gunnison's (*Cynomys gunnisoni*), black-tailed (*C. ludovicianus*), white-tailed (*C. leucurus*), and Utah prairie dogs (*C. parvidens*), at 29 paired sites (58 prairie dog colonies) across the Intermountain West. We present the basic framework for our integration of field data into models of plague dynamics within prairie dog populations inhabiting prairie ecosystems.

Specifically, we use our estimated abundances and locations of activity centers estimated using the R package SCRbayes, which formulates SCR models in a Bayesian framework, as input into our mechanistic models of plague transmission. We demonstrate how the use of these integrated models of disease transmission can distinguish various patterns of disease processes such as epizootic or enzootic disease conditions based on empirical data, and can assist in designing effective disease management strategies.

References

Spatial-Capture Recapture, J.A. Royle, R.B. Chandler, R. Sollmann, and B. Gardner. 2014. Academic Press, Waltham, MA.

Modelling resource selection across multiple spatial scales using varying coefficient regression

T.Cornulier [a] and A. Villers [b]

[a] School of Biological Sciences
University of Aberdeen
Scotland
cornulier@abdn.ac.uk

[b] CEBC-CNRS
France
villers.alexandre@gmail.com

Keywords: resource selection; species distribution; generalized additive model; smoothing; spatial scale; scale-dependence.

Abstract: Species distribution models are widely used to predict the presence or abundance of animals in space in relation to combinations of environmental predictors – such as important habitat or food resources – perceived across a range of spatial scales. Identifying the combination of environmental variables and spatial scales that best predict animal distribution is difficult, due to often large numbers of plausible predictors and collinearity across spatial scales. Indeed, dependence on a particular resource may stretch out over a range of distances, its strength may vary with distance, and the strength-distance relationship is likely to be resource-specific, depending on the resource's biological function.

Here we propose a solution to these statistical problems grounded in ecological principles. The availability of resource r at location s is measured in D non-overlapping nested concentric rings (representing nested spatial scales) forming a S x D predictor matrix X_r. We use varying-coefficient regression to estimate the vector of coefficients for X_r, constraining effects to vary as a smooth function of spatial scale d. By applying appropriate parameter constraints, this approach resolves multiple scale and collinearity issues while providing ecologically meaningful inference about the scale-specific profile of effect for each resource.

We compare several methods of estimation and smoothing functions including splines and random walks. Using simulations we show the excellent performance of the method in terms of bias, precision and power even when test data are generated by a different model. We finally illustrate the application and the ecological insights gained from the approach by analysing real data on bird nest distribution (Montagu's harrier – *Circus pygargus*) in agricultural landscapes. The model has intuitive biological interpretation and is straightforward to implement with standard statistical software, making it accessible to most ecologists.

Integrating multiple data sources in species distribution modeling: A framework for data fusion

Krishna Pacifici[1], Brian J. Reich[2], David A.W. Miller[3], Beth Gardner[4], Glenn Stauffer[3], Susheela Singh[2], Alexa McKerrow[5], and Jaime Collazo[6]

[1]Department of Applied Ecology, North Carolina State University, Raleigh, NC 27695, USA
[2]Department of Statistics, North Carolina State University, Raleigh, NC 27695, USA
[3]Department of Ecosystem Science and Management, Pennsylvania State University, University Park, Pennsylvania, 16802, USA
[4]Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC 27695, USA
[5]U.S. Geological Survey, Core Science Systems, Biodiversity and Spatial Information Center, North Carolina State University, Raleigh, NC 27695, USA
[6]U. S. Geological Survey, North Carolina Cooperative Fish and Wildlife Research Unit and Department of Applied Ecology, North Carolina State University, Raleigh, NC, USA

Keywords: Brown-headed Nuthatch, data fusion, multivariate CAR, species distribution modeling

Abstract: The last decade has seen a dramatic increase in the use of species distribution models (SDMs) to characterize patterns of species' occurrence and abundance. Efforts to parameterize SDMs often create a tension between the quality and quantity of data available to fit models. Estimation methods that integrate both standardized and non-standardized data types offer a potential solution to the trade-off between data quality and quantity. We outline a continuum of strategies for integrating data, each of which treat data from standardized sampling protocols as a benchmark while integrating information available for non-standardized data types. We use three unifying principles to ensure robustness of our framework. First, estimation should permit sharing of information across space and among different sources of data. Second, strategies used to integrate across data sources should be motivated by an understanding of the ecological, sampling, and observational processes generating the different data sources. Finally, the general approaches should be flexible enough to accommodate different data structures and to account for sources of contamination and variability (e.g., false positives, false negatives, and covariates). We develop three general classes of models based on the Multivariate CAR model that permit varying levels of complexity in both the understanding of the ecological process and the degree of contamination of the different data sources. Through the use of simulations and a case study of the Brown-headed Nuthatch in the Southeastern U.S. we demonstrate a range of approaches, all of which improved predictions relative to a single data set. Methods that allow for both data types to be used will maximize the useful information available for estimating species distributions.

Dealing with many, possibly correlated, covariates in capture-recapture models

O. Gimenez [a]

[a] National Center for Scientific Research
Montpellier, France
olivier.gimenez@cefe.cnrs.fr

Keywords: capture-recapture models; dimension-reduction techniques; multivariate statistics.

Abstract: Capture-recapture models are widely used in statistical ecology for estimating demographic parameters. These models allow covariates to be incorporated to better understand population demography. The challenge is then to determine the relevant covariates that explain most of the variation in survival (Grosbois et al. 2008). When dealing with several covariates, one can include every one of them in a multiple regression modeling framework, therefore allowing for a possible joint effect of the variables on the parameters of interest. However, this approach is hampered by two main issues: first, because it increases the number of parameters to be estimated, incorporating many covariates in a model results in a loss of power in detecting a significant effect as well as a decrease in the precision associated with its magnitude; second, it has been long acknowledged that correlation among the set of covariates, usually referred to as multicollinearity, may alter reliability and interpretation of multiple regression analyses.

Here, we explore two dimension-reduction techniques to properly assess the effect of many and possibly correlated covariates in capture-recapture models. First, we implement a procedure based on forward selection and principal component analyses (PCA) to determine significant covariates (Aguilera et al. 2005). Second, we investigate a variant to PCA known as partial least squares (PLS) that has been specifically developed in a regression context (Bastien et al. 2005).

The performance of both approaches is analyzed by conducting simulation studies in which different scenarios are considered in terms of the number of covariates and the level of collinearity in covariates. We also illustrate these approaches with a real case study on the effect of climatic conditions on the survival of Snow petrels in Antarctica.

With the improvement of technology, the amount of collected information keeps growing, and consequently more and more complex models can be supported by the data. To overcome both issues of high-dimensionality and multicollinearity, we recommend using PCA or PLS in conjunction with capture-recapture models. Our methods requires standard statistical tools, which permits an efficient and easy implementation using standard software.

References

Aguilera, A.M., Escabias, M., Valderrama, M.J. (2006). Using principal components for estimating logistic regression with high-dimensional multicollinear data. *Computational Statistics and Data Analysis*, 50:1905-1924.

Bastien, P., Esposito Vinzi, V., Tenenhaus, M. (2005). PLS generalised linear regression. *Computational Statistics and Data Analysis*, 48:17-46.

Grosbois V., Gimenez O., Gaillard, J.-M., Pradel, R., Barbraud, C., Clobert, J., Møller, A. P. and H. Weimerskirch (2008). Assessing the impact of climate variation on survival in vertebrate populations. *Biological Reviews*, 83:357-399.

Spatial capture-recapture from the air with high-definition video

D. Borchers [a], P. Nightingale [b], R. Fewster [c] and S. Menze [a]

[a] Centre for Research into Ecological and Environmental Modelling
University of St Andrews, The Observatory, Buchanan Gardens,
Fife, St Andrews, KY16 9LZ, Scotland
dlb@st-andrews.ac.uk, mail@sebastian-menze.com

[b] School of Computer Science,
University of St Andrews, Jack Cole Building, North Haugh,
St Andrews, Fife, KY16 9SX, Scotland
pwn1@st-andrews.ac.uk

[c] Department of Statistics, University of Auckland
Private Bag 92019, Auckland, New Zealand
r.fewster@auckland.ac.nz

Keywords: aerial survey; double-observer; availability bias; animal movement.

Abstract: Aerial surveys using high-definition cameras instead of human observers are seeing growing use. While camera surveys have some clear advantages over human surveys, they also present some analytic challenges. Here we develop a model and estimator for aerial surveys of marine mammals using high-definition video. The survey involves two passes over transect strips. Individuals are not identifiable from the video. We formulate the survey as a spatial capture-recapture area search model with stochastic availability and movement, and unknown recaptures. Animals' dive cycles are modelled using a Markov modulated Poisson process and their movement is modelled by a diffusion process. Estimation is by maximisation of the marginal likelihood, marginalising over all possible capture histories, using times and locations of pairs of detections made on different passes to quantify the probability that the same animal could generate the pair. Marginalising is computationally challenging and constraint programming is used to improve efficiency. We consider use of the method to monitor harbour porpoise abundance in and around offshore windfarms. The method is shown to work well when density is low, and we investigate the effect of time lag between the two passes (a key design parameter) on estimator performance.

Modelling temporal and spatial variability in tag reporting rates for Newfoundland cod (*Gadus morhua*)

C. Konrad[a1], **J. Brattey**[b] and **N. Cadigan**[a2]

[a]Centre for Ecosystem Fisheries Research
Fisheries and Marine Institute
Memorial University of NL
St John's, NL, Canada
[a1]christoph.konrad@mi.mun.ca
[a2]noel.cadigan@mi.mun.ca

[b]Science Branch
Department of Fisheries and Oceans
St John's, NL, Canada
john.brattey@dfo-mpo.gc.ca

Keywords: Mark-Recapture; Tagging; Tag reporting rate; Newfoundland Cod; Template Model Builder; TMB; Random Walk

Abstract: Mark-Recapture experiments can be used to estimate the exploitation rate of a fishery; however, the estimate is influenced by the tag reporting rate by the fishers. We present two methods to estimate the reporting rates in high/low reward ($100 and $10 CAD respectively) long-term cod tagging experiments. We fit three Binomial logistic mixed effect models, two with temporal auto-correlation in the reporting rate year effects and one with independent year effects. The two auto-correlated models vary in their complexity: one estimates a constant fixed effect (i.e. intercept), the other one estimates a linear regression (i.e. intercept and slope) to which the random year effects are added. Due to the complexity of the fishery, our models account for factors such as fisheries-type, tag-type, tagging experiment, release length, recapture length, and recapture region. The commercial fishery showed spatial and temporal variation, with the reporting rate estimates lying between 0.60 and 0.78 for the independent year effect model. The temporal auto-correlation model had slightly higher estimates, ranging from 0.67 to 0.84. Furthermore, using simulations we assessed the performance of these models, paying particular attention to data requirements of the auto-correlated models. The coverage probability of nominal 95% confidence intervals was also investigate using simulations.

Trace-contrast models for capture-recapture without capture histories

R. M. Fewster[a], **B. C. Stevenson**[b], and **D. L. Borchers**[b]

[a]Department of Statistics, University of Auckland, New Zealand
r.fewster@auckland.ac.nz

[b]Centre for Research into Ecological and Environmental Modelling
University of St Andrews, Scotland
bcs5@st-andrews.ac.uk, dlb@st-andrews.ac.uk

Keywords: mark-recapture; natural tags; Neyman-Scott process; camera traps

Abstract: Capture-recapture studies increasingly rely upon natural tags that allow animals to be identified by intrinsic features such as coat markings, DNA profiles, acoustic profiles, or spatial locations. These innovations greatly broaden the scope of capture-recapture estimation and the number of capture samples achievable. However, they are invariably imperfect measures of identity, effectively sacrificing sample quality for quantity and accessibility. Drawing on ideas from a new Palm likelihood approach to parameter estimation in clustered point processes, we propose a new framework for drawing inference from capture-recapture studies based on comparing pairs of samples. Importantly, no reconstruction of capture histories is needed. We show that we can achieve accurate, precise, and computationally fast inference. We illustrate the methods with a camera-trap study of a partially-marked population of ship rats (*Rattus rattus*) in New Zealand.

Multivariate State Hidden Markov Models for Mark-Recapture Data

D. S. Johnson[a] and J. L. Laake

Alaska Fisheries Science Center, NOAA National Marine Fisheries Service,
7600 Sand Point Way NE, Seattle, WA 98115 USA
[a]devin.johnson@noaa.gov

Keywords: capture-recapture, Cormack-Jolly-Seber, Hidden Markov Model, multivariate, partial observation, state uncertainty

Abstract: State-based Cormack-Jolly-Seber (CJS) models have become an often used method for assessing states or conditions of free-ranging animals through time. Although originally envisioned to account for differences in survival and observation processes when animals are moving though various geographical strata, the model has evolved to estimate vital rates in different life-history or diseased states. We further extend this useful class of models to the case of multivariate state data. Researchers can record values of several different states of interest; e.g., geographic location and reproductive state. Traditionally, these would be aggregated into one state with a single probability of state uncertainty. However, by modeling states as a multivariate vector, one can account for partial knowledge of the vector as well as dependence between the state variables in a parsimonious way. A hidden Markov model (HMM) formulation allows straightforward maximum likelihood inference. The proposed HMM models are demonstrated with a case study using data from a California sea lion (*Zalophus californianus*) vital rates study. The marked sea lions were resighted by unique hot brand numbers, however, flipper tags were also applied to each side. The resight data were composed of three states (1) location (San Miguel I. or Año Nuevo I., California), (2) tag status for left flipper (present or absent) and (3) tag status for right flipper. The use of the brand mark for primary resighting allowed the rare opportunity to investigate dependence of pinniped flipper tag loss over the lifetime of the animal. The location portion of the state was always known upon resight, however, the status of either one or both tags was sometimes missed by the observer. We illustrate use of the multivariate state model to estimate age-specific survival, tag loss rates and dependence, and movement between islands.

Trace-contrast models to account for identification uncertainty in aerial surveys of cetacean populations

B. C. Stevenson[a], **D. L. Borchers**[a], and **R. M. Fewster**[b]

[a]Centre for Research into Ecological and Environmental Modelling
School of Mathematics and Statistics
University of St Andrews
United Kingdom
bcs5@st-andrews.ac.uk, dlb@st-andrews.ac.uk

[b]Department of Statistics
University of Auckland
New Zealand
r.fewster@auckland.ac.nz

Keywords: Mark-recapture, Palm intensity, point process

Abstract: Aerial surveys are an efficient and cheap means of monitoring cetacean populations. Line transects are flown by planes, from which detections of surfacing individuals can be made either by human observers or high-definition cameras. Absolute abundance cannot be estimated if only a single plane is used as there is no information available about the proportion of the population that evaded detection. Instead, two planes can be flown in tandem with some separation between them; this results in capture-recapture data, with each plane acting as a capture 'occasion'. However, individuals cannot be identified and so it is not always clear whether or not two sightings are detections of the same animal, and this precludes the use of traditional inference approaches. Here we show how detection data from wildlife surveys can often be considered as a clustered point pattern, whereby points corresponding to two detections of the same individual are typically nearby. This point pattern therefore provides noisy information about animal identities. In such cases estimation can be achieved using recently developed trace-contrast models. Here we extend these to account for a variety of features present in data collected on aerial surveys—in particular, partial individual identifiability (i.e., when some detection pairs are either known to be of the same animal, or known to be of different animals). An existing abundance estimation method for this kind of survey involves the enumeration of all plausible combinations of identity allocations and the calculation of their respective likelihood contributions; we show that the trace-contrast approach achieves inference at a far cheaper computational cost.

Distance sampling with camera traps

E.J. Howe [a], S.T. Buckland [a], M-L. Després-Einspenner [b], D.L. Borchers [a], and H.S. Kühl [b]

[a] Centre for Research into Ecological and Environmental Modelling
University of St Andrews
St Andrews, Scotland, UK
ejh20@st-andrews.ac.uk
steve@st-andrews.ac.uk
dlb@st-andrews.ac.uk

[b] Department of Primatology
Max Planck Institute for Evolutionary Anthropology
Leipzig, Germany
marie_despres@eva.mpg.de
kuehl@eva.mpg.de

Keywords: animal abundance; camera trapping; density; distance sampling.

Abstract: Reliable estimates of animal density and abundance are essential for effective wildlife conservation and management. Forests, particularly tropical forests, are rapidly dwindling, and the animals living there are subject to various anthropogenic impacts and stressors, so there is an urgent need to monitor the diverse assemblages of wildlife these habitats support. However, forests also pose challenges to researchers, notably limited visibility and accessibility, which prohibit aerial surveys and make data collection by human observers costly and time consuming. Remote, motion-sensitive photography ("camera trapping") has proven efficient for detecting multiple species of animals in such habitats, but statistical estimators of density from camera trapping data are still in development for species that cannot be individually identified. We present a point-transect distance sampling approach for estimating animal density from camera trapping data. Cameras are deployed at randomly- or systematically-selected points that are independent of animal density, and set to take photographs or record video for as long as animals are present to trigger them. From the images, we measure the radial distance between animals and cameras every *t* units of time while their midpoint is visible in the field of view. We regard the distance data as a series of snapshots, and select *t* to ensure that the distance an animal can move between successive snapshots is small relative to the range of the camera sensor. Sampling effort at a point is quantified in terms of the number of potential images obtained while the camera was operating, and the angle of the field of view of the camera; density estimation follows by standard point transect methods. This method allows researchers to exploit the various advantages and extensions of the distance sampling framework, including software. We tested it by simulation, and used it to estimate densities of Maxwell's duikers (*Philantomba maxwellii*) in Tai National Park, Côte d'Ivoire. Large samples of distance observations were obtained in a relatively short time period. Densities estimated from simulated data were unbiased, but bootstrap estimation of variances was required to achieve nominal confidence interval coverage. The method requires additional field-testing and validation, but we expect it to enhance the information obtained from camera trapping studies, and to be applicable to many large- and medium-sized animals in forest habitats.

Using visual survey data to estimate passive acoustic detection parameters for harbor porpoise abundance estimates

E.K. Jacobson [a], K.A. Forney [b], and J. Barlow [c]

[a] Scripps Institution of Oceanography,
University of California San Diego,
La Jolla, CA, USA
ekjacobs@ucsd.edu

[b] NOAA Southwest Fisheries Science Center,
Santa Cruz, CA, USA
karin.forney@noaa.gov

[c] NOAA Southwest Fisheries Science Center,
La Jolla, CA, USA
jay.barlow@noaa.gov

Keywords: density estimation, passive acoustic monitoring, harbor porpoise

Abstract: Passive acoustic monitoring is a promising approach for monitoring long-term trends in harbor porpoise *(Phocoena phocoena)* abundance along the U.S. West Coast. Before passive acoustic monitoring can be implemented to estimate absolute abundance, ancillary information is needed to convert recorded numbers of echolocation clicks to harbor porpoise densities. In particular, an echolocation click rate for individual animals and the effective radius monitored by each sensor are needed. In the present study, we used paired data from an array of 11 passive acoustic click detectors (C-PODs, Chelonia Ltd.) and three simultaneous aerial line-transect visual surveys at our 295 sq km study area in Monterey Bay, California to estimate the product of the echolocation rate and detection radius. During these aerial surveys, 245 groups of harbor porpoise were detected and we calculated an average density of 0.77 harbor porpoise per sq km. On the dates when aerial surveys were flown, the C-PODs detected an average of 112 harbor porpoise clicks per instrument-hour. We used an objective map of the aerial survey data to estimate the densities of harbor porpoises at each of the acoustic monitoring sites on each of the days of aerial survey effort. By substituting these visually estimated porpoise densities along with the observed number of echolocation clicks at each monitoring site into a cue-counting density estimation framework, we solved for the product of the two unknown passive acoustic detection parameters. We considered the visual density estimates and previously published harbor porpoise click rates and C-POD detection radii as informative priors in a Bayesian hierarchical model. This model-based approach resulted in a posterior distribution of the product of echolocation rate and detection radii comparable to previously published values. This technique may be a viable alternative for obtaining passive acoustic detection parameters when intensive experimental approaches are not feasible.

Spatial point process models for distance sampling surveys of animals that occur in groups

F. E. Bachl[a], **Y. Yuan**[b], **Janine B. Illian**[b], **Stephen T. Buckland**[b]
Håvard Rue[c], **Tim Gerrodette**[d], **D. L. Borchers**[b] and **F. Lindgren**[a]

[a]Department of Mathematics
University of Bath
Bath, United Kingdom
f.e.bachl@bath.ac.uk, f.lindgren@bath.ac.uk

[b]School of Mathematics and Statistics
University of St Andrews
St Andrews, Scotland
yy84@st-andrews.ac.uk, dlb@st-andrews.ac.uk

[c]Department of Mathematical Sciences
NTNU
Trondheim, Norway
Havard.Rue@math.ntnu.no

[c]NOAA National Marine Fisheries Service
Southwest Fisheries Science Center
3333 North Torrey Pines Court
La Jolla, California 92037, USA
tim.gerrodette@noaa.gov

Keywords: Distance sampling; Bayesian inference; Cox processes; stochastic partial differential equations

Abstract: Distance sampling is a widely used method of estimating wildlife population abundance. However, the underlying spatial structure of the surveyed population as well as the process governing group size distribution are often not taken into account by conventional design-based distance sampling methods. As a remedy we formulate distance sampling data as a thinned spatial point process and show how a log-Gaussian Cox process model allows us to simultaneously model animal abundance, spatial group formation properties and the effect of group size on detection probability. Our method uses a flexible stochastic partial differential equation (SPDE) approach to account for the spatial autocorrelation, and the Integrated Nested Laplace Approximation (INLA) approach for Bayesian inference. We illustrate the method using distance sampling data from a series of shipboard line transect surveys of dolphins in the eastern tropical Pacific (ETP).

Incorporating animal movement into distance sampling

R. Glennie, **S.T. Buckland** and **R. Langrock**

Centre for Research into Ecological and Environmental Modelling
University of St Andrews
St Andrews, Fife
United Kingdom
rg374@st-andrews.ac.uk

Keywords: abundance, distance sampling, animal movement, hidden Markov model

Abstract: Conventional distance sampling assumes that surveyed animals do not move whilst each transect is sampled (Buckland et al., 2001, p. 31). Movement can cause substantial bias in estimates of abundance (Glennie et al., 2015) and can prohibit the application of distance sampling altogether for fast-moving species. Double-observer methods have been used to eliminate bias when animals move in response to the observer (Buckland et al., 2010); however, no method exists to use distance sampling for animals that move independently. In this talk, we introduce a flexible continuous-time framework that incorporates movement modelling with distance sampling. We use auxiliary information, such as GPS tag data or expert elicitation, on target species movement together with the distance sampling survey data to estimate detectability and abundance. The continuous-time likelihood is approximated using the efficient hidden Markov model forward algorithm (Cappé et al., 2006); therefore, the method can account for all possible paths of each detected individual and so diminish the bias caused by movement. We illustrate the applicability of this model, discuss the computational difficulties involved and consider future development.

References

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., & Thomas, L. (2001). Introduction to distance sampling: estimating abundance of biological populations. Oxford University Press.

Buckland, S. T., Laake, J. L., & Borchers, D. L. (2010). Double-Observer Line Transect Methods: Levels of Independence. Biometrics, 66(1), 169-177.

Cappé, O., Moulines, E., & Rydén, T. (2006). Inference in hidden Markov models. Springer Science & Business Media.

Glennie, R., Buckland, S. T., & Thomas, L. (2015). The effect of animal movement on line transect estimates of abundance. PloS one, 10(3), e0121333.

Animal movement on point transect sampling abundance estimates

R. Prieto González[a], **T.A. Marques**[a] and **L. Thomas**[a]

[a]CREEM
University of St Andrews
St Andrews, UK
rpg2@st-andrews.ac.uk

Keywords: distance sampling; animal movement; abundance

Abstract: Distance sampling is a widely used method for estimating the abundance of wild animal populations. One key assumption of this method is that animals do not move while within detection range; in other words, the survey is a snapshot in time. Wildlife movement has long been recognized as an important problem for distance sampling methodology. We concentrate in point transect sampling, for which the observer is stationary and the problem more serious. Although distance sampling relies upon the idea that wildlife are observed at one instant of time, in practice detections occur over some interval. Bias arising for wildlife movement increases with time at the point.

Here, we present a novel method which incorporates animal movement into point transect estimates of abundance. We describe how animal movement leads to bias and show how that bias can be corrected for, leading to more robust density estimates.

Estimating abundance by integrating different sampling methods with varying spatial scales

N.J. Hostetter [a], B. Gardner [a], A.T. Gilbert[b], K.A. Williams[b], R.R. Veit[c]

[a] Dept. of Forestry and Environmental Resources
North Carolina State University
Raleigh, NC, USA
njhostet@ncsu.edu
bagardne@ncsu.edu

[b] Biodiversity Research Institute
Gorham, ME
andrew.gilbert@briloon.org
kate.williams@briloon.org

[c] Department of Biology
College of Staten Island/City University of New York,
New York, NY, USA
richard.veit@csi.cuny.edu

Keywords: distance sampling; abundance; seabirds; data integration

Abstract: Recent advances in integrated models (i.e., simultaneous analysis of multiple sources of information) have improved the estimation of demographic parameters. Here, we focus on a type of integrated model for data sources from different sampling methods with varying spatial scales. Motivated by a study on marine birds, where data were collected using boat-based distance sampling and high definition aerial videography, we develop a model to combine the data sources to predict species abundance across a large region. Abundances are allowed to correlate spatially through a multivariate CAR model, while incorporating spatially indexed environmental covariates. Additionally, imperfect detection is modeled using distance sampling for the boat-based surveys. To evaluate the utility in jointly modeling the two datasets, we compared this model with two other variations (1) a hierarchal distance sampling model with environmental covariates using the boat data only and (2) a model where aerial counts were smoothed across the study area and incorporated as an additional covariate in model. All models were evaluated and compared using 10-fold cross validation to determine bias and mean squared error. We applied the model to two species groups of marine birds in the mid-Atlantic Coast, USA – loons and alcids. These species were selected due to their different spatial patterns and habitat relationships. For loons, the integrated model identified abundance hotspots that were not predicted in boat only models and resulted in lower bias and mean squared error. Results differed for alcids, where correlations between survey types were lower and abundance estimates were not improved by the data integration model. The model developed here provides a flexible framework to analyze complementary data sources and can be extended to multiple data types with varying quality and spatial mismatch.

Nonparametric Bayesian model for functional equivalence of species

N.G. Johnson[a] and **J.C. von Fischer**[b] and **C.T. Webb**[c]

[a]National Institute for Mathematical and Biological Synthesis
Knoxville, TN, USA
nels.johnson@nimbios.org

[b]Department of Biology
Colorado State University
Fort Collins, CO, USA
joe.von_fischer@colostate.edu

[b]Department of Biology
Colorado State University
Fort Collins, CO, USA
colleen.webb@colostate.edu

Keywords: community data; nonparametric Bayes; functional equivalence; multiple species; model selection;

Abstract: Measurements of the composition of a community are frequently very sparse, with many species, for large numbers of samples. Particularly for measurements collected over a broad spatial range. Species interact with each other and the ecosystem itself. Understanding how species impact ecosystem functions is challenging when the community is sparse with many species. Reducing the number of species by combining species with the same impact on the ecosystem function into functionally equivalent groups would dramatically reduce computational needs and improve inference. We define criteria for what a useful model selection tool in this framework should look like and we propose a nonparametric Bayesian model which fits the criteria. We illustrate performance of the proposed model on several different sets of community data. We look at abundance, relative-abundance, and presence-absence community data for communities of ants, microbes, and plants, respectively.

Using copulas to model ecological communities

F. G. Blanchet [a] and B. M. Bolker [b]

[a] Department of Mathematics &Statistics
McMaster University
Hamilton, Ontario, Canada
guillaume.blanchet@math.mcmaster.ca

[b] Departments of Mathematics & Statistics and Biology
McMaster University
Hamilton, Ontario, Canada
bolker@mcmaster.ca

Keywords: Bayesian statistics, community ecology, copula

Abstract:

In this study, we propose to use Gaussian copulas to model ecological communities, i.e. the distribution of multiple species (with a combination of known and unknown characteristics) across multiple sites. Gaussian copulas provide a powerful alternative to current approaches such as numerical ecology (distance-based approaches, which depend sensitively on the distances chosen) or parametric joint species distribution models (which depend on the choices of conditional distributions of species). Copulas provide equal flexibility and interpretability to existing approaches (they have the same underlying structure as joint species distribution models), but copula users are freed from the need to specify distances or distributions *a priori*, making it easier to construct accurate and relevant species community models. Our approach also allows modellers to combine different response data types together in a single analysis (e.g. abundance data, percentage coverage, biomass), thus making it possible to approach new ecological questions.

We will illustrate our finding using simulated as well as real ecological data to show the potential of our approach.

Joint host–brood parasite occurrence models can improve predictions and reveal ecological traps.

Philip T. Patton[a], **Krishna Pacifici**[a], and **Jaime Collazo**[a,b]

[a]Department of Applied Ecology
North Carolina State University
Raleigh, North Carolina, USA
ptpatton@ncsu.edu
jkpacifi@ncsu.edu

[b]USGS North Carolina Cooperative Fish and Wildlife Research Unit
North Carolina State University
Raleigh, North Carolina, USA
jcollazo@ncsu.edu

Keywords: brood parasitism; hierarchical models; co–occurrence models; Shiny Cowbird; Puerto Rican Vireo

Forest edges can be an ecological trap for certain songbird species because they have higher rates of nest failure from brood parasitism than forest interiors. These forest edges dominate in forest/agricultural matrices, which are common throughout the globe. For example, the montane region of Puerto Rico—home to the invasive Shiny Cowbird and its preferred host, the endemic Puerto Rican Vireo—is an extensive coffee growing region pockmarked with forest patches. Although these regions, which often include agroforests or shade coffee, are considered a boon ecologically and economically, they could be poor habitat for hosts. Accurate predictions of where hosts occur, where parasites occur, and where the two co–occur could offer clues about the habitat quality for the host in these regions. If a brood parasite is more likely to occur where its host occurs, then joint occurrence models for the host and parasite may produce better predictions of parasite occurrence than standard parasite occupancy models. Joint models provide the added benefit of estimating host occurrence. These estimates can be used to produce improved predictions of co–occurrence. Using the example from Puerto Rico, we explore different ways of incorporating host information into joint occurrence models. These included the conditional occurrence model (Waddle et al. 2010), a two-stage hierarchical Bayesian model that propagates the joint uncertainty from the host and the parasite through to estimates of co-occurrence, and a commonly used model that includes the naïve occurrence of the host as a covariate. We assessed the predictive performance of each model using AUC, a pseudo–R^2 (coefficient of discrimination), and DIC. The two–stage model performed best by DIC and AUC. The presence of a vireo increased the probability of occurrence of the cowbird, which preferred coffee plantations over forests. Shade coffee plantations had high rates of vireo occurrence and cowbird-vireo co-occurrence, suggesting that coffee plantations could be an ecological trap for vireos. Using co-occurrence models improved predictions of cowbird occurrence and, therefore, provided finer detail about the distribution of brood parasitism. Cowbird parasitism of vireo nests, in this region, may be restricted to coffee plantations.

Waddle, J. H., R. M. Dorazio, S. C. Walls, K. R. Rice, J. Beauchamp, M. J. Schuman, and F. J. Mazzotti. 2010. A new parameterization for estimating co-occurrence of interacting species. Ecological Applications, **20**:1467-1475.

Estimating occupancy from variable effort surveys of species detections

David A.W. Miller

Department of Ecosystem Science and Management

Pennsylvania State University

University Park, PA 16802, USA

dxm84@psu.edu

Keywords: citizen science, detection, occupancy, species distribution modeling

Abstract - Estimating true occurrence probabilities (as opposed to relative occurrence) from detection non-detection data not collected under a standard survey design such as standard occupancy designs with repeat visits poses severe challenges. My goal is to describe a general approach for estimating true occurrence probabilities for species distribution models and other applications in the case where data consists only of whether or not the species is detected at a location and some measure of survey effort. This data structure is common for many citizen science programs, including eBird and state run breeding bird atlases, as well as other non-standardized survey efforts such as museum samples. The standard approach for dealing with these types of data are to include effort as an explanatory variable for the probability of observing a species alongside other covariates explaining occurrence in the same model to help control for unequal effort. However, true occurrence probability is not identifiable and as effort increases the model will generally predict the probability of occurrence will go to 1, even if the species does not occur in all locations. I show that in many cases true occurrence probability is identifiable and develop the maximum likelihood estimator for estimating occurrence when effort varies among locations. The approach works because as effort increases, the probability of detecting the species actually approaches the true occurrence probability and the probability of not detecting a species should generally follow an exponential decay function as effort increases. The full maximum likelihood estimator includes occurrence probability, detection probability, and effort in the formulation allowing a flexible framework to model heterogeneity for each. I will use simulations and analyses of Pennsylvania Breeding Bird Atlas and eBird data to show how the method can be applied and approaches to deal with heterogeneity and when effort itself is weakly measured. While data intensive, the approach show promise as a framework for accounting for effort in species distribution models fit to non-standard data.

Integrating data from multiple sources to improve species distribution models

D.L. Miller

Integrated Statistics
Woods Hole, MA
United States of America
dave@ninepointeightone.net

Keywords: integrating data; survey data; incidental data; species distribution modelling

Abstract: Increasing quantities of and access to both wildlife survey data and non-designed incidental or citizen science data have left us with a rather big problem: how to we put all of these disparate pieces together and build species distribution models that use as much of the available data as possible? This leads us to a series of sub-questions that I will address in this talk: should we combine data then model it all at once or, build multiple models and figure out how to combine their outputs (or couple their fitting)? How can we find equivalences in recorded effort (and what can we do when no effort is recorded)? I'll illustrate these issues and offer some solutions using example data from aerial and shipboard surveys of seabirds in New England, as well as from large-scale surveys of marine mammals in the North Atlantic.

The Zoön Project: Reproducible, Remixable and Shareable Species Distribution Modelling

N Golding.[a], T.A. August[b], T.C.D. Lucas[c], D.J. Gavaghan[d], N.J.B. Isaac[b],
R.B. O'Hara[e], E.E. van Loon[f] and G.J. McInerny[g]

[a] School of BioSciences, University of Melbourne, Australia. nick.golding.research@gmail.com
[b] Centre for Ecology & Hydrology, Wallingford, UK. tomaug@ceh.ac.uk; njbisaac@gmail.com
[c] CoMPLEX, University College London, UK. timcdlucas@gmail.com
[d] Department of Computer Science, University of Oxford, UK. david.gavaghan@dtc.ox.ac.uk
[e] Senckenberg Biodiversity and Climate Research Centre, Germany. rni.boh@gmail.com
[f] IBED, University of Amsterdam, The Netherlands. e.e.vanloon@uva.nl
[g] Centre for Interdisciplinary Methodologies, University of Warwick, UK.
gmcinerny@hotmail.com

Keywords: species distribution models; software.

Abstract: Species distribution modelling (SDM) is very widely used in ecology and SDM methods are the subject of thousands of papers every year. Despite the huge amount of effort invested in SDM research, the field is suffering a reproducibility crisis. Benchmark data and metrics to evaluate methods have not yet been developed, each new method is evaluated against a different dataset, and there are significant barriers to users implementing most new methods. This impedes the large-scale methodological comparisons which are crucial to advancing SDM research. The Zoön Project is developing software and an online platform to provide a solution to these problems.

Zoön encodes SDM analyses as a simple workflow composed of a series of modular processes. Each module is a small, self-contained snippet of R code to carry out one of five key SDM steps: obtaining occurrence or covariate data; pre-processing these data; fitting a statistical model; or generating interpretable outputs. This modular framework ensures that researchers can:rapidly create and share new modules, evaluate them against benchmark workflows, and easily slot new modules into an existing SDM analysis.

The first stage of the project developed an R package (cran.r-project.org/package=zoon) which provides a standard interface to SDM analyses and enables users to draw from an online repository of community-contributed SDM modules (github.com/zoonproject/modules). Modules have already been developed to obtain data from local files or online repositories (including GBIF and Bioclim), generate pseudo-absence data, carry out spatially-stratified cross-validation, fit common SDM models and over 200 machine learning algorithms, and generate interactive maps, model diagnostics and projections to novel climates.

Zoön workflows return a self-contained data object representing the entire analysis, which can be shared, explored, edited and re-run to enable fully reproducible SDM. The second stage of the project is underway; developing an online platform to share these workflow objects and provide a forum to discuss and evaluate new methods.

Modeling benthic fauna prevalence and abundance in a tropical tidal river

A. Lim [a], S. Chesoh [b] & D. McNeil [c]

[a] Department of Mathematics and Computer Science
Faculty of Science and Technology, Prince of Songkla University
Pattani 94000, Thailand
apiradee.s@psu.ac.th

[b] Department of Science
Faculty of Science and Technology, Prince of Songkla University
Pattani 94000, Thailand
chesoh.s@hotmail.com

[c] Department of Statistics
Faculty of Science, Macquarie University
Sydney 2109, Australia
don.mcneil@mq.edu.au

Keywords: prevalence; abundance; estuarine benthic fauna; factor analysis; regression model

Abstract: Benthic fauna play a fundamental role in sustaining aquatic ecosystem health by grazing organic detritus and thus providing nutrients for plankton and other organisms (Herman et al, 1999). The objective of this study was to illustrate an appropriate statistical method for analysing prevalence and abundance of communities of benthic fauna taxa in tidal rivers. Samples were collected from ten sites at bi-monthly intervals from 2005 to 2015 using Ekman 15x15 cm^2 420 micrometre mesh dredges, yielding 62 separate identifiable taxa (mostly genera), subsequently reduced to 45 by aggregating 18 infrequently occurring taxa. A factor analysis model was fitted by maximum likelihood to the correlation matrix of these data after adding 1 and log-transforming organism counts to remove skewness, yielding four interpretable community groups comprising, eight, seven, six and four different taxa, respectively, and 37 taxa (including the group of 18 rarer taxa) that did not fit this model. For these five groups, we then fitted additive three-factor (month, year and site) models to estimate prevalence (using logistic regression) and abundance (using linear regression on logarithms of populated samples). Q-q plots indicated plausibility of normality assumptions, and model results were displayed by graphing confidence intervals for comparing levels of each factor (adjusted for other factors) with overall means. Group 4 was found only in the five downstream sites, and overall prevalence in these groups decreased upstream, ranging from 20% for groups 1, 2 and 4, 40% for group 3 (found only at the five sites nearest the source) and 60% for group 5. There were also substantial variations in levels and patterns of prevalence and abundance between the five groups with respect to season, trend, and location upstream.

Reference

Herman, P.M.J. et al (1999) Ecology of estuarine macrobenthos. *Advances in Ecological Research*, 29:195-240.

From trajectories to animals' use of space maps: comparison of different approaches

M.P. Etienne and **P. Gloaguen** *

UMR AgroParisTech/INRA 518 Team MORSE
AgroParisTech
16 rue Claude Bernard
75 231 Paris Cedex 05, France
marie.etienne@agroparistech.fr
pierre.gloaguen@gmail.com

Keywords: Home-range; Utilization Distribution; Kernel Methods; Brownian Bridge; Ornstein Ulhenbeck process.

Abstract: A main topic in analysis of animals trajectories is the study of individual's home range. Home range of individuals are mainly computed by estimated Utilization Distributions (UD) of individuals. Two definitions of Utilization distribution are found in the literature. The first one will be named Global Utilization Distribution (GUD) in this work and consists in the time-averaged probability density of presence of the individual.* The second definition will be named Conditionnal Utilization Distribution (CUD) and proposes to consider the probability density of presence of the individual conditionally to its observed positions.

Two methods are linked to these two distributions to derive maps from observed trajectories. The GUD is mostly estimated thanks to non parametric kernel methods. In order to derive the GUD, the individual's movement has to exhibit some stationarity properties.

On the other hand, the CUD is classically estimated thanks to Brownian Bridges approaches and assumes that the Brownian motion is a good local approximation of the movement but does not require any stationarity assumptions.

In most of home-range studies, the formal definitions of these distributions are not clearly stated, and because they both produce maps, confusions might occur for the user regarding their differences.

We therefore propose to provide some mathematical formal definitions of these two fundamental concepts that are GUD ans CUD in order to highlight the fondamental conceptual differences between them.

Using the classical Ornstein-Uhlenbeck (OU) process as an example, we compare the two corresponding Utilization Distribution, estimated by classical methods, and illustrates their differences. The OU process is chosen because it verifies the assumption of stationarity (under some condition on its parameters), it is not far from a Brownian motion (at least around the attraction point) and some closed formula for GUD and CUD could be derived.

The differences between the two approaches on some more complex movement might be studied numerically (for example through Monte Carlo simulations) and should illustrate the importance of considering one or another according to the objective of the expected maps.

Diving effects on surface movement: a state-space model approach

M. Auger-Méthé [a], C.M. Albertsen [b], D.J. Yurkowski [c], K. Bøe [d], I.A. Fleming [e], and J. Mills Flemming [f]

[a,f] Department of Mathematics and Statistics
Dalhousie University
Halifax, Canada
[a] auger-methe@dal.ca
[f] Joanna.Flemming@dal.ca

[b] National Institute of Aquatic Resources
Technical University of Denmark
Charlottendlund, Denmark
cmoe@aqua.dtu.uk

[c] Great Lakes Institute for Environmental Research
University of Windsor
Windsor, Canada
dyurkow@uwindsor.ca

[d,e] Department of Ocean Sciences
Memorial University
St. John's, Canada
[d] kbe144@mun.ca
[e] ifleming@mun.ca

Keywords: Ocean Tracking Network; movement model; Template Model Builder.

Abstract: Diving activity can change the surface movement of marine animals and can be used to help characterize their behavioral state (Bestley et al. 2014). However, few 2D movement models incorporate diving information, and one of the few that does (Bestley et al. 2014), discretized movement paths into a limited number of behavioral states. Modeling discrete states assumes that animal behavior can be classifiable into a small number of categories, usually two (foraging and travelling), and that the characteristics of each behavior remain constant through time. Recent studies have shown that modeling changes in behavior through time-varying parameters is a powerful approach (e.g., Breed et al. 2012). Here, we describe a movement model that incorporates continuous behavioral changes and dive data. The model was develop for and applied to Argos movement data from ringed seals and light geolocation data from Atlantic salmons. We use a state-space model to account for the large measurement errors in these two datasets. Unlike most movement state-space models in the literature (e.g., Breed et al. 2012, Bestley et al. 2014), we use a likelihood-based non-Bayesian framework to estimate the parameters. This is possible through the use of the fast and efficient R package Template Model Builder (TMB). We discuss the advantage of our model and of using TMB when estimating the behavior of marine animals.

References

Bestley, S. et al. (2014) Taking animal tracking to new depths: synthesizing horizontal-vertical movement relationships for four marine predators. *Ecology*, 96:417-427.

Breed, G.A. et al. (2012) State-space methods for more completely capturing behavioral dynamics from animal tracks. *Ecological Modelling*, 235-236:49-58.

An individual-based model to examine mechanisms of long-distance movements of terrestrial mammals

C. Bracis [a] and T. Mueller [b]

[a] Senckenberg Biodiversity and Climate Research Centre (BiK-F)
Goethe Universität Frankfurt
Frankfurt (Main), Germany
cbracis@uw.edu

[b] Senckenberg Biodiversity and Climate Research Centre (BiK-F)
Goethe Universität Frankfurt
Frankfurt (Main), Germany
muellert@gmail.com

Keywords: cognition; individual-based model; movement; migration; memory; perception; ungulate.

Abstract: One of the key questions for addressing the underlying mechanisms of long-distance animal movement is how animals select where to go and, in particular, which cues they use. Most modelling efforts of the underlying drivers of mammalian land migration simply assume perception of favourable resources as the navigational mechanism. The possible role of memory that would allow mammals to forecast conditions at distant locations and times based on information about environmental conditions from previous years has generally not been considered. We compare these alternative navigation mechanisms using a spatially-explicit individual-based model that includes different possible information sources animals could utilize from their environment. In the model, animals integrate their expectation of habitat quality in different directions using a spatial kernel. They then probabilistically select a preferred direction that drives a continuous autocorrelated movement process. We compare movement decisions based on different movement models including a random null model, perception, and memory. Perceptually-guided individuals use currently sensed resource quality within their perceptual range, while memory-guided individuals use long-term averages of past resource quality to forecast future conditions. We apply this model to terrestrial mammals making long-distance movements, which are important to allow animals to access alternate locations such as feeding opportunities or breeding grounds and are crucial to the long-term survival of many species. Mechanisms underlying these movements are significantly less studied compared to bird migration, where a variety of navigational mechanisms, including perception and memory, are well established. Using remotely-sensed vegetation data, we perform simulations of foragers moving through landscapes of time-varying resource quality guided by perception or memory. We compare those simulations to GPS data on large-scale movements of grazing ungulates to reveal the underlying navigation mechanisms at work.

A multivariate mixed hidden Markov model to analyze blue whale diving behaviour during controlled sound exposures

Roland Langrock[a,b], **Stacy L. DeRuiter**[b,c], **Tomas Skirbutas**[b] , **Jeremy A. Goldbogen**[d] , **John Calambokidis**[e] , **Ari S. Friedlaender**[f,g] , **Brandon L. Southall**[g]

[a]Bielefeld University, Germany

[b]University of St Andrews, UK

[c]Calvin College, USA

[d]Stanford University, USA

[e]Cascadia Research Collective, USA

[f]Oregon State University, USA

[g]SEA, Inc., USA

Keywords: behavioural response; multivariate time series; numerical maximum likelihood; random effects; state-switching model

Abstract: We apply hidden Markov models (HMMs) to characterize multivariate time series data on the movement and diving behaviour of 37 blue whales, identifying latent states corresponding to three main underlying behaviour states: shallow feeding, travelling, and deep feeding. The model formulation accounts for inter-whale differences via a computationally efficient discrete random effect, and measures potential effects of experimental acoustic disturbance on between-state transition probabilities. We identify clear differences in blue whale disturbance response depending on the behavioural context during exposure. Findings are consistent with earlier studies using smaller samples, but the HMM approach provides a more nuanced characterization of behaviour changes.

General tools to model movement in highly fragmented patch networks

Juan Manuel Morales
Laboratorio ECOTONO
INIBIOMA
Universidad Nacional del Comahue
Quintral 1250, (8400) Bariloche, Argentina
jm.morales@conicet.gov.ar

Abstract:

For many theoretical and applied questions it is important to understand how animals move in a patch network. However, this is no simple task because the probability that an individual moves from one particular patch to another one is affected in a non-trivial way by the characteristics and location of other patches in the network.

Here we present some simple and flexible statistical movement models that take into account the spatial structure of a patch network. For this we follow Ovaskainen and Cornell (2003) derivation of the formulae for diffusion in highly fragmented landscapes but replace some of their analytical results with general "statistical" functions with parameters that need to be estimated from data, leaving diffusion as a special case.

We start by considering a matrix **H** holding the probabilities of eventually going to patch *j* before dying or emigrating from the patch network given that the animal has just left patch *i*. The elements of this matrix are a function of the distance between patches d_{ij} but can also depend on other attributes such as patch size, quality, etc. These H_{ij} values can be thought of as the probabilities that an individual starting at a distance d_{ij} from patch *j* will eventually reach it before dying considering that there are no other patches in the landscape.

From matrix **H** we want to obtain the probabilities P_{ij} of visiting *next* patch *j* given that the individual has just left patch i. If we assume that P_{ij} depends only on the animal just leaving patch *i* but not on the full history of previous movements, we can write H_{ij} as a combination of these P_{ij}: $H_{ij} = P_{ij} + \sum_{k\neq j} P_{ik}H_{kj}$. We then use a linear solver to obtain the P_{ij}s. This is the main "trick" that we rely on for the purposes of estimating the effect of the spatial structure of the patch network. The P_{ij} estimated in this way depend on movement properties and on the structure of the habitat network but are valid only for the case when animals forget where they were coming from once they arrive to a patch. However, after we have solved for P_{ij} we can modify these probabilities using weights that depend on the visitation history.

We present both a MCMC and ABC toolbox to estimate the relevant parameters for the case of data collected through tracking animals (MCMC) and mark-recapture protocols (ABC).

Reference:

Ovaskainen, O., and S. J. Cornell. 2003. Biased movement at a boundary and conditional occupancy times for diffusion processes. Journal of Applied Probability **40**:557-580.

Estimating negative binomial parameters from occurrence data with detection times

W.H. Hwang[a], **R. Huggins**[b] and **J. Stoklosa**[c]

[a] Institute of Statistics
National Chung Hsing University
Taichung, Taiwan
wenhan@nchu.edu.tw

[b]Department of Mathematics and Statistics
The University of Melbourne
Melbourne, Australia
[c] School of Mathematics and Statistics and Evolution & Ecology Research Centre
The University of New South Wales
Sydney, Australia

Keywords: aggregation index; cost analysis; mis-identification; presence-absence data.

Abstract: The negative binomial distribution is a common model for the analysis of count data in biology and ecology. In many applications, we may not observe the complete frequency count in a quadrat but only that a species occurred in the quadrat. If only occurrence data are available then the two parameters of the negative binomial parameters, the aggregation index and the mean, are not identifiable. This can be overcome by data augmentation or through modelling the dependence between quadrat occupancies. Here we propose to record the (first) detection time while collecting occurrence data in a quadrat. We show that under what we call proportionate sampling, where the time to survey a region is proportional to the area of the region, that both negative binomial parameters are estimable. When the mean parameter is larger than two, our proposed approach is more efficient than the data augmentation method developed by Solow and Smith (2010), and in general is cheaper to conduct. We also investigate the effect of mis-identification when collecting negative binomialy-distributed data, and conclude that, in general, the effect can be simply adjusted for provided that the mean and variance of mis-identification probabilities are known. The results are demonstrated in a simulation study and illustrated in several real examples.

References

Solow, A.R. and Smith, W.K. (2010). On predicting abundance from occupancy. *American Naturalist*, 176:96-98.

National biodiversity monitoring schemes: setting standards

S.T. Buckland

Centre for Research into Ecological and Environmental Modelling
University of St Andrews
St Andrews, United Kingdom
steve@st-andrews.ac.uk

Keywords: citizen science; Convention for Biological Diversity targets; temporal trends; turnover measures.

Abstract: In 2010 in Aichi, Japan, ‘20 targets for 2020’ (https://www.cbd.int/sp/targets/) were agreed under the Convention for Biological Diversity (CBD). These targets are considerably wider-ranging and more vague than the 2010 target, ‘to reduce the rate of loss of biodiversity’. Hence determination of whether targets are met is (even) more problematic. As with the 2010 target, little thought seems to have been given to what data are required if such determinations are to be made with any confidence. Some capability to monitor national trends in biodiversity would seem essential. We explore statistical issues for national monitoring schemes, and use UK schemes to illustrate good and bad practice. Issues include: survey design; spatial variation in temporal trends; quantifying effects of climate change; regression to the mean; bias from monitoring known colonies only; the effects of varying detectability; and ‘citizen science’ monitoring schemes.

Taxonomic diversity, functional redundancy and resilience of nocturnal lepidopteran communities in plantation woodlands.

Lucinda Kirkpatrick[1], Lauren Fuller[2], Kirsty Park[2]

[1]Presenting author:
Biological and Ecological Sciences,
University of Stirling
Scotland
Lucinda.kirkpatrick@stir.ac.uk

[2] Biological and Ecological Sciences,
University of Stirling
Scotland

Keywords: Taxonomic diversity, functional diversity, community composition, nocturnal lepidoptera, resilience, plantation

Abstract: Recently several studies have demonstrated the potential for measures of functional diversity to better explain ecosystem functioning, resistance to disturbance (functional redundancy) and resilience (response diversity) compared to traditional measures of taxonomic diversity. These tools may be invaluable for ecosystem managers requiring more mechanistic explanations of the complex ecosystem processes maintaining ecosystem states (Standish et al., 2014).

As plantation woodlands now represent a major world forest cover type, forest managers are increasingly looking to more sustainable management approaches. Understanding the impacts of disturbance on nocturnal lepidoptera provides an opportunity to investigate the impacts of forest management processes on an ecologically important and diverse order, and allows the development of management techniques to mitigate future functional diversity loss. We assessed 170 stands across three large plantation forests in the UK to determine current levels of moth biodiversity, how moth functional and taxonomic diversity varied with disturbance such as clearfelling, and assessed moth community resilience (recovery from disturbance; using effect traits) and resistance (ability to withstand disturbance; functional redundancy, using response traits) at multiple spatial scales (Laliberté & Legendre, 2010). We found significant differences between stand types, with ecosystem resilience impacted by presence of broadleaved woodland within the landscape. Maintaining native woodland within a plantation matrix should benefit moth community composition and related ecosystem service delivery.

Laliberté, E., & Legendre, P. (2010). A distance-based framework for measuring functional diversity from multiple traits. *Ecology*, *91*(1), 299–305. http://doi.org/10.1890/08-2244.1

Standish, R. J., Hobbs, R. J., Mayfield, M. M., Bestelmeyer, B. T., Suding, K. N., Battaglia, L. L., … Thomas, P. a. (2014). Resilience in ecology: Abstraction, distraction, or where the action is? *Biological Conservation*, *177*(SEPTEMBER), 43–51. http://doi.org/10.1016/j.biocon.2014.06.008

Studies on the molecular profiling of tasar silkworm, *Antheraea mylitta*. D

Shamitha Gangupanthulu*, G. Renuka, M. Sreenivas and G. Shiva Kumar
Department of Zoology
Kakatiya University, Warangal
Telangana, India-506009
*corresponding author : vodithalashamitha@gmail.com

Key words: *Antheraea mylitta*; molecular markers; *ex-situ*; diversity.

Abstract: The polyphagous, sericigenous lepidopteran tasar silkworm *Antheraea mylitta* Drury found in tropical forests mostly feeding on *Terminlia arjuna* has a wide range of distribution. In India, it covers more than twenty states as ecoraces, with variations in phenotypic traits like fecundity, voltinism, cocoon weight, silk ratio etc. These populations are very difficult to separate based on morphological traits. Molecular markers have been widely used in studies of the genetics of wild populations to understand genotypic diversity within and between populations and evolutionary processes. The molecular diversity was assessed among seven ecoraces of *A.mylitta* with SSR and ISSR primers on agarose gel. For the present study, PCR-SSR and PCR-ISSR based phylogenetic analysis using POPGENE 1.32 was carried out, which revealed the genetic closeness of the ecoraces based on geographic regions and altitude.

Quantifying temporal changes in ecological communities

H. Shimadzu

Centre for Biological Diversity and Scottish Oceans Institute
University of St Andrews
St Andrews, Scotland
hs50@st-andrews.ac.uk

Keywords: beta diversity; temporal turnover.

Abstract: Ecological communities vary through time (and space). Range migrations in response to climate change, invasive species and the emergence of novel ecosystems highlight the importance of temporal change in ecological communities as a fundamental part of global change in the Anthropocene. Although temporal changes are usually quantified using a variety of metrics initially developed to capture spatial differences, they are the consequence of unidirectional community dynamics resulting from processes such as population growth, colonisation and local extinction.

This talk introduces a statistical framework of quantifying temporal changes in ecological communities accounting for community dynamics, and proposes a new temporal turnover measure which encompasses those matrices widely used for both presence/absence and abundance data in the spatial context. A simulation study and an analysis of an estuarine fish community both clearly demonstrate that our proposed turnover measure offers additional insights relative to spatial context-based metrics. It reveals whether community turnover is due to shifts in community composition or in community abundance and identifies the species and/or environmental factors that are responsible for any change.

References

Shimadzu, H. Dornelas, M. and Magurran, A.E. (2015). Measuring temporal turnover in ecological communities. *Methods in Ecology and Evolution* 6(12): 1384-1394.

National biodiversity monitoring interactive reporting

Paul van Dam-Bates[a] **Meredith McKay**[b]

[a]Department of Conservation
Science Advisor - Statistician
Christchurch, New Zealand
pbates@doc.govt.nz

[b]Department of Conservation
Technical Advisor - National Biodiversity Monitoring & Reporting Programme
Christchurch, New Zealand
mmckay@doc.govt.nz

Keywords: Automated reporting; Biodiversity; Public data

Abstract: In 2011 The Department of Conservation (DOC) began a systematic sampling programme for all public conservation land (PCL) in New Zealand. The goals are to report on national and regional trends in biodiversity, inform on resource allocation within PCL, evaluate the effectiveness of management, and provide an early warning system for emerging threats. Roughly two hundred and eighty randomly selected locations on an 8-km grid are sampled in a 5-year cycle for a total of ≈1400 sites across the country. Surveys include vegetation (20×20 m plots), birds (occupancy, distance sampling, and acoustic arrays) and mammals (trapping and pellet counts). All of this data is publicly available, but due to the complexity and scope, the analysis and reporting has been limited leaving a lot of potential for future work. We have developed automated plot-level reports to allow users to explore the data at each individual site and put that into the context of national PCL biodiversity. R is used for data processing and modelling, and then knitr and Pandoc embed the R outputs into HTML pages running JavaScript. By having HTML as an output document we are able to create dynamic reports with interactive graphics at multiple spatial scales. They can then be hosted online for public access along with the actual data.

Flexible discrete space models for animal movement

E.M. Hanks[a]

[a] Department of Statistics
The Pennsylvania State University
University Park, PA, USA
hanks@psu.edu

Keywords: animal movement; stochastic processes; spatial ecology.

Abstract: Animal movement behavior is complex, and often exhibits changing behavior over time, response to local environmental conditions, long-term memory, and dependence between conspecifics. We propose a flexible discrete-space model for animal movement based on a stochastic continuous-time Markov chain process. Within this framework, both the absolute movement rate and directional bias in movement can be modeled using environmental and other covariates. We show how this framework can capture correlated random walk behavior, central-place foraging, seasonally-varying response to environmental gradients, and other behaviors. We also present an auxillary variable representation of this movement model that allows for model fitting using standard GLM software, and introduce the 'ctmcmove' R-package that facilitates fitting this model to observed telemetry data.

References

Hanks, E.M., Hooten, M.B., and Alldredge, M.W. (2015) Continuous time discrete space models for animal movement. *The Annals of Applied Statistics*, 9:145-165.

Inference in models of dynamic fitness optimization based on observed animal behavior

U.H. Thygesen[a], **T.A. Patterson**[b]

[a]Centre for Ocean Life
Technical University of Denmark
National Institute for Aquatic Ressources
Charlottenlund, Denmark
uht@aqua.dtu.dk

[b]CSIRO, Marine and Atmostpheric Research
Hobart, Tasmania, Australia
Toby.Patterson@csiro.au

Keywords: Behavioral ecology, Dynamic fitness optimization, Time series analysis, Maximum Likelihood.

Abstract: Mathematical models, in general, have highest credibility and fidelity when they build on established first principles in combination with available data. In behavioral ecology, such principles derive from evolutionary arguments, i.e. animal behavior should be fitness-optimizing. However, evolutionary principles rarely enter in a direct mechanistic fashion in statistical models, largely due to the lack of a suitable estimation framework. Here, we present a statistical framework for inference in dynamic models of fitness optimizing behavior, which allows maximum likelihood estimation of fitness components and trade-offs based on observation of animal behaviour - in this case collected by data storage tags. The central component in the framework is that behavior is assumed to be random, but such that the probability of a behavior increases with the fitness resulting from that behavior. We demonstrate different methods for inference in such models, for situations where dynamic states and decision variables are observed or partly observed. We illustrate the framework and the methods with a simulation example involving patch selection, as well as with real data involving vertical strategies of migrating eel and foraging tuna.

Quantifying uncertainty in spatio-temporal forest composition changes inferred from fossil pollen records

A. Dawson[a,b] and **C.J. Paciorek**[a] and **J.S. McLachlan**[c]
and **S. Goring**[d] and **J.W. Williams**[d] and **S.T. Jackson**[b,e]

[a]Department of Statistics
University of California, Berkeley
Berkeley, California
andria.dawson@gmail.com

[b]Department of Geosciences
The University of Arizona
Tucson, Arizona
[c]Department of Biological Sciences
University of Notre Dame
South Bend, Indiana
[d]Department of Geography
University of Wisconsin - Madison
Madison, Wisconsin
[e]Department of the Interior Southwest Climate Science Center
U.S. Geological Survey
Tucson, Arizona

Keywords: spatio-temporal, predictive process, Bayesian, HMC, pollen, forest, paleoecology

Abstract: Understanding past compositional changes in vegetation provides insight about ecosystem dynamics in response to changing environments. Past vegetation reconstructions rely predominantly on fossil pollen data from sedimentary lake cores, which acts as a proxy record for the surrounding vegetation. Stratigraphic changes in these pollen records allow us to infer changes in composition and species distributions. Pollen records collected from a network of sites allow us to make inference about the spatio-temporal changes in vegetation over thousands of years. To reconstruct vegetation composition, we build a Bayesian hierarchical model for the Upper Midwestern USA that links vegetation composition to fossil pollen data via a dispersal model. First, we estimate the relationship between vegetation and pollen for the settlement era using United States Public Land Survey data and a network of pollen records. Parameter estimates from the settlement era and the fossil pollen proxy records can then be used to estimate the latent species distributions and relative abundances over the last 2500 years. To avoid estimating the latent vegetation process for all spatial locations in the domain, we use a predictive process framework which requires that this process be estimated only for a selection of knots. In both the calibration and prediction phases of the model, we use an efficient implementation of the No-U-Turn Sampler (NUTS), a variant of the Hamiltonian Monte Carlo gradient-based algorithm. This work highlights the importance of efficient statistical and computational methods in cases such as this where the model is sufficiently complex and the domain of inference is large.

References

Paciorek, C.J. and McLachlan, J.S. (2008) Mapping ancient forests: Bayesian inference for spatio-temporal trends in forest composition. *JASA*, 104:608-622.

A Spatio-Temporal Model to Infer Colonization Dynamics of Sea Otters in Glacier Bay, Alaska

P.J. Williams[a], **M.B. Hooten**[b], **J.N. Womble**[c], **G.G. Esslinger**[d], and **M.R. Bower**[c]

[a]Colorado Cooperative Fish and Wildlife Research Unit
Colorado State University
Fort Collins, Colorado
perry.williams@colostate.edu

[b]U.S. Geological Survey
Colorado Cooperative Fish and Wildlife Research Unit
Colorado State University
Fort Collins, Colorado
mevin.hooten@colostate.edu

[c]National Park Service
Southeast Alaska Inventory and Monitoring Network
Juneau, Alaska
jamie_womble@nps.gov
michael_bower@nps.gov

[d]U.S. Geological Survey
Alaska Science Center
Anchorage, Alaska
gesslinger@usgs.gov

Keywords: diffusion, optimal monitoring, population trends

Abstract: Methods for inferring the dynamics of population spread through time are important for many ecological applications including the re-introduction of extirpated species and invasive species management. Spatio-temporal statistical models for describing and predicting spatially-explicit processes that evolve over time are being increasingly used by ecologists and provide a framework for understanding the dynamics of population spread. Sea otters (*Enhydra lutris*) were introduced to the outer coast of southeastern Alaska in 1965 and first detected in Glacier Bay around 1993. Since 1993, sea otters have increased in both abundance and distribution in Glacier Bay; preliminary abundance estimates have increased from 5 otters to >8,000 otters. We developed a Bayesian hierarchical spatio-temporal model and fit it to aerial survey data to better understand the ecological processes governing changes in sea otter abundance and distribution. Understanding the processes governing the population expansion provides a template for developing an optimal, dynamic monitoring framework that can be used to both monitor abundance and distribution of a spreading population, and also to reduce uncertainty in our understanding of the ecological processes associated with population spread.

Exploring the genetic consequences of dispersal patterns and varying life history parameters; a simulation based on North Sea and Skagerrak cod

Ingrid Spies*, Lorenz Hauser, Per Erik Jorde, Halvor Knutsen, André E. Punt, Lauren Rogers, Nils Chr. Stenseth

Abstract

Genetic data are commonly used to draw conclusions about dispersal. However, even significant genetic differentiation cannot be directly translated into a dispersal rate. Interpretation of genetic data can be even more problematic when genetic differentiation is not significant, because it is difficult to distinguish between moderate dispersal rates and complete panmixia (Waples 1998). A model incorporating genetics and population dynamics was used to estimate dispersal between the North Sea Atlantic cod (*Gadus morhua*) and the Norwegian Skagerrak. A fjord and the adjacent coastal population were examined, which are connected to the North Sea by an unknown amount of dispersal. Results suggest that significant genetic differentiation observed between the North Sea and the fjord population (Søndeled) was most consistent with 54-80 age-0 immigrants into the fjord each year from the North Sea. Observed genetic differentiation between the North Sea and the coast adjacent to the Søndeled fjord (Risør) was small and not significant, and dispersal estimates were fairly high (1,300-2,800) indicating that migration from the North Sea impacts the outer coast demographically. The estimated number of migrants resulting in empirical levels of genetic differentiation was sensitive to maturity-at-age and Skagerrak population size, depending on the level of genetic differentiation. Estimates of migration were also sensitive to the extent of fishing mortality and fishery selectivity. Dispersal is likely to be affected by some level of stochasticity; therefore, we applied several different patterns of migration to examine its effect on genetic differentiation.

Spatio-temporal modelling using geostatistics as a tool to manage Portuguese bivalve small-scale fisheries

Marta M. Rufino[a,b], Fábio Pereira[a], Paula Moura[a], Paulo Vasconcelos[a] & Miguel B. Gaspar[a,b]

[1] *Instituto Português do Mar e da Atmosfera (IPMA, I.P), Rua Alfredo Magalhães Ramalho Nº 6, 1495-006 Algés, Lisboa, Portugal.* marta.m.rufino@gmail.com

[2] *Centro de Ciências do Mar (CCMAR), Universidade do Algarve, Campus de Gambelas, 8005-139 Faro, Portugal*

Keywords: spatial ecology; fisheries; abundance; species distribution models

Abstract:

Most marine species are aggregated into spatial patches. Those patches vary through time, both in biomass and location, as a response to a set of drivers. Thus, to understand the ecological, environmental and anthropological drivers of species biomass, long-term spatial analyses of species dynamics are essential. In particular, target species of small-scale fisheries are further influenced by governance and market drivers, with a significant impact on their dynamics, which in turn have considerable economic and social impacts. In the coast of the Algarve (southern Portugal) an important small-scale fishery targeting four commercial bivalve species takes place on the sandy bottoms. In the current work, almost three decades (1986-2015) of annual surveys of the striped venus clam (*Chamelea gallina*) were analysed using spatial-temporal geostatistical methods. We suggest that such historical survey data (commonly produced by fisheries institutes) may be modelled in function of space and time using a spatial-temporal variogram. The model is then used to produce predictions over the study area using spatial-temporal kriging. This approach produces regular interpolated maps of species biomass, overcoming issues of different samples taken every year and providing an attractive analysis to stakeholders.

Furthermore, spatial-temporal regular prediction maps can then be used to determine the prevailing favourable areas for the species through time-series, the spatial patterns of variation through time and estimate the location of the main source of fishable stock. Additionally, Empirical Orthogonal Function analysis (EOF) can then be used to identifying the main patterns of spatial distribution and temporal evolution of the species, which can then be associated with environmental variables.

The results will be used in the framework of the European research project SAFI (http://www.safiservices.eu/; REA Grant Agreement 607155), that aims to develop indicators to support fisheries and help in management decisions, thus contributing for the sustainable exploitation of the living marine resources by using satellite-derived information.

Estimating abundance from detections of marked and unmarked animals in camera traps

R.M. Dorazio[a] and **D.P. Onorato**[b]

[a]Wetland and Aquatic Research Center
United States Geological Survey
Gainesville, Florida, USA
bdorazio@usgs.gov

[b]Fish and Wildlife Research Institute
Florida Fish and Wildlife Conservation Commission
Naples, Florida, USA
dave.onorato@myfwc.com

Keywords: abundance; camera-trap survey; spatial capture-recapture; species distribution model

Abstract: We propose a new class of spatial models for analyzing the detections of marked and unmarked animals in camera-trap surveys. In these models the spatial distribution of individual activity centers is assumed to follow a spatial point process parameterized to include the effects of spatially varying covariates (e.g., habitat measurements). Similarly, the sequence of individual detection times is assumed to follow a temporal point process during the continuous period of each camera's operation. This process is parameterized to include the effects of trap-specific covariates on an individual's baseline rate of detection so that factors influencing animal movements or behaviors may be estimated. As an illustration we analyzed the detections of Florida panthers in camera-trap surveys of Big Cypress National Preserve and Florida Panther National Wildlife Refuge.

Value of trait-based Bayesian priors in survival and viability analysis for microbats

P. E. Lentini[a], T. J. Bird[b], S. R. Griffiths[c], L. N. Godinho[d] and B. A. Wintle[a]

[a]School of BioSciences
The University of Melbourne
Parkville, Australia
pia.lentini@unimelb.edu.au
b.wintle@unimelb.edu.au

[b]Department of Geography and Environment
University of Southampton
Southampton, England
tomasbird@gmail.com

[c]Department of Zoology
La Trobe University
Bundoora, Australia
S.Griffiths@latrobe.edu.au

[d]Biology Laboratory
The University of Melbourne
Parkville, Australia
lisangodinho@gmail.com

Keywords: Chiroptera, demography, mark-recapture, mortality, population viability, state-space model.

Abstract: Accurate survival estimates are needed to construct robust population models, which are a powerful tool for understanding and predicting the fates of species under scenarios of environmental change. Microbats make up 17% of the global mammalian fauna, yet the processes which drive differences in demographics between species are poorly understood. We collected survival estimates for 44 microbat species from the literature and constructed a Bayesian mixed-effects model to determine the effects of reproductive, feeding, and demographic traits on survival. Our trait-based model indicated that bat species that produce more young per year exhibit lower apparent annual survival, as do males and juveniles compared with females and adults respectively. Using eight years of monitoring data for two Australian species, we demonstrate how knowledge about the effect of traits on survival can be incorporated into Bayesian survival analyses. This approach can be applied to any group and is not restricted to bats or even mammals. The incorporation of informative priors based on traits can allow for more timely construction of population models to support management decisions and actions.

References

Lentini, P. E., Bird, T., Griffiths, S. R., Godinho, L. N. and Wintle, B. A. (2015) A global synthesis of survival estimates for microbats. *Biology Letters* 11, 20150371.

Multistate release-recapture modeling using a graphically driven software interface

John R. Skalski[a], Trisha Lockhart[a], and Adam Pope[b,c]

[a] Columbia Basin Research
School of Aquatic & Fishery Sciences
University of Washington
1325 Fourth Avenue, Suite 1515
Seattle, Washington 98101-2540
skalski@uw.edu, tlockhar@uw.edu

[b] Quantitative Ecology and Resource Management
University of Washington
1325 Fourth Avenue, Suite 1515
Seattle, Washington 98101-2540

[c] Current affiliation: U.S. Geological Survey
Western Fisheries Research Center
Columbia River Research Laboratory
5501-A Cook-Underwood Road
Cook, Washington 98605
apope@usgs.gov

Keywords: release-recapture; multistate model

Abstract:

In the Pacific Northwest, multistate release-recapture models are used to monitor the survival of outmigrating juvenile salmon and returning adult salmon in tributary systems throughout the Columbia Basin. A complex system of over 200 adult and juvenile PIT-tag detection arrays have been installed in tributary systems in Washington, Oregon, and Idaho to monitor migration success. Each tributary system requires a unique multistate model tailored to the number and location of the detectors, the geometry of the river system, and can change as detection arrays enter and leave service.

The burden of model development in the past has often meant tagging studies going without proper design or analysis. To provide regional statistical support, a graphically driven software program has been developed that translates user-created schematics of release-recapture designs into multistate models. Model development previously taking weeks or months can now be accomplished in hours or minutes. Salient survival summaries can be easily calculated by interacting with model schematics on a display screen. The software has applications to a wide variety of terrestrial and aquatic systems. Examples of model construction and analyses will be illustrated using tagging studies from Columbia and San Joaquin rivers.

Efficient MCMC Sampling for Hidden Markov Models and Multistate Capture-Recapture

Daniel Turek[a], Perry de Valpine[b]*, and Christopher J. Paciorek[c]

[a] Department of Statistics
Department of Environmental Science, Policy, and Management
University of California
Berkeley, CA, USA
dturek@berkeley.edu

[b] Department of Environmental Science, Policy, and Management
University of California
Berkeley, CA, USA
pdevalpine@berkeley.edu

[c] Department of Statistics
University of California
Berkeley, CA, USA
paciorek@stat.berkeley.edu

Keywords: Capture-recapture, Effective sample size, Hidden Markov model, Hierarchical model, MCMC, NIMBLE, Sampling efficiency.

Abstract: Traditional Markov chain Monte Carlo (MCMC) sampling of hidden Markov models (HMMs) involves latent states underlying an imperfect observation process, and generates posterior samples for top-level parameters concurrently with nuisance latent variables. When potentially many HMMs are embedded within a hierarchical model, this can result in prohibitively long MCMC runtimes. We study combinations of existing methods, which are shown to vastly improve computational efficiency for these hierarchical models while maintaining the modelling flexibility provided by embedded HMMs. The methods include discrete filtering of the HMM likelihood to remove latent states, reduced data representations, and a novel procedure for dynamic block sampling of posterior dimensions. Using the NIMBLE package for R, we develop and test combined computational approaches using three examples from capture-recapture, although our methods are generally applicable to any embedded discrete HMMs. These combinations provide up to three orders of magnitude improvement in MCMC sampling efficiency, defined as the rate of generating effectively independent posterior samples. In addition to being computationally significant for this class of hierarchical models, this result underscores the potential for vast improvements to MCMC sampling efficiency which can result from combinations of known algorithms.

Integrating stable isotope and parasite data into mark-recovery models to estimate migratory connectivity in Barn Swallows

F. Korner-Nievergelt [a], M.U. Grüebler [a], J. v. Rönn [a]

[a] Swiss Ornithological Institute
Sempach, Switzerland
fraenzi.korner@vogelwarte.ch

Keywords: integration of different data sources; multi-state model; large-scale movement; migratory connectivity.

Abstract: Migratory connectivity has implications both for the migrating individuals and populations (Norris and Marra 2007), and for the ecosystems that are connected by migrating individuals (Bauer and Hoye 2014). Therefore, the knowledge of migratory connectivity is fundamental for understanding many aspects in ecology (Webster et al. 2002).

It is still difficult to measure migratory connectivity because this requires knowledge about the non-breeding distribution of many individuals of many populations. Methods allowing for the study of large numbers of individuals usually provide only imprecise information, either because of spatial heterogeneous sampling probability (e.g., reencounters of marked individuals) or because the location information is vague (e.g., stable isotopes, parasites). By formally combining data from different data sources, the precision and reliability of connectivity measures increases.

We combine ring-recovery, stable isotope and blood parasite data of Barn Swallows *Hirundo rustica* in a single model to estimate migratory connectivity for breeding populations in central Europe and Scandianavia. To this end, we integrated three different two-level models for each of the three data sources. For the ring recovery data, we used a non-Markovian multi-state mark-recovery model. The stable isotope data was included using a normal-mixture model. For the parasite data, we developed a binomial mixture model that connected parasite data from the non-breeding areas with parasite data from the breeding areas.

The three breeding populations clearly showed different non-breeding distributions. Of the Swedish population, the majority (74%) of individuals migrated to southern Africa and the rest of the population mainly stayed in central Africa for the non-breeding period. Of the northern German population around a third migrated to each of western, central and southern Africa. The southern Germany population stayed in western and central Africa (44 and 43%, respectively). Eastern Africa does not seem to be an important non-breeding area for any population (used by less than 6% of each population).

The integration of different data sources within a mark-recapture modelling framework enables the quantification of migratory connectivity on a large-scale. The results show a strong connectivity between Scandinavia and southern Africa as well as between southern Germany and western-central Africa through Barn swallow migration. Further, the northern German population seems to behave intermediately. This result illuminates important questions regarding the evolution of bird migration.

References

Bauer, S. and Hoye, B.J. (2014) *Science* 344:1242552.
Norris, D.R. and Marra, P. P. (2007) *The Condor* 109:535-547.
Webster, M.S., Marra, P.P., Haig, S.M., Bensch, S. and Holmes, R. (2002) *Trends in Ecology and Evolution* 17:76-83.

Novel application of a quantitative spatial comparison tool to species distribution data

Esther L. Jones[a,b], Luke Rendell[b] , Enrico Pirotta[c] and Jed A. Long[d]

[a] Centre for Research into Ecological and Environmental Modelling
University of St Andrews
St Andrews, UK
el298@st-andrews.ac.uk

[b] Sea Mammal Research Unit, Scottish Oceans Institute
University of St Andrews
St Andrews, UK
ler4@st-andrews.ac.uk

[c] Institute of Biological and Environmental Sciences
University of Aberdeen
Aberdeen, UK
pirotta.enrico@gmail.com

[d] Department of Geography and Sustainable Development,
School of Geography and Geosciences
University of St Andrews
St Andrews, UK
jed.long@st-andrews.ac.uk

Keywords: edge effects, map comparison, moving window, *Physeter microcephalus,* reflection algorithm, spatial heterogeneity, sperm whale, SSIM index, uncertainty

Abstract: Comparing geographically referenced maps has become an important aspect of spatial ecology (e.g. planning for conservation when detecting and quantifying change in species distributions). Whilst humans are adept at recognising and extracting structure from maps (i.e. identifying spatial patterns), quantifying these structures can be difficult. Here, we show how the Structural Similarity (SSIM) index (Wang *et al.* 2004), a map comparison method adapted from techniques developed in computer science to determine the quality of image compression, can be used to compare spatial ecological data. We extend the SSIM index to incorporate uncertainty from the underlying spatial models, and provide an algorithm to correct for internal edge effects so that loss of spatial information from the map comparison is limited. The SSIM index uses a spatially-local window to calculate statistics based on local mean, variance, and covariance between the maps being compared. A number of statistics can be calculated using the SSIM index, ranging from a single summary statistic to quantify similarities between two maps, to maps of similarities in mean, variance, and covariance that can provide additional insight into underlying biological processes. We demonstrate the applicability of the SSIM approach using a case study of sperm whales (*Physeter microcephalus*) in the Mediterranean Sea and identify areas where local-scale differences in space-use between groups and singleton whales occur (Pirotta *et al.* 2011). We show how novel insights into spatial structure can be extracted, which could not be obtained by visual inspection or cell-by-cell subtraction. As an approach, SSIM is applicable to a broad range of spatial ecological data, providing a novel, implementable tool for map comparison.

Identification of influential predictors in 'large-data' contexts with spatial and temporal dependencies

K. Shuvo Bakar [a,b] and Ralph Mac Nally [a]

[a] Institute for Applied Ecology
The University of Canberra
Bruce ACT 2617, Australia
shuvo.bakar@gmail.com
Ralph.MacNally@gmail.com

[b] Present address: Research School of Social Sciences
The Australian National University
Canberra, Australia.

Keywords: adaptive shrinkage, hierarchical Bayesian models, species distribution models, stochastic variable selection

Abstract: Analyzing and modelling massive data sets is a demanding task for ecological applications. Identifying the best species distribution model is very often a key concern to ecological researchers, especially the identification of important predictors of species' occurrence or abundance in complex spatial and temporal contexts. To address these challenges, we developed a new technique for spatio-temporal modelling problems that incorporates stochastic variable selection (George & McCulloch, 1993) with adaptive shrinkage (Tibshirani, 1996). The resulting model is specified within a hierarchical Bayesian framework and Markov chain Monte Carlo techniques are used to make inferences. The work is novel in solving the complex hierarchical modelling structures with a simple and fast adaptive solution that includes model uncertainty and importance for each predictor variables. We applied the method to gain insights on the environmental and anthropogenic factors appearing to drive spatial and temporal dynamics of North American and Australian birds

References

George, E. I., & McCulloch, R. E. (1993). Variable selection via Gibbs sampling. Journal of the American Statistical Association, 88(423), 881-889.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B (Methodological), 267-288.

Spatial function estimation on regions with irregular boundaries

R.P. Barry[a] and **J. McIntyre**[b]

[a]Department of Mathematics and Statistics
University of Alaska Fairbanks
Fairbanks, AK 99775
rpbarry@alaska.edu

[a]Department of Mathematics and Statistics
University of Alaska Fairbanks
Fairbanks, AK 99775
jpmcintyre@alaska.edu

Keywords: boundary bias; diffusion; kernel density estimation; kernel regression; point process

Abstract: Estimation of a density or regression function from geospatial data is a common goal in many studies. For example ecologists use such data to construct maps of the amounts or characteristics of plants and animals based on location within a region. Popular methods for estimating these spatial functions include nonparametric kernel density estimation and kernel regression estimation, as well as parametric spatial regression (e.g. kriging). Often however the region of interest includes irregular boundaries beyond which the function is undefined. For example a function describing water pH should be well-defined for locations within a lake, but not for locations on shore. Traditional estimators ignore irregular boundaries and holes in a region, leading to biases. We present an alternative estimator that accounts for boundaries and holes in the estimation process. The estimator is a type of kernel estimator that is based on the density of random walks of length k on a lattice that is constrained to stay within the region's boundaries. Unbiased cross validation is used to find the optimal walk length k. Simulations show that the estimator is superior to other estimators in the presence of boundaries, and comparable in the absence of boundaries. We present several examples to illustrate the method.

Monitoring Through Many Eyes

E.E. Peterson[a], K. Anthony[b], T. Bednarz[a], R. Brown[c], J. Caley[b], B. Christensen[d], A. James[d],
K. Mengersen[a], A. Siliato[d], J. Vercelloni [a], G. Winter[d]

[a] ARC Centre for Excellence in Mathematical and Statistical Frontiers,
Queensland University of Technology
Brisbane, Australia
Erin.Peterson@qut.edu.au

[b] Australian Institute of Marine Science,
Townsville, Australia

[c] Department of Information Systems,
Queensland University of Technology
Brisbane, Australia

[d] Visualisation and eResearch Lab, Institute for Future Environments
Queensland University of Technology
Brisbane, Australia

Keywords: Citizen science, immersive elicitation, geostatistical modeling, coral cover, Great Barrier Reef

Abstract: There are many critical environmental issues facing the Great Barrier Reef including crown-of-thorns starfish outbreaks, loss of habitats for endangered or vulnerable species, water quality impacts on coral cover, coral bleaching, the impact of cyclones and the effectivness of management strategies and interventions. Monitoring data are relatively sparse given the size of the Reef, but citizen-contributed data has the potential to significantly increase the spatio-temporal coverage of data used to inform management decisions. We have developed an immersive, predictive, visual environment that can make use of multiple data sources to monitor the health of the Great Barrier Reef. The infrastructure allows crowd-sourced 2-D and 360-degree imagery taken by recreational and professional divers to be uploaded and geo-located within an existing digital map. Elicitation modules have been developed to allow regular citizens to classify benthic habitat in the images, while marine ecologists can be immersed into the digital environment to elicit information and interpretations of reef health issues. This expert "prior information" can be combined with existing monitoring data to build spatio-temporal models used to gain a better understanding of the ecosystem, inform spatially explicit management decisions, and guide future data collection. Our case study focuses on coral cover near Heron Island in the Great Barrier Reef, but the modelling infrastructure is generally applicable to other citizen-science based monitoring efforts in the terrestrial, freshwater, or marine environments.

Attributing changes in the distribution of species abundance to weather variables using the example of British breeding birds

Cornelia S. Oedekoven[1*], Philip J. Harrison[1], David A. Elston[2], Mark J. Brewer[2], Stephen T. Buckland[1], Alison Johnston[3] and James W. Pearce-Higgins[3]

[1]Centre for Research into Ecological and Environmental Modelling, University of St. Andrews, The Observatory, St Andrews, Fife, KY16 9LZ, UK;

[2]Biomathematics & Statistics Scotland, Craigiebuckler, Aberdeen, AB15 8QH, UK;

[3]British Trust for Ornithology, The Nunnery, Thetford, Norfolk, IP24 2PU, UK

[*]cso2@st-andrews.ac.uk

Keywords: generalized additive models, generalized linear models, geometric mean of relative abundance, spatio-temporal modelling, climate change, UKCP09 climate projections.

Abstract:

Understanding the drivers of spatio-temporal variation in species abundance or biodiversity can be challenging. A further challenge is understanding how this variation might be driven by changes in climate. The 'standard' approach for including climate information into spatio-temporal models of species abundance or biodiversity is to include weather variables where each variable is included as a single covariate (as a linear term, low order polynomial or as some kind of smoother). This, however, may lead to confounding issues of the spatial and temporal effects of the covariates as the effect of a particular change in the weather covariate is assumed constant, regardless of whether the change takes place over space or over time. We have developed a novel approach to separating out different types of change in any particular weather covariate. We decompose the initial weather covariate into three new covariates, separating out spatial variation in weather (averaging over years), temporal variation in weather (averaging over space) and a space-time anomaly term. These three new covariates are each fitted separately in the model.

We illustrate the approach using generalized additive models applied to count data for a selection of species from the UK's Breeding Bird Survey collected between 1994 and 2013. The weather covariates included in our models are the centred mean temperatures during the preceding winter and the centred mean temperatures and rainfall in the preceding breeding season. We compare models that include these covariates directly with models including decomposed components of the same covariates. We consider both linear and smooth terms of all covariates. Furthermore, we investigate the necessity of the space-time smooth in the model after fitting the weather covariates as smooth effects. A quasi-Poisson error distribution is used for the bird count data in order to allow for over- or underdispersion. Our results indicate that the decomposed weather variables can have very different effects in the model, with the spatial term often being dominant. AIC tended to select the most complex model fitted, i.e. including smooth effects of two centred and one decomposed weather covariates as well as the space-time smooth.

Estimation of vital rates of unmarked breeding birds at large scales

Edwige Bellier[a,b], **Marc Kéry**[a] and **Michael Schaub**[a]

[a]Swiss Ornithological Institute
Sempach, Switzerland

[b]The Arctic University of Norway
Tromsø, Norway
edwige.bellier@uit.no

Keywords: counts, survival, recruitment, strength of density-dependence, environmental stochasticity, detection probability, short-lived species

Abstract: Linking spatio-temporal variation in environmental factors to variation in demographic rates is essential for a mechanistic understanding of the dynamics of a population. We investigate how the variation in vital rates is related to variation in environmental factors for four common breeding birds (robin, blackcap, great tit and mistle trush). We use new dynamic N-mixture models that simultaneously account for effects of intraspecific density and environmental fluctuations and accommodate imperfect detection. We propose a multicriteria model selection procedure to determine if density-dependence occurs either in survival or recruitment or in both vital rates. Once the model structure for density-dependence is selected, environmental covariates are included either in recruitment or survival or both vital rates. Our approach is useful to shed light on how density-dependent vital rates vary as functions of environmental covariates at large scales.

Statistical Ecology and Management: Boundary Issues

L. New [a]

[a] Department of Mathematics and Statistics
Washington State University, Vancouver
Vancouver, WA, USA
leslie.new@wsu.edu

Keywords: Bayesian statistics, golden eagles, government, management

Abstract: For many, statistical ecology has two goals: the development of new, useful statistical approaches and their application in the context of conservation and management. Statistical tools, such as distance sampling and mark-recapture, have been invaluable in enabling more informed decision making, and are always evolving and being improved upon. When combined with a literature full of new tools and models, managers and practitioners in the field may find it difficult to keep up with current best practice. This raises the question as to the role of statistical ecologists in the application of their research. If we want others to adopt our approaches, what responsibility do we have to them? To ensure our work is being used correctly? At what point does the responsibility transfer to the user? To stimulate discussion, I will present a case study on golden eagles (*Aquila chrysaetos*). I helped the US Fish and Wildlife Service (USFWS) construct a Bayesian collision risk model to predict eagle fatalities on wind facilities prior to construction (New et al. 2015). The model is used in the permitting process for US wind facilities, and is regularly implemented by USFWS biologists. I will discuss my role in the continued, practical application of this work, how the boundaries between research and application have blurred, and when those boundaries have needed to be enforced.

References

New, L., Bjerre, E., Millsap, B., Otto, M.C. and Runge, M.C. (2015) Incorporating uncertainty into a collision risk model for predicting avian fatalities at wind facilities: An example using golden eagles, *Aquila chrysaetos. PLOS ONE,* 10(7): e0130978. doi: 10.1371/journal.pone.0130978

A goodness-of-fit test for parametric mean functions

Y.H. Huang and **C. C. Wen**

Department of Mathematics
Tamkang University
New Taipei City, Taiwan
yhhuang@mail.tku.edu.tw

Keywords: goodness-of-fit test, Hilbert space, inner product, orthogonality, parametric model

Abstract: In the probability Hilbert space, which consists of functions of response variables and covariates with finite second moments, the conditional mean can be viewed as the projection of the response variable onto the subspace of functions of covariates. As a result of the projection, the random errors, which are defined as differences between the response variable and its conditional mean, are orthogonal to this subspace. If the mean function is correctly specified, then the residuals of the fitted model are close to these random errors and are approximately orthogonal to this subspace. But if the model is misspecified, one is unable to find a suitable projection, and the residuals are not orthogonal to this subspace. Given this geometric motivation, we develop a goodness-of-fit test based on whether certain inner products have zero means. This testing procedure is novel and easy to implement. Using a simulation study, we show that the proposed test has favorable features compared with many existing methods.

deBInfer: Bayesian inference for dynamical models of biological systems in R

P.H. Boersch-Supan[a] and **L.R. Johnson**[b]

[a]Department of Integrative Biology
University of South Florida
Tampa, FL, USA
phb@usf.edu

[b]Department of Integrative Biology
University of South Florida
Tampa, FL, USA
lrjohnson@usf.edu

Keywords: Bayesian inference; differential equations; R package; energy budget models; population dynamics;

Abstract:
Differential equations (DEs) are commonly used to model the temporal evolution of biological systems, but statistical methods for comparing DE models to data and for parameter inference are relatively poorly developed. This is especially problematic in the context of biological systems where observations are often noisy and only a small number of time points may be available.

Bayesian approaches offer a coherent framework for parameter inference that can account for multiple sources of uncertainty, while making use of prior information. This approach further offers a rigorous methodology modeling the link between unobservable model states and parameters, and observable quantities.

We present `deBInfer`, an R package implementing a Bayesian framework for parameter inference in DEs. `deBInfer` provides templates for the DE model, the observation model, the data likelihood, and the prior distributions of the parameters, as well as a Markov chain Monte Carlo (MCMC) procedure to estimate the posterior distributions of the parameters and any derived quantities, including the model trajectories. Further functionality is provided to facilitate MCMC diagnostics and the visualisation of the posterior distributions of model parameters and trajectories. The templating approach makes our code applicable to a wide range of DE models.

We demonstrate the application of `deBInfer` to ordinary and delay DE models drawn from the fields of metabolic ecology and disease ecology, respectively, by fitting a dynamic energy budget model to individual growth data of wandering albatross (*Diomedea exulans*) chicks, and a population growth model to zoospore counts of the chytrid fungus *Batrachochytrium dendrobatidis*, a pathogen that has devastated amphibian populations around the globe.

Displaying complex ecological data

S. Wongsai [a], N. Kongchouy [b] & D. McNeil [c]

[a] Department of Computer Engineering
Faculty of Technology and Environment, Prince of Songkla University
Phuket 83120, Thailand
feb_dao@hotmail.com

[b] Department of Mathematics and Statistics
Faculty of Science, Prince of Songkla University
Hat Yai 90110, Thailand
nootchanath.k@psu.ac.th

[c] Department of Statistics
Faculty of Science, Macquarie University
Sydney 2109, Australia
don.mcneil@mq.edu.au

Keywords: Graphical methods; statistical models; data transformation; ecology

Abstract: Ecological data are complex and diverse, involving many correlated variables with highly skewed distributions at different levels in environmental food-webs, and thus present difficult challenges for statisticians. And the modern explosion of environmental data from remote sensing makes it even more important to develop appropriate models and graphs that environmental scientists can understand and conveniently use. In this paper we present a logical framework for modelling and graphing such data that involves no new statistical methods or computer software, but focuses on appropriate use of standard methods (McNeil, 2015). Illustrative examples use data from a sizable ongoing study involving monthly observations since 2005 on five food-web levels – water characteristics, phytoplankton and zooplankton, benthic fauna, fish larvae and fish – from a tropical estuary-river system. Outcomes include taxa prevalence, abundance and diversity, modelled separately using factor analysis, logistic regression, and linear regression with transformed data to efficiently satisfy statistical assumptions, but graphed with the object of highlighting the information content, quality of scientific evidence, and interpretability of results.

Reference

McNeil, D. (2015) Epidemiological statistics – Further issues. *Wiley StatsRef: Statistics Reference Online*, DOI: 10.1002/9781118445112.stat00321.pub2

Validation of state-space models fitted as mixed effects models using Template Model Builder

U.H. Thygesen[a], **C.M. Albertsen**[a], **C.W. Berg**[a],
K. Kristensen[a], **A. Nielsen**[a], **M.W. Pedersen**[a]

[a]Technical University of Denmark
National Institute for Aquatic Ressources
Charlottenlund, Denmark
uht@aqua.dtu.dk

Keywords: State-space models, maximum likelihood, model validation, residual analysis.

Abstract:

Statistical models in ecology often follow the state space paradigm, while recent computational methods for estimation in these models model do not rely on the state space structure but rather on general methods for nonlinear mixed effects models. Validation is an important part of statistical modeling, which does not always receive as much attention as for example hypothesis testing. Here, we discuss principles, practicalities and implementation of validating state-space models, which have been fitted with Template Model Builder (TMB), a tool for estimation in general nonlinear and non-Gaussian mixed effects models. It is widespread practice to base model validation on residuals such as the difference between observations and posterior means, but it is known that this approach is flawed, and we demonstrate these flaws. More correct approaches are to base validation on one-step prediction errors, or on a single sample of unobserved random variables from the posterior distribution. We discuss these principles in the context of simulated examples, including nonlinear relationships, non-Gaussian distributions, and discrete observations, and we demonstrate implementations in TMB which require minimal coding effort. Finally we apply the methods to two real data sets, one related to geolocation of seals, and one related to a single-species unstructured biomass model of population dynamics.