Post-Congress Symposia

Advances in statistical philosophy and experimental design in behavioral ecology

Discussing a new "ABC" of statistical practices and beyond

Organizers: László Zsolt Garamszegi & Shinichi Nakagawa

Contributors: Sara Calhim, Ned Dochtermann, Gergely Hegyi, Peter Hurd, Christian Jørgensen, Nobuyuki Kutsukake, Marc Lajeunesse, Kimberly Pollard, Holger Schielzeth & Matthew Symonds
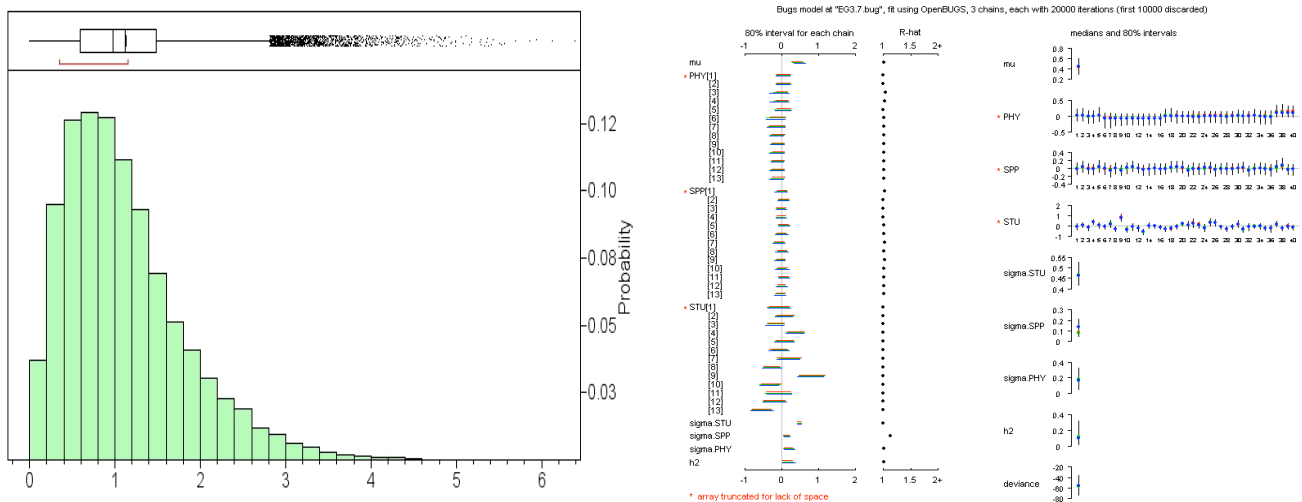
Time: August 15th – 9.00am to 3.00pm

Place: will be assigned, please check the information desk

Registration: please sign up at the information desk

PROGRAM OVERVIEW

09.00-09.05 – General introduction (Garamszegi)

Part A/B (chair: Garamszegi)

09.05-09.25 – Talk 1: The use of Akaike's information criterion in behavioral ecology (speaker: Dochterman)

09.25-09.45 – Talk 2: Stepwise selection and information theory in ecology and behavior (speaker: Hegyi)

09.45-10.05 – Talk 3: AIC's next top model: model averaging as a solution to model selection uncertainty (speaker: Symonds)

10.05-10.30 – Talk 4: The ABCs of Bayesian methods for behaivoural ecologists (speaker: Nakagawa)

10.30-10.50 – coffee break

Part B/C (chair: Nakagawa)

10.50-11.10 – Talk 5: X and Y w/ Z: analyses of the zero-inflated data in behavioural ecology (speaker: Kutsukake)

11.10-11.30 – Talk 6: A model based approach to account for phylogenetic history and test evolutionary hypotheses with meta-analysis (speaker: Lajeunesse)

11.30-11.50 – Talk 7: Conclusions beyond support: over-confident estimates in mixed models (speaker: Schielzeth)

11.50-12.10 – Talk 8: Bootstrapping confidence intervals on effect sizes (speaker: Hurd)

12.10-13.00 – lunch

…and beyond (chair: Nakagawa)

13.00-13.20 – Talk 9: The influence of sampling design on individuality calculations (speaker: Pollard)

13.20-13.40 – Talk 10: Collecting comparative data - a look into avian testes size (speaker: Calhim)

13.40-14.00 – Talk 11: Within-species variation and heterogeneity in data quality in phylogenetic and comparative studies (speaker: Garamszegi)

14.00-14.20 – Talk 12: Mechanistic modeling as a method for testing and generating hypotheses (speaker: Jørgensen)

14.20-14.25 – Conclusion & take home message (Nakagawa)

14.25-15.00 - General discussion

ABSTRACTS

Overall theme: **Advances in statistical philosophy and experimental design in behavioural ecology**

László Zsolt Garamszegi and Shinichi Nakagawa

Department of Biology, University of Antwerp, Wilrijk, Belgium (e-mail: laszlo.garamszegi@ua.ac.be)
Department of Zoology, University of Otago, Dunedin, New Zealand (e-mail: shinichi.nakagawa@otago.ac.nz)

We aim to bring together researchers with different interests in behavioural ecology to discuss recent significant developments in the interpretation of behavioural and ecological data. Analytical tools that incorporate statistical philosophies not relying on statistical significance have been highlighted in recent years, but are still in limited use in our field, because of common beliefs and familiarity with classical approaches. We feel that an exhaustive revision of the classical and new methods is timely. We would like to focus on the basic philosophy behind different approaches by considering pro- and contra arguments, and provide a broad array of examples of biological questions that can be tested in correlative or experimental designs, in intra- or inter-specific contexts, and in different taxonomic groups. We propose a discussion along an "ABC" framework, in which "A" stands for AIC (Akaike's Information Criterion), a popular tool of information theoretic approaches that address the trade-off between model complexity and goodness of fit. This allows an objective assessment of the potentially large number of competing models that could describe the data equally well. "B" refers to Bayesian inference, in which new empirical evidence is combined with past knowledge to update or newly infer the probability of the hypotheses under test. Bayesian approaches also harness us with capabilities of parameter estimation in a way that gives an alternative to the classical framework. "C" emphasizes the importance of confidence intervals, which give the precision of parameter or effect size estimates. Effect size estimates and their confidence intervals should be at the heart of statistical inference because they relate to biological importance in a way that statistical significance does not. Effect size estimates also facilitate meta-analysis, which has recently established itself as an essential tool for quantitative review in the field. The three components of our ABC framework complement each other, but are in contrast with conventional hypothesis testing based on statistical significance. With the symposium, we hope to help the objective decision of behavioural ecologists looking for a suitable way of analyzing their data.

Talk 1: **The use of Akaike's Information Criterion in behavioral ecology**

Ned A. Dochtermann

Program in Ecology, Evolution and Conservation Biology, Department of Biology, University of Nevada, Reno (email: dochterm@unr.nevada.edu)

Behavioral ecologists deal with complex systems and seek to understand how multiple interacting evolutionary factors have shaped behavioral phenotypes. Due to the complexity of questions being asked, a conventional null hypothesis testing framework will not always provide the most effective route to increase our understanding of the evolutionary causes and consequences of observed behaviors. When tied explicitly to existing theory, model selection methods using Akaike's information criterion (AIC) offer a flexible alternative with which to distinguish between a priori hypotheses. Model selection using AIC scores is demonstrated with examples from the evolutionary ecological literature and by focusing on an example of AIC use in the evaluation of hypotheses of behavioral syndrome structure.

Talk 2: **Stepwise selection and information theory in ecology and behavior**

Gergely Hegyi and László Zsolt Garamszegi

Department of Systematic Zoology and Ecology, Eötvös Loránd University
Pázmány Péter sétány 1/C, H-1117 Budapest, Hungary (email: gehegyi@yahoo.com)
Department of Biology, University of Antwerp, Wilrijk, Belgium (e-mail: laszlo.garamszegi@ua.ac.be)

Research in ecology and behavior has been increasingly focusing on multiple potential determinants of phenomena, so the use of statistical models with multiple predictors has become general in the literature. However, drawing reliable conclusions usually necessitates the consideration of initially complex models and subsequent simplification. The traditional way of model simplification in ecological and behavioral research has been the use of threshold-based removal-reintroduction algorithms, that is, stepwise selection. This method has recently been strongly criticized for multiple reasons, including the biased distribution of the resulting parameter estimates, the incongruence of different selection algorithms, the redundancy of repeated parameter testing, and the potentially incorrect reliance on a single final model. A more recent alternative solution to stepwise selection is an information theoretic (IT) approach that quantifies the relative suitability of multiple, potentially non-nested models based on a balance of model fit and the accuracy of estimates. This method is now increasingly propagated. Here we attempt to provide an objective view of stepwise and IT approaches. First, we explain the necessity of model simplification in ecological and behavioral studies. Second, we discuss the multiple ways of stepwise selection, pinpoint its inherent weaknesses, and suggest answers to some of the recent criticism against it. Third, we emphasize the advantages of the IT approach, but also identify problems which currently make it a difficult or even dangerous method in studies of ecology and behavior. Finally, we suggest avenues for future development in the field of model simplification.

Talk 3: **AIC's next top model: Model averaging as a solution to model selection uncertainty**

Matthew Symonds

Department of Zoology, University of Melbourne, Victoria 3010, Australia (email: symondsm@unimelb.edu.au)

The ability to find the best model by comparing the relative likelihoods of all possible models is one of the strengths of AIC analyses. However, frequently, many alternative models are almost as likely as the best-supported model (i.e. have approximately equal AIC values), thus reducing confidence in conclusions. In a recent analysis of environmental factors predicting species diversity in Australian birds, I identified hundreds of 'almost equally likely' models using AIC. I show how model averaging helped to identify the parameters that were most strongly represented among these models, and discuss the advantages of this technique in ecological analyses.

Talk 4: **The ABCs of Bayesian methods for behavioural ecologists**

Shinichi Nakagawa

Department of Zoology, University of Otago, PO Box 56, Dunedin, New Zealand (email: shinichi.nakagawa@otago.ac.nz)

Bayesian statistics have recently gained its popularity in many areas including phylogeny construction and complex ecological modelling. Behavioural ecologists seem to be among the last to employ this flexible framework in their routine analysis. Here, I start with explaining fundamental difference between frequentist and Bayesian approaches. Then, I discuss the advantages of Bayesian methods especially in two contexts: 1) mixed-effects modelling and 2) multi-response variable analysis. In the frequentist framework, the calculation of the degree of freedom is problematic in mixed-effects models whereas in the Bayesian framework, this problem can be effectively handled. I explain the concept of effective parameter numbers and deviance information criterion (DIC), which is a Bayesian equivalent of Akaike's information criterion (AIC). Then, I describe one of major strengths of the Bayesian framework, which allows implementations of statistical models with multiple response variables (this includes the modelling of correlations between these response variables while fitting multiple predictors). Also, I will introduce available Bayesian software and books which may be readily accessible to Behavioural Ecologists. Therefore, Behavioural Ecologists can try and start using Bayesian statistics from the next day.

Talk 5: **X and Y w/ Z: analyses of the zero-inflated data in behavioural ecology**

Nobuyuki Kutsukake

Department of Evolutionary Studies of Biosystems, The Graduate University for Advanced Studies, Hayama, Miura-gun, Zushi, Kanagawa, 240-0193, Japan (email: kutsukake_nobuyuki@soken.ac.jp)

In behavioural ecology, it is common that many zero values are contained in the data (so called the zero-inflated data). Because the data transformation is not useful for normalization, it is required to employ special statistical methods to analyze the zero-inflated data. In this talk, I will present two methods to deal with the zero-inflated data. The first method is to conduct two separate analyses - the former analysis investigates factors predicting whether the dependent term is zero or non-zero and the latter one investigates factors predicting the dependent term of nonzero. The second method is to model the zero-inflated data distribution by the Bayesian method. I will present the example from the analyses of allogrooming behaviour among cooperatively breeding meerkats.

Talk 6: **A model based approach to account for phylogenetic history and test evolutionary hypotheses with meta-analysis**

Marc J. Lajeunesse

Dept. of Ecol. & Evol. Biol., Cornell University, E145 Corson Hall, Ithaca, NY 14853-2701 (email: mjL63@cornell.edu)

Approaches to meta-analysis in biology have drawn heavily from the medical, social and educations sciences. However, a unique and emerging challenge for biology is the lack of statistical methods and protocol to synthesize studies while simultaneously accounting for phylogenetic non-independence of taxa. Given that an explicit goal of meta-analysis is to generalize across published research from a broad range of taxa, then the phylogenetic non-independence of these taxa may threaten the validity of these generalizations. Here, I present a statistical framework that integrates phylogenetic information into conventional meta-analysis when a) taking a weighted average effect size using fixed- and random-effects models and b) testing for homogeneity of variances. In addition, I describe a protocol for testing evolutionary hypotheses—a first for meta-analysis—by introducing methods that a) evaluate phylogenetic conservatism and b) contrast neutral and adaptive models of evolution based on an Ornstein-Ulehlenbeck process. This model based approach uses AIC scores to evaluate the fit of different evolutionary hypotheses on meta-analysis. Finally, I illustrate these methods by integrating phylogenetic information in a meta-analysis on why mating with virgin males can result in greater reproductive output for lepidopteran females, and how the magnitude of this reproductive success can relate to the origin of monandrous or polyandrous mating.

Talk 7: **Conclusions beyond support: over-confident estimates in mixed models**

Holger Schielzeth* and Wolfgang Forstmeier

*Max Planck Institute for Ornithology, Postfach 1564, 82305 Starnberg, Seewiesen, Germany (*email: schielz@mail.orn.mpg.de)*

Mixed effect models are frequently used to control for the non-independence of data points (e.g. repeated measures taken from the same individuals), when the aim is to estimate fixed effects and to test their significance. This is often done by including individual-specific (random) intercepts. The widespread believe is that this controls for all types of pseudoreplication within individuals. However, this is often not the case. If the aim is to estimate effects that vary within as well as between individuals (i.e. there are individual-specific slopes), random-intercept models are likely to give overconfident estimates leading to conclusions that are not supported by the data. By allowing slopes as well as intercepts to vary between individuals, it is possible to account for the non-independence of slopes. Such random-slope models are easily implemented in standard statistical software and will give the appropriate standard errors for the population slopes.

Talk 8: **Bootstrapping confidence intervals on effect sizes**

Peter L. Hurd

Department of Psychology, University of Alberta, Edmonton, Alberta, Canada, T6G 2E9 (email: phurd@ualberta.ca)

Bootstrap re-sampling allows confidence intervals to be estimated for statistics that do not have simple sampling distributions. Relatively robustness and simple, this technique is becoming more and more popular as its usefulness becomes more and more apparent. For those unfamiliar with bootstrapping I will provide a short introduction explaining the underlying logic and methodology required to teach the topic to undergraduate students. For those with a passing familiarity, I will then describe the use of the bootstrap to estimate population parameters in two situations, one successful, the other more of a cautionary tale. A brief discussion of some of the caveats and limitations of the technique will follow. Bootstrapping has applications beyond simple descriptive statistics. The ability to generate probability distributions for parameters may be fruitfully applied to problems in inferential statistics. For those whose firm grasp on statistics extends further than my own, I will finish by attempting to explain how the bootstrap may be used in some pernicious hypothesis testing problems.

Talk 9: **The influence of sampling design on individuality calculations**

Kimberly A. Pollard*, Daniel T. Blumstein and Suzanne C. Griffin

*University of California, Los Angeles (UCLA); University of California, Los Angeles (UCLA); University of Montana-Missoula (*email: kpollard@ucla.edu)*

Quantifying individuality is important to many behavioral studies, and a variety of methods have been used. However, it is not yet known how robust certain individuality metrics are to perturbations in sampling design. Using alarm call data from social sciurid rodents, we test the effects of sampling effort on the calculation of Beecher's information statistic for individuality. Results indicate the statistic is influenced by the number of observations per session, but only slightly influenced by the number of animals and number of recording sessions. The statistic's comparative value remains strong; it can indicate one species (or modality, trait, etc.) is more individualistic than another. However, sampling design must be controlled to ensure accurate comparison.

Talk 10: **Collecting comparative data – a look into avian testes size**

Sara Calhim

Department of Biology, Queen's University, Kingston, ON, K7L 3N6, Canada (email: calhims@queensu.ca)

Across a range of taxa, relative testes size is considered to be a reliable index of sexual selection. Avian testes are located internally and show marked seasonal change, which makes estimation of their maximum size during breeding difficult. Previous comparative studies have used different methods, several sources of data, and relied on hitherto unchecked methodological assumptions, to estimate testes size. First, we demonstrate how and why the testes size estimates vary across studies, and suggest more precise guidelines for increased accuracy. Second, using a smaller, but more reliable dataset, we show that some of the results reported in comparative analyses may have been affected by the low quality of the large datasets used in those previous studies. In this particular trade-off, more is certainly not better.

Talk 11: **Within-species variation and heterogeneity in data quality in phylogenetic and comparative studies**

László Zsolt Garamszegi

Department of Biology, University of Antwerp, Wilrijk, Belgium (e-mail: laszlo.garamszegi@ua.ac.be)

Comparative analyses aim to explain variations in phenotypes that occur at the across species level. Such approaches inherently rely on the assumption that species-specific means that are usually calculated from intraspecific samples of modest sample size are biologically meaningful. However, variation within-species is not negligible. Measurement errors, fluctuations in behaviour, seasonal variations, population or individual differences all reduce the repeatability of a trait. In a non-phylogenetic context, low repeatability can raise type II error (random noise) only, while in phylogenetic study, it can also lead to type I error (bias). Therefore, within-species variation should be accounted for in phylogenetic comparative studies, which has been suggested by a number of theoretical papers. However, common practice shows that although researchers exercise great care in accounting for type I errors caused by phylogenetic relationships, the problems posed by within-species variation are usually neglected. A meta-analyses of more then 200 comparative analyses, all using phylogenetic corrections, revealed that only few studies report within-species repeatabilities, but many of them are using species-specific estimations from samples with considerably different sample size and variation. I will demonstrate that within-species variation and heterogeneous sampling may affect comparative results, and will also review some potential analytic solutions to eliminate the problem. These problems and solutions equally apply to analysing features of individuals (or other sampling units) when sampling effort or variation differs among individuals.

Talk 12: **Mechanistic modeling as a method for testing and generating hypotheses**

Christian Jørgensen

Department of Biology, University of Bergen, Box 7803, N-5020 Bergen, Norway
(email: christian.jorgensen@bio.uib.no)

This talk will not be about statistics but about modeling the processes that generate patterns in data. While statistical analysis identifies relationships that are suggestive of underlying mechanisms, process-based modeling can help ascertain whether the suggested mechanism is indeed the causative link. Science aims at explaining how nature works. I will argue that to do so, we need two approaches in parallel. First, we need data collection and subsequent statistical analysis to identify relationships that suggest mechanisms or causation. In a second step, we need process-based models to ascertain whether those mechanisms are sufficient to explain the observed patterns. These models will need to include biological mechanisms operating at the individual level. Scaling these processes up will reveal emergent patterns at higher levels such as the population. This approach of modeling individuals and mechanisms to generate higher-level patterns is often referred to as pattern-oriented modeling. A role for such modeling is not only to test hypotheses generated by statistical data-mining, but also to reveal new patterns and generate new hypotheses for further statistical analysis. I will illustrate the talk with examples from fish behavior and fish population dynamics.