

# Modelling animal social networks: New solutions and future directions

Damien R. Farine<sup>1,2,3</sup> 

<sup>1</sup>Division of Ecology and Evolution, Research School of Biology, Australian National University, Canberra, Australian Capital Territory, Australia

<sup>2</sup>Department of Evolutionary Biology and Environmental Science, University of Zurich, Zurich, Switzerland

<sup>3</sup>Department of Collective Behaviour, Max Planck Institute of Animal Behavior, Konstanz, Germany

## Correspondence

Damien R. Farine  
Email: [damien.farine@anu.edu.au](mailto:damien.farine@anu.edu.au)

## Funding information

Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: PCEFP3\_187058; H2020 European Research Council, Grant/Award Number: 850859

**Handling Editor:** Mariano Rodriguez-Cabal

## Abstract

**Research Highlight:** Ross, C. T., McElreath, R., & Redhead, D. (2023). Modelling animal network data in R using STRAND. *Journal of Animal Ecology*. <https://doi.org/10.1111/1365-2656.14021>. One of the most important insights in ecology over the past decade has been that the social connections among animals affect a wide range of ecological and evolutionary processes. However, despite over 20 years of study effort on this topic, generating knowledge from data on social associations and interactions remains fraught with problems. Redhead et al. present an R package—STRAND—that extends the current animal social network analysis toolbox in two ways. First, they provide a simple R interfaces to implement generative network models, which are an alternative to regression approaches that draw inference by simulating the data-generating process. Second, they implement these models in a Bayesian framework, allowing uncertainty in the observation process to be carried through to hypothesis testing. STRAND therefore fills an important gap for hypothesis testing using network data. However, major challenges remain, and while STRAND represents an important advance, generating robust results continues to require careful study design, considerations in terms of statistical methods and a plurality of approaches.

## KEYWORDS

animal social network analysis, Bayesian statistics, generative network models, social behaviour

The social environment plays a major role in shaping ecological processes. For example, the size of a group that an individual experiences will affect its risk of being eaten upon encountering a predator; the number of contacts it has will affect its exposure to pathogens; and who it interacts with (or their traits) can affect the behaviour it expresses (Cantor et al., 2021; Ward & Webster, 2016). These effects can, in turn, impact the social strategies that individuals express, generating socio-eco-evolutionary feedbacks (Cantor et al., 2021; Farine et al., 2015). It is therefore unsurprising that the study of social behaviour and its consequences has become a major topic of research. However, while technological advances, from high-resolution GPS (He et al., 2023) to the analysis of drone footage with

computer vision and deep learning (Koger et al., 2023), are rapidly facilitating our capacity to collect data on social behaviour, our ability to draw inferences from social data remains plagued by problems.

A number of issues arise when confronting animal social network data with statistical models. Two particularly pressing and inter-related problems are the need to deal with a multitude of factors that can confound measures of the social environment and the fact that existing methods that can deal with confounds do not provide reliable estimates of effect sizes. Specifically, a wide range of factors can affect estimates of social connections among individuals, including methodological (e.g. the number of observations) and some biological (e.g. weather conditions), with consequences

for within- or between-population contrasts (Ogino, Maldonado-Chaparro, et al., 2023). These have long been dealt with by using permutation tests—randomisation procedures that swap elements in the data—to generate a distribution of expected values against which a network-based metric can be compared and a  $p$ -value calculated. However, while permutation tests can be very effective for dealing with confounding factors (Farine & Carter, 2022; Puga-Gonzalez et al., 2021), they come at the cost of being unable to generate estimates of corrected effect sizes beyond those of the fitting model itself. Further, while some forms of permutation tests can account for uncertainty in the data collection when generating  $p$ -values (Farine, 2017), these estimates of uncertainty are not explicitly carried through in a reportable form.

Ross et al. (2023) highlight the overlooked potential for generative network models as a tool for hypothesis testing in animal social network analysis. As the name suggests, generative network models involve producing random networks with some guiding rules that represent the hypothesis of interest. In doing so, they can evaluate the relative importance of these rules (or hypothesised effects) for reproducing the main structural features of the observed network. In principle, these rules can include both the biological factors of interest (e.g. whether females are more strongly connected in the network than males) while also capturing the process of data collection (e.g. how often each individual was observed). Ross et al. (2023) then present the R package *STRAND* that implements generative models. Notably, they build their implementation in a Bayesian framework, which allows dyadic estimates of edge weights to be represented by distribution and avoids the need to reduce these down to a single number. This much-needed package comes almost 10 years after the first calls to use Bayesian methods to carry uncertainty in the estimates of edge weights through to the construction of the network itself (Farine & Strandburg-Peshkin, 2015; Nightingale et al., 2015). Such functionality is important because networks otherwise cannot encode contextual information alongside the value of edges. For example, two individuals could have an edge weight of 0 because they avoid one another or because some sampling artefacts mean that they never had the opportunity to be observed together. These two types of zeros represent the outcome of very different (biological and methodological) processes. Thus, instead of having a single zero, the Bayesian approach can represent likely true zeros as those with low uncertainty and potential false zeros as those with high uncertainty.

The first step in developing or refining analytical methods is to confirm that, under expected conditions, they perform to the same or better standard than existing methods. Using simulations, Redhead et al. compared the performance of Bayesian generative modelling (implemented in *STRAND*) with long-standing permutation-based methods (MRQAP, using node permutations implemented in the R package *asnipe*, Farine, 2013). They find that, in the absence of confounding effects, the results are largely comparable across different approaches, with *STRAND* being more sensitive to detecting small effect sizes. This is generally good news on two fronts. First,

the results of Ross et al. (2023) provide the platform for a more in-depth examination of the performance of different models under the conditions we expect from field data. Second, as in the analysis of many other types of data, we can now benefit from a plurality of approaches, each with their own strengths (e.g. sensitivity, provision of effect sizes) and limitations (e.g. the assumptions that they violate, such as non-independence, and sampling equality; see Weiss et al., 2021 for more discussion).

While the *STRAND* package provides important new functionality in social network analysis, major hurdles remain. The current simulations used by Ross et al. (2023) largely conform to the expectations of the data from the perspective of the model. Yet, confounding factors in animal social networks are likely to universally violate these expectations, even when individuals are studied in controlled conditions and extremely well sampled (Ogino, Maldonado-Chaparro, et al., 2023). As noted by the authors, *STRAND* (and the similar package *BISON*, Hart et al., 2023) have only been implemented and tested with minimal functionality. They therefore require that the network is well (or equally importantly, evenly) sampled and free of other, non-social drivers of structure (e.g. non-overlapping home ranges). While, in theory, these functionality in *STRAND* can be extended to also include confounding factors, doing so would need to be specific to each study (Ross et al., 2023). Increasing proficiency among researchers in using these methods will make this an appealing solution for some studies, and explicitly modelling effects such as spatial behaviours will also be of substantial interest (Webber et al., 2023). However, it is unlikely that modelling frameworks such as *STRAND* will be suitable 'out-of-the-box' for many studies.

As with many methods, developing a new tool is just the start, and further studies are needed to understand the strengths and weaknesses of any approach. One interesting question is whether the models implemented by *STRAND* can overcome some existing biases. For example, Redhead et al. highlight the study of reciprocity in cooperation networks when demonstrating *STRAND*'s functionality. The hypothesis that reciprocity is a driver of cooperation is often contrasted with relatedness as an alternative hypothesis. However, Carter et al. (2019) found that models that included both were generally biased towards detecting an effect of relatedness. This is because reciprocity networks are usually incomplete (e.g. we can observe A feed B but might miss B feeding A and rarely have the full history of interactions). By contrast, genetic data are always complete across all dyads containing genotyped individuals. A good test of the power of generative models will be to determine whether they are robust when incomplete sampling affects some networks (e.g. reciprocity) but not others (e.g. relatedness).

Beyond advancing the state of our statistical tools, we can also ride the technological wave as a means of improving our ability to generate new knowledge. One inherent limitation of the network approach is that it collapses a large amount of data into a single network. While the Bayesian approach implemented by *STRAND* and some permutation tests can make use of the underlying data, they do

not give us repeated measures. However, novel tracking techniques can generate much more finely resolved data on the patterns of contact and interactions among individual animals. Tracking groups of animals using markers (Alarcón-Nieto et al., 2018), deep learning (Ferreira et al., 2020; Schofield et al., 2023) and GPS (He et al., 2023) can provide an almost continuous record of the social behaviour of animals. From these data, we can generate replicated—daily (Ogino, Maldonado-Chaparro, et al., 2023), weekly (Aplin et al., 2015) or monthly (Ogino, Strauss, et al., 2023)—snapshots of networks. Future work should look at analytical workflows that can detect and model consistent confounds (e.g. weather, resource availability) and biases (e.g. personality, tag performance; see He et al. 2023) that arise within and across network replicates. A key question to answer is whether using more replicates of networks constructed with fewer data (i.e. more uncertainty in each network measure but more repeated measurements) versus having fewer networks constructed using more data (or one network from the whole dataset, producing just one measurement for each individual) provides the most robust inference. Such questions are ripe for examination.

Finally, we should avoid succumbing to the risk of having the available statistical tools dictate the science that we do. All models are flawed, even if some are useful. Current knowledge, including the simulations provided by Ross et al. (2023), suggests that—with some notable exceptions (Weiss et al., 2021)—results should not be particularly sensitive to the choice of method, as long as the data are collected relatively evenly across individuals, that confounding effects are removed from the data and that the biological effect is of a meaningful size. It is also worth remembering that many questions about social behaviour can be answered without needing a network approach, especially when the data are highly resolved in time, by conducting event-based analyses (e.g. Cantor et al., 2023). Such analyses can avoid many of the issues that arise from pooling data within individuals and across dyads, making it easier to account for, and model, multiple different drivers and consequences of social behaviour.

## CONFLICT OF INTEREST STATEMENT

The author declares no conflicts of interest.

## DATA AVAILABILITY STATEMENT

Data have not been archived because this article does not use data.

## ORCID

Damien R. Farine  <https://orcid.org/0000-0003-2208-7613>

## REFERENCES

- Alarcón-Nieto, G., Graving, J. M., Klarevas-Irby, J. A., Maldonado-Chaparro, A. A., Mueller, I., & Farine, D. R. (2018). An automated barcode tracking system for behavioural studies in birds. *Methods in Ecology and Evolution*, 9, 1536–1547.
- Aplin, L. M., Firth, J. A., Farine, D. R., Voelkl, B., Crates, R. A., Culina, A., Garroway, C. J., Hinde, C. A., Kidd, L. R., & Psorakis, I. (2015). Consistent individual differences in the social phenotypes of wild great tits, *Parus major*. *Animal Behaviour*, 108, 117–127.
- Cantor, M., Farine, D. R., & Daura-Jorge, F. G. (2023). Foraging synchrony drives resilience in human–dolphin mutualism. *Proceedings of the National Academy of Sciences of the United States of America*, 120, e2207739120.
- Cantor, M., Maldonado-Chaparro, A. A., Beck, K. B., Brandl, H. B., Carter, G. G., He, P., Hillemann, F., Klarevas-Irby, J. A., Ogino, M., & Papageorgiou, D. (2021). The importance of individual-to-society feedbacks in animal ecology and evolution. *Journal of Animal Ecology*, 90, 27–44.
- Carter, G. G., Schino, G., & Farine, D. (2019). Challenges in assessing the roles of nepotism and reciprocity in cooperation networks. *Animal Behaviour*, 150, 255–271.
- Farine, D. R. (2013). Animal social network inference and permutations for ecologists in R using *asnipe*. *Methods in Ecology and Evolution*, 4, 1187–1194.
- Farine, D. R. (2017). A guide to null models for animal social network analysis. *Methods in Ecology and Evolution*, 8, 1309–1320.
- Farine, D. R., & Carter, G. G. (2022). Permutation tests for hypothesis testing with animal social network data: Problems and potential solutions. *Methods in Ecology and Evolution*, 13, 144–156.
- Farine, D. R., Montiglio, P.-O., & Spiegel, O. (2015). From individuals to groups and back: The evolutionary implications of group phenotypic composition. *Trends in Ecology & Evolution*, 30, 609–621.
- Farine, D. R., & Strandburg-Peshkin, A. (2015). Estimating uncertainty and reliability of social network data using Bayesian inference. *Royal Society Open Science*, 2, 150367.
- Ferreira, A. C., Silva, L. R., Renna, F., Brandl, H. B., Renoult, J. P., Farine, D. R., Covas, R., & Doutrelant, C. (2020). Deep learning-based methods for individual recognition in small birds. *Methods in Ecology and Evolution*, 11, 1072–1085.
- Hart, J., Weiss, M. N., Franks, D., & Brent, L. (2023). Bison: A Bayesian framework for inference of social networks. *Methods in Ecology and Evolution*, 14, 2411–2420.
- He, P., Klarevas-Irby, J. A., Papageorgiou, D., Christensen, C., Strauss, E. D., & Farine, D. R. (2023). A guide to sampling design for GPS-based studies of animal societies. *Methods in Ecology and Evolution*, 14, 1887–1905.
- Koger, B., Deshpande, A., Kerby, J. T., Graving, J. M., Costelloe, B. R., & Couzin, I. D. (2023). Quantifying the movement, behaviour and environmental context of group-living animals using drones and computer vision. *Journal of Animal Ecology*, 92, 1357–1371.
- Nightingale, G., Boogert, N. J., Laland, K. N., & Hoppitt, W. (2015). Quantifying diffusion in social networks: A Bayesian approach. In J. Krause, R. James, D.W. Franks & D.P. Croft (Eds.), *Animal Social Networks* (pp.38–52). Oxford University Press.
- Ogino, M., Maldonado-Chaparro, A. A., Aplin, L. M., & Farine, D. R. (2023). Group-level differences in social network structure remain repeatable after accounting for environmental drivers. *Royal Society Open Science*, 10, 230340.
- Ogino, M., Strauss, E. D., & Farine, D. R. (2023). Challenges of mismatching timescales in longitudinal studies of collective behaviour. *Philosophical Transactions of the Royal Society B*, 378, 20220064.
- Puga-Gonzalez, I., Sueur, C., & Sosa, S. (2021). Null models for animal social network analysis and data collected via focal sampling: Pre-network or node network permutation? *Methods in Ecology and Evolution*, 12, 22–32.
- Ross, C. T., McElreath, R., & Redhead, D. (2023). Modelling animal network data in R using STRAND. *Journal of Animal Ecology*. <https://doi.org/10.1111/1365-2656.14021>
- Schofield, D. P., Albery, G. F., Firth, J. A., Mielke, A., Hayashi, M., Matsuzawa, T., Biro, D., & Carvalho, S. (2023). Automated face recognition using deep neural networks produces robust primate social networks and sociality measures. *Methods in Ecology and Evolution*, 14, 1937–1951.
- Ward, A., & Webster, M. (2016). *Sociality: The behaviour of group-living animals*. Springer.

- Webber, Q. M., Albery, G. F., Farine, D. R., Pinter-Wollman, N., Sharma, N., Spiegel, O., Vander Wal, E., & Manlove, K. (2023). Behavioural ecology at the spatial-social interface. *Biological Reviews*, *98*, 868–886.
- Weiss, M. N., Franks, D. W., Brent, L. J., Ellis, S., Silk, M. J., & Croft, D. P. (2021). Common datastream permutations of animal social network data are not appropriate for hypothesis testing using regression models. *Methods in Ecology and Evolution*, *12*, 255–265.