ANIMAL SOCIAL NETWORKS

Methods in Ecology and Evolution | BRITISH ECOLOGICAL SOCIETY

Null models for animal social network analysis and data collected via focal sampling: Pre-network or node network permutation?

Ivan Puga-Gonzalez[1,2] (ID) | Cédric Sueur[3,4] (ID) | Sebastian Sosa[3] (ID)

[1]Institute for Global Development and Planning, University of Agder, Kristiansand, Norway

[2]Center for Modeling Social Systems at NORCE, Kristiansand, Norway

[3]Université de Strasbourg, CNRS, IPHC, UMR 7178, Strasbourg, France

[4]Institut Universitaire de France, Paris, France

Correspondence
Ivan Puga-Gonzalez
Email: ivanpuga@gmail.com

Handling Editor: Carolyn Kurle

Abstract

1. In social networks analysis, two different approaches have predominated in creating null models for hypothesis testing, namely pre-network and node network permutation approaches. Although the pre-network permutation approach appears more advantageous, its use has mainly been restricted to data on associations and sampling methods such as 'group follows'.

2. The pre-network permutation approach has recently been adapted to data on interactions and the focal sampling method, but its performance in different scenarios has not been thoroughly explored. Here, we assessed the performance of the pre-network and node network permutation approach in several simulated scenarios based on proneness to false positive or false negatives and with or without observation bias.

3. Our results showed that the pre-network permutation was sensitive to false positives in scenarios with or without observation bias. The node network permutation approach produced fewer false positives and negatives than the pre-network approach, but only in scenarios without observation bias. In scenarios with observation bias, the node network permutation approach was outperformed by pre-network permutation.

4. Caution should be taken when using the pre-network and node network permutations to create null models with data collected via focal sampling. This study provides future methodological research perspectives for social network analyses.

KEYWORDS
group-living, null models, permutation tests, simulations, social network analysis

1 | INTRODUCTION

Social network analysis (SNA) has become a standard toolbox to describe and quantify the social structure of group-living animals, identify the mechanisms driving group-level properties and elucidate how group-level properties may, in turn, feed back into individuals' social behaviour and interactions with their environment (Croft, Madden, Franks, & James, 2011; Farine & Whitehead, 2015; Fisher & McAdam, 2017; Sosa, 2018; Sueur, Romano, Sosa, & Puga-Gonzalez, 2019). For instance, SNA has shown that variation in the social phenotype or behaviour of individuals may affect overall group structure (Aplin et al., 2013; Croft et al., 2009), has provided

evidence that individuals' phenotypes are correlated to their position in the social network (Firth et al., 2018; Firth, Sheldon, & Brent, 2017; Sih, Spiegel, Godfrey, Leu, & Bull, 2018; Sosa, 2016) and has revealed that this position may ultimately affect the fitness of individuals (Formica et al., 2012; Silk, 2007; Stanton & Mann, 2012). SNA has thus become an essential tool for researchers in evolutionary biology and behavioural ecology (Croft et al., 2011; Farine & Whitehead, 2015).

One of the main challenges of SNA is the testing of hypotheses. The main problem is that the data represented in social networks are not independent (Croft et al., 2011; Farine & Whitehead, 2015). This non-independence of the data forbids the application of conventional parametric statistical methods (unable to deal with non-independence) to test the hypotheses (but see Cranmer, Leifeld, McClurg, & Rolfe, 2017). One of the solutions proposed to circumvent this problem is network permutation (randomization of data; Manly, 1995; Whitehead & Dufault, 1999). Network permutation creates a large number of randomly generated networks by shuffling the original data while keeping certain features of the original data set constant (e.g. number of observations per individual). This creates a null distribution of values for the statistic of interest (i.e. a null model) against which the measured statistic or metric can be compared and its significance value calculated (Farine, 2017).

Two of the most commonly used permutation methods to build null models are node permutations and pre-network permutations. Pre-network permutations were first developed by Bejder, Fletcher, and Brager (1998) as an extension of a method developed by Manly (1995) to test for the co-occurrence of species on islands. This method was initially used by researchers studying whether indices of associations among pairs of individuals were purely random or not in a social population (Bejder et al., 1998; Whitehead, 1999). Data on associations are often captured via the gambit of the group, which assumes that all individuals observed within a group at a given location and at a certain point in time are associated (Whitehead & Dufault, 1999). Associations are not therefore direct interactions between two individuals but are rather co-occurrences of individuals in the same group. The data collected are then represented in a group-by-individual matrix, where rows represent the subgroups or clusters of individuals observed and columns represent the individuals; values of 1 or 0 in each cell of the matrix indicate whether the individual (column) was observed or absent in that specific subgroup (row). To test whether association indices appear by chance or not, pre-network permutations reshuffle the data in the group-by-individual matrix in such a way that the row and column totals remain unaffected (number of individuals per subgroup and number of observations per individual remain constant). At each time step, two individuals are swapped between subgroups, on the condition that individual A occurring in subgroup 1 but not in subgroup 2 is swapped to subgroup 2, and individual B occurring in subgroup 2 but not in subgroup 1 is then swapped to subgroup 1. After each permutation, the network is reconstructed and the statistic of interest is recalculated. One of the advantages of this method is that it can control swaps for different factors,

for example, location, and thus disentangle whether non-random associations are due to social or other factors (Whitehead, 1999).

The node network permutation approach is the other most commonly used method to test network-related hypotheses. In animal research, node permutations have mainly been used to compare two matrices (or networks) involving the same group of individuals. In this case, the values entered in the matrix cells are based on direct behavioural observations (e.g. grooming) and the tests have been widely used to test for reciprocity and behavioural interchange (Hemelrijk, 1990a, 1990b), especially in primate studies (Puga-Gonzalez, 2017). In contrast to the gambit of the group, direct observations are usually collected via focal sampling, scan sampling or ad libitum sampling (Altmann, 1974). The data are then entered in an $n \times n$ matrix (n = number of individuals) where rows are actors and columns receivers of a behaviour (e.g. aggression). These data are then used to calculate a specific node metric (e.g. degree). Node permutation is achieved by redistributing the identity of the nodes at each time step while keeping the node metric (value) constant. This makes it possible to test whether a specific network metric is associated with a specific node attribute (e.g. whether females groom more than males). One advantage of this method is that it is simple to implement since permutations are carried out in the adjacency matrix of the original network; however, unless additional constraints are added to the swap of the node labels (Pinter-Wollman et al., 2014), the test can only tell whether the network structure is different from a random configuration since it cannot control for other factors such as time or location (Farine & Whitehead, 2015).

The effectiveness of pre-network permutations on association data has been explored at length (Bejder et al., 1998; Farine, 2014; Farine & Whitehead, 2015; Sundaresan, Fischhoff, & Dushoff, 2009; Whitehead, 1999; Whitehead, Bejder, & Ottensmeyer, 2005). Few studies, however, have compared the effectiveness of pre-network and node network permutations. Farine (2014) used simulations to test the effectiveness of weighted associations in detecting phenotypic assortment under different sources of noise (e.g. sampling errors). Using pre-network and node network permutations, he showed that both approaches appeared to qualitatively yield the same results in all cases tested (n = 10) except one, in which the node permutation approach failed to reject the null hypothesis (type II error/false negative). More recently, Farine (2017) adapted the pre-network permutation approach to interaction data collected via focal sampling (Altmann, 1974). He used simulated data that mimicked focal sampling data collection, a female bias social phenotype (producing a higher average weighted degree among females than males) and an observation bias (females 20% less likely to be observed than males) to compare the ability of pre-network and node network permutations to correctly identify a difference between the weighted degree of females and males, therefore avoiding false negatives (type II errors). Only the pre-network permutation approach rejected the null hypothesis, correctly identifying a stronger social phenotype in females than males, despite females being observed less frequently than males (Farine, 2017). He concluded that the pre-network permutation approach, adapted to focal

TABLE 1 Range of variation of initial parameters

Parameter	Variable type	Range
(a) Group size	Discrete	[10–100]
(b) Female sex ratio	Continuous	[0.2–0.8]
(c) Female observation bias[a]	Continuous	[0.5–1.0]
(d) Number of focal samples	Discrete	[100–2,000]

[a]In simulations with certainty in observations this variable was kept constant at 1.

sampling data, was a better choice and recommended the use of this approach (Farine, 2017; Farine & Whitehead, 2015).

However, several factors were left unexplored in Farine's (2017) study. No attention was paid to the parameter space of the simulations: conclusions were based on one simulation, with a single group size, one observation bias value and no repetitions; the avoidance of false positives (type I errors) was not explored, and nor was the effect of other factors (e.g. sampling effort) that may potentially affect the performance of the permutation tests. This study uses simulations to make a thorough exploration of the parameter space and study the effect of two additional factors on the ability of pre-network and node network permutation tests to avoid false positives (type I) and false negatives (type II errors). Given that pre-network permutations have recently been adapted for this type of data collection (Farine, 2017), we focused on simulations mimicking focal sampling data collection. To make our results comparable, we used the same R code used in the study by Farine (2017), with some slight modifications (see methods). We explored the ability of the pre-network and node network permutation tests to avoid false positives (type I) and negatives (type II errors) under four different scenarios, namely two scenarios with no observation bias and with equal or different sex social phenotype (SSP); and two scenarios with observation bias and with equal or different SSP. These scenarios tested the robustness of the permutation approaches to false positives (equal SSP) and negatives (different SSP) with and without the presence of observation bias. In all, 500 simulations were run per scenario, and simulations varied in the value of four parameters: group size, sex ratio, number of samplings and degree of observation bias (Table 1). We hypothesized that both pre-network and node network permutations would perform equally well in scenarios without observation bias, and that pre-network permutations would outperform node network permutations in scenarios with observation bias.

2 | MATERIALS AND METHODS

2.1 | Simulation design

We followed the simulation approach described by Farine and Whitehead (2015) and Farine (2017). To generate the simulated data, we used a slightly modified version of the R code published by Farine (2017). The first modification corrected a small problem in the code that was creating a slightly higher observation bias than expected; the second modification allowed us to run simulations while automating the variation of the initial conditions (Table 1). To test the effect of observation bias, simulations were run with an observation bias (where females had a lower probability than males of being observed); or without an observation bias (where males and females were equally likely to be observed). In the wild, for instance, observation bias may occur in species where males have brighter colours or ornaments than females or where bold individuals are more active than their shy counterparts; in these cases, individuals with the former attribute are more easily observed than those displaying the latter (Klaich, Kinas, Pedraza, Coscarella, & Crespo, 2011). In our simulations, observation bias consisted of deliberately overlooking females during samplings and only recording their presence in a percentage of them (range [50%–100%], Table 1). In simulations with no observation bias, all males and females were recorded in samplings.

To test the sensitivity of the permutation approaches to false positives (type I error) and false negatives (type II error), simulation scenarios differed in the social phenotype displayed by females and males. In one type of scenario, both males and females were equally social and thus had an equally weighted degree. In this scenario, if the permutation approach detected a significant difference in weighted degree between the sexes, it would be erroneously rejecting the null hypothesis (type I error). In the other type of simulation scenarios, females were more social than males and thus had a higher weighted degree than males. In this scenario, the permutation approach must detect a significant difference between the sexes; failing to do so would be a false negative (type II error). The difference in social phenotype was generated by allocating the females to larger subgroups and the males to smaller ones during the focal samplings. When no difference was present, males and females were equally likely to be in any given subgroup. We also investigated the effect of three socio-demographic factors on the ability of the statistical test to avoid false positives and false negatives, namely group size, sex ratio and sampling effort (number of focal samples); their range of variation is shown in Table 1. The data collected from the simulations were analysed through either a pre-network or a node network permutation procedure. Pre-network and node network permutations were carried out using the same R code published by Farine (2017) with a slight modification to correct for the way the swapping of individuals occurred between focal samples and the way females were assigned to subgroups. See Supporting Information for a more detailed description of the simulation, the modifications to Farine's (2017) R code, and the overall R code used to generate the simulated data.

2.2 | Parameters, data collection and statistical analysis

Four different parameters were varied for each simulation scenario, namely group size, female sex ratio, female observation bias

and number of focal samples (Table 1). In the scenarios with no observation bias, female observation bias was kept constant at 1. Note that the lower the value of observation bias, the higher the likelihood of 'overlooking' females will be. We sampled the parameter space (variables a–d in Table 1) using Latin hypercube sampling (Stein, 1987) with the 'lhs' R library (Carnell, 2018). Five hundred different combinations of input parameter values were run per simulation scenario, that is, a total of 2,000 simulations. From the *observed* data per simulation, we constructed social networks using simple ratio index (Cairns & Schwager, 1987; Whitehead & Dufault, 1999) and calculated the weighted degree of all individuals in the network. We then ran a linear model (weighted degree ~sex) and obtained significance values using two different network permutation methods: pre-network and node network permutation. Significant values were set at 0.05, were two-tailed and were estimated by comparing our 'observed' statistical metric (the β estimate of the sex factor in the linear model) to the null distribution created from 1,000 permutations. Because the significance level (α) was set to 0.05, we expected a rate of false positives of ~5% (i.e. ~25 cases out of 500). It was impossible to calculate the expected rate of false negatives because this rate is conditional on the value of α (0.05) and the values of μ (mean), σ (SD) and n (group size), all of which are simulation specific. We therefore reported the percentage of false negatives found in each set of 500 simulations. Linear models met the assumptions of normality, homoscedasticity and independence of residuals.

By categorizing simulations into those with no difference in SSP and those with different SSP (females stronger than males), we were able to discern between type I and type II errors, respectively. When the SSP is equal between the sexes, the statistical tests should find no difference in weighted degree between sexes; if found, this result is a false positive (type I error). On the other hand, when SSP is stronger among females and the statistical test fails to reject the null hypothesis, this result is a false negative (type II error). By categorizing simulations into with/without observation bias, we tested the influence of 'overlooking' individuals on the robustness of the statistical tests. Finally, we assessed the effect of each parameter on the likelihood of type I or II error by running logistic regression models in which the presence of false positives/negatives was the response variable, and the input parameters (a–d in Table 1) were the predictors. This made it possible to

assess which factors were more likely to drive false positives/negatives. Logistic regression models were checked for overdispersion by calculating the ratio of residual deviance to degrees of freedom. In all cases, the ratio was ~1 (no overdispersion). All simulations and statistical analysis were carried out in R, version 3.5.2 (R Core Team, 2018).

3 | RESULTS

3.1 | No observation bias and no difference in social phenotype between sexes (false positives)

The pre-network and node network permutation detected a significant ($p < 0.05$) difference between the weighted degree of the sexes in 37% of cases (185/500) and 5.6% of cases (28/500), respectively. The pre-network permutation procedure therefore had a high rate of false positives (type I error), whereas the node network permutation procedure had an expected rate, that is, ~5%. Figure S1 shows that as expected, the difference in the median degree between males and females appears normally distributed around 0 when there is no difference in social phenotype between the sexes (Figure S1). The logistic regression model showed that for the pre-network permutation procedure, the likelihood of false positives decreased with decreasing values of group size and increased with increasing number of focal samples (Table 2; Figure S2). These results, however, should be taken with caution since only 6.8% of the variance was accounted for by these factors (Nagelkerke pseudo-R^2 index, Table 2). When the data were analysed with results of the node permutation procedure, the likelihood of false positives increased with increasing number of focal samples (Figure S2). Note that the model accounted for only 4.4% of the variance (Table 2) and that the rate of false positives (5.6%) was close to what was expected by chance (5%); these results should therefore be taken with caution.

3.2 | No observation bias and females with stronger social phenotype (false negatives)

Both procedures had a low rate of false negatives, that is, 9.4% (47/500) and 3.2% (16/500) for the pre-network and node network

TABLE 2 Logistic regression models according to (A) pre-network or (B) node network permutation procedure

	(A) Pre-network permutation				(B) Node network permutation			
	Estimate	SE	z	p	Estimate	SE	z	p
Intercept	−0.5907	0.3994	−1.479	0.139	−4.6971	0.9077	−5.175	**<0.001**
Group size	−0.0156	0.0037	−4.205	**<0.001**	0.0057	0.0076	0.747	0.455
Sex ratio	0.9994	0.5528	1.808	0.071	0.8624	1.1535	0.748	0.455
Num focals	0.0004	0.0002	2.134	**0.033**	0.0009	0.0004	2.492	**0.013**
Nagelkerke R^2	0.0688				0.0436			

*Note:* The dependent variable is the presence (1) or absence (0) of false positives (Type I error).
Bold significant *p*-values (<0.05).

TABLE 3 Logistic regression models according to (A) pre-network or (B) node network permutation procedure

	(A) Pre-network permutation				(B) Node network permutation			
	Estimate	SE	z	p	Estimate	SE	z	p
Intercept	1.4599	0.6352	2.298	**0.022**	8.5789	2.8050	3.058	**0.002**
Group size	−0.0231	0.0067	−3.458	**0.001**	−0.6092	0.1582	−3.851	**<0.001**
Sex ratio	−3.2976	0.9889	−3.334	**0.001**	1.2810	2.1294	0.602	0.547
Num focals	−0.0011	0.0003	−3.662	**<0.001**	−0.0015	0.0008	−2.036	**0.042**
Nagelkerke R^2	0.1611				0.7001			

Note: The dependent variable is the presence (1) or absence (0) of false negatives (Type II error).
Bold significant *p*-values (<0.05).
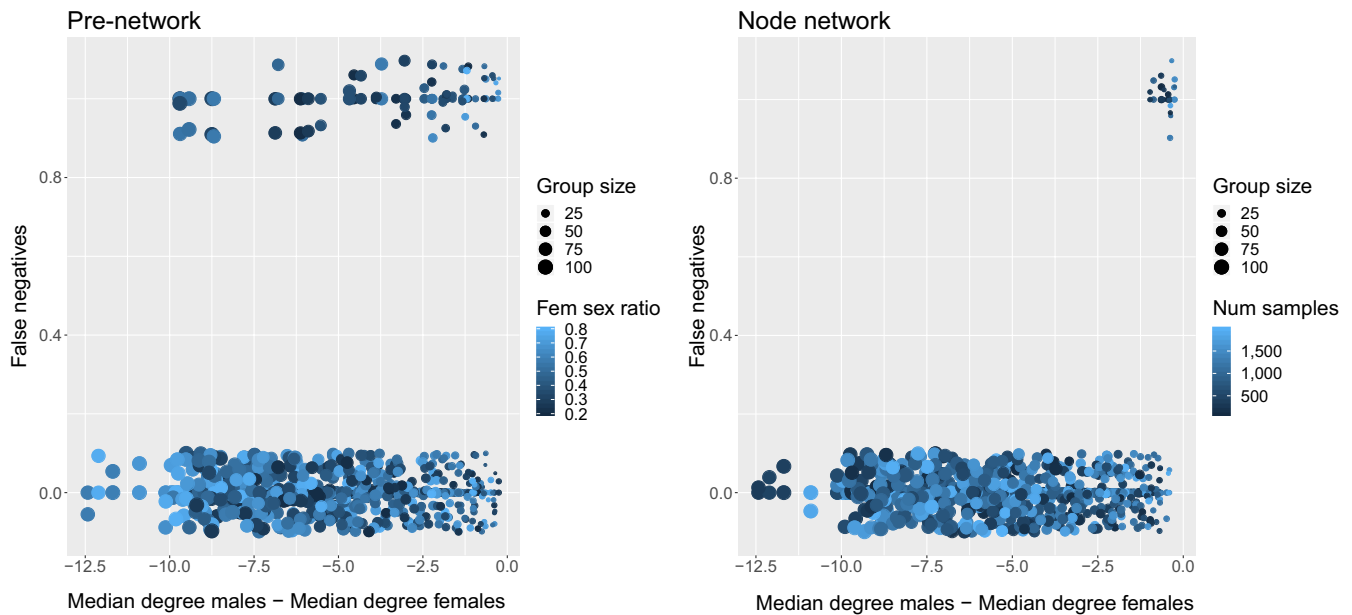
FIGURE 1 Four-dimensional plot of the presence (1) or absence (0) of false negatives (*y*-axis) according to the difference between males' weighted degree minus females' weighted degree (*x*-axis), group size (point size) and female sex ratio (point colour) for left panel (Pre-network); and number of samples (point colour) for right panel (Node network permutation procedure). Data points are jittered along the *y*-axis for optimal visualization

permutation procedure, respectively. The logistic regression model showed that, for the pre-network permutation procedure, the likelihood of false negatives decreased with increasing values of group size, sex ratio and number of focal samples (Table 3; Figure S3). The model, however, accounted for only 16.1% of the variance (Nagelkerke pseudo-R^2 index, Table 3). For the node network permutation procedure, the logistic regression model showed that the likelihood of false negatives decreased with increasing values of group size and number of focal samples (Table 3; Figure S3). This model explained 70% of the variance observed and group size had the biggest effect (Table 3; Figure S3).

Figure 1 presents the presence/absence of false negatives when data are plotted against the difference between the median degree of males minus that of females. As expected, in all simulations, females had a higher median weighted degree than males (all data points lay on the negative side of the *x*-axis, Figure 1). A clear pattern

can be observed in both panels: in the left panel, false positives appear to be driven by groups with a small size and a low female sex ratio (Figure 1; Figure S3). In the right panel, on the other hand, false negatives appear to be mainly driven by small group size and a low number of focal samples (Figure 1; Figure S3).

## 3.3 | Observation bias and no difference in social phenotype between sexes (false positives)

The pre-network and node network permutation procedures detected significant differences (*p* < 0.05) in 35.6% (178/500) and 60.8% of cases (304/500), respectively. Although both procedures had a high rate of false positives, it was substantially higher in the node network permutation. The logistic regression model showed that the likelihood of false positives decreased with increasing values

TABLE 4 Logistic regression models according to (A) pre-network or (B) node network permutation procedure

	(A) Pre-network permutation				(B) Node network permutation			
	Estimate	SE	z	p	Estimate	SE	z	p
Intercept	−0.0250	0.6337	−0.040	0.968	10.9000	1.1130	9.792	**<0.001**
Group size	−0.0150	0.0037	−4.028	**<0.001**	0.0290	0.0056	4.097	**<0.001**
Female obs Bias	−0.0256	0.6633	−0.039	0.969	−17.0200	1.4600	−11.657	**<0.001**
Sex ratio	−0.2028	0.5520	−0.367	0.713	1.9550	0.7866	2.485	**0.013**
Num focals	0.0003	0.0002	1.935	0.053	0.0005	0.0003	2.204	**0.028**
Nagelkerke R^2	0.0555				0.6326			

Note: The dependent variable is the presence (1) or absence (0) of false positives (Type I error).
Bold significant *p*-values (<0.05).
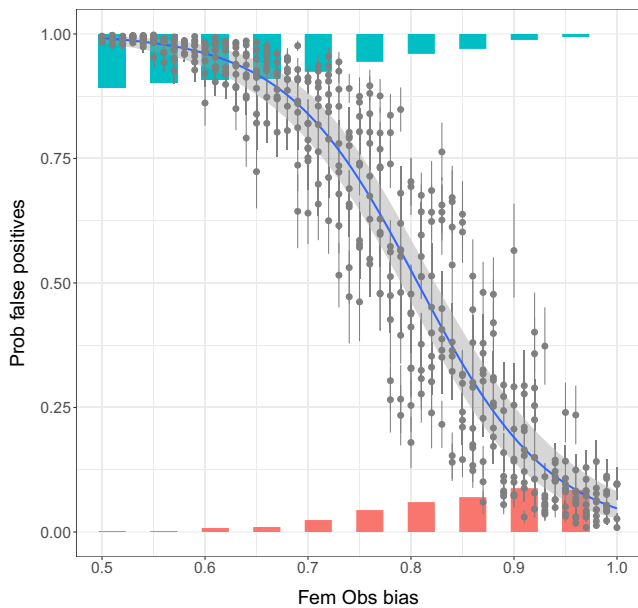
FIGURE 2 Predicted probability of false positives according to the logistic regression model obtained with data of the node network permutation procedure. Each data point represents the probability (±*SE*) according to each combination of parameter values (*n* = 500). The *x*-axis shows the value of the parameter with the highest effect, female observation bias. Histograms represent the 'observed' frequency of the presence (1, blue) and absence (0, pink) of false positives

of group size for the pre-network permutation procedure (Table 4; Figure S4). However, the model accounted for only 5.5% of variance (Nagelkerke pseudo-R^2 index, Table 4). For the node network permutation procedure, the logistic regression model showed that the likelihood of false positives increased with increasing likelihood of overlooking females (lower values of female observation bias) and increasing values for group size, sex ratio and number of samples (Table 4; Figure 2; Figure S5). This model explained 63.26% of the variance observed, with female observation bias having the biggest effect (Nagelkerke pseudo-R^2 index, Table 4; Figure 2).
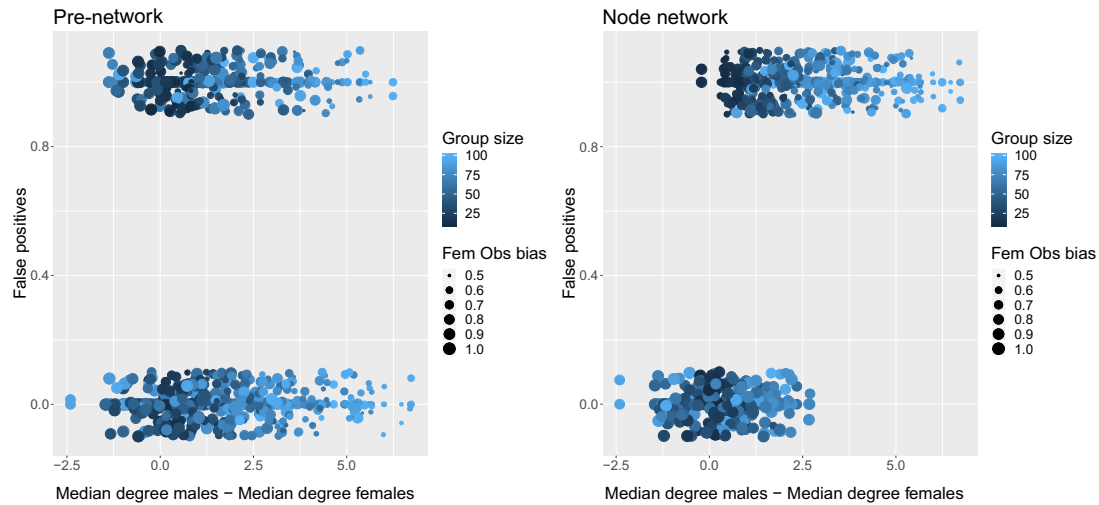
Figure 3 presents the presence/absence of false positives when data are plotted against the difference between the median weighted degree of males minus that of females. As expected, the lower the

female observation bias (size of data points in Figure 3), the higher the overlooking of females will be, and thus the higher the median weighted degree of males will be compared to that of females (data points lying on the positive side of the *x*-axis). No clear pattern can be observed for the pre-network permutation procedure (Figure 3, left). A clear pattern can be observed for the node network permutation procedure (Figure 3, right); the lower the female observation bias (data point size) and the bigger the group size is (lighter colour), the more likely it is that the procedure will incorrectly reject the null hypothesis.

3.4 | Observation bias and a stronger social phenotype in females (false negatives)

The pre-network and node network permutation procedure failed to detect significant differences (*p* > 0.05) in 12.6% (63/500) and 36.6% of cases (183/500), respectively. The pre-network procedure thus clearly outperformed the node network permutation procedure. The logistic regression model showed that, for the pre-network permutation procedure, the likelihood of false negatives decreased as the values of group size, female observation bias and female sex ratio increased (Table 5; Figure S6). The model, however, accounted for only 7.8% of variance (Nagelkerke pseudo-R^2 index, Table 5). For the node network permutation procedure, the logistic regression model showed that the likelihood of false negatives positive increased with the increasing likelihood of overlooking females (lower values of female observation bias) and decreased as the values of group size, female sex ratio and number of focal samples increased (Table 5; Figures 4 and 5; Figure S7). This model explained 69.1% of variance observed, with the largest effect attributable to female observation bias (Nagelkerke pseudo-R^2 index, Table 5; Figure 4).

Figure 5 presents the presence/absence of false negatives when data are plotted against the difference between the median degree of males minus that of females. As expected, females had a higher median weighted degree than males (most data points lay on the negative *x*-axis, Figure 5), and the higher the overlooking of females (i.e. lower values of female observation bias, darker

FIGURE 3 A four-dimensional plot of the presence (1) and absence (0) of false positives (*y*-axis) according to the difference between males' weighted degree minus females' weighted degree (*x*-axis), females' observations bias (data points' size) and group size (data points' colour). Left, Pre-network and right, Node network permutation procedure. Data points are jittered along the *y*-axis for a better visualization

TABLE 5 Logistic regression models according to (A) pre-network or (B) node network permutation procedure

|	**(A) Pre-network permutation**				**(B) Node network permutation**			
	Estimate	*SE*	*z*	*p*	**Estimate**	*SE*	*z*	*p*
Intercept	1.7120	0.8687	1.971	**0.049**	17.5300	1.6820	10.422	**<0.001**
Group size	−0.0147	0.0054	−2.709	**0.007**	−0.0457	0.0066	−6.869	**<0.001**
Female obs Bias	−2.2119	0.9828	−2.251	**0.024**	−19.9800	1.8440	−10.839	**<0.001**
Sex ratio	−2.1933	0.8220	−2.668	**0.008**	−1.6620	0.8320	−1.998	**0.046**
Num focals	−0.0002	0.0003	−0.856	0.392	−0.0006	0.0003	−2.104	**0.035**
Nagelkerke R^2	0.0782				0.6914			

Note: The dependent variable is the presence (1) or absence (0) of false negatives (Type II error).
Bold significant *p*-values (<0.05).
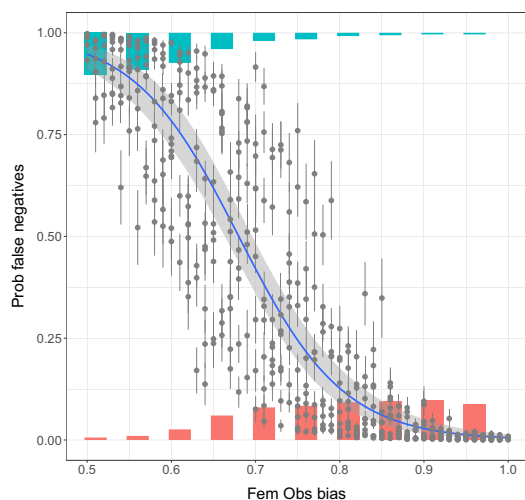
FIGURE 4 Predicted probability of false negatives according to the logistic regression model obtained with data of the node network permutation procedure. Each data point represents the probability (±*SE*) according to each combination of parameter values (*n* = 500). The *x*-axis shows the value of the parameter with the highest effect, that is, female observation bias. Histograms represent the 'observed' frequency of the presence (1, blue) or absence (0, pink) of false negatives

colour in Figure 5), the closer the data points are to 0 or positive on the *x*-axis (Figure 5). No clear pattern can be observed for the pre-network permutation procedure (Figure 5, left). A clear pattern can be observed for the node network permutation procedure; the procedure failed to reject the null hypothesis when there was a high likelihood of overlooking females. Note that in some simulations, despite females having a stronger social phenotype, the median degree of males was higher than that of females (points lying on the positive side of the *x*-axis of Figure 5) when the probability of overlooking females was very high (female observation bias ~0.5). In the 75 simulations where this occurred, the pre-network and node network permutation procedures detected a significant difference between the weighted degree of the sexes (albeit in the wrong direction, i.e. males > females) in 72% (54/75) and 5.3% (4/75) of cases, respectively. Hence, when the probability of overlooking females was very high, the pre-network procedure showed a strong tendency to erroneously detect a higher weighted degree among males than among females. When we recalculated the percentage of false negatives without including these 75 cases, we found that pre-network and node network permutation procedures failed to detect significant
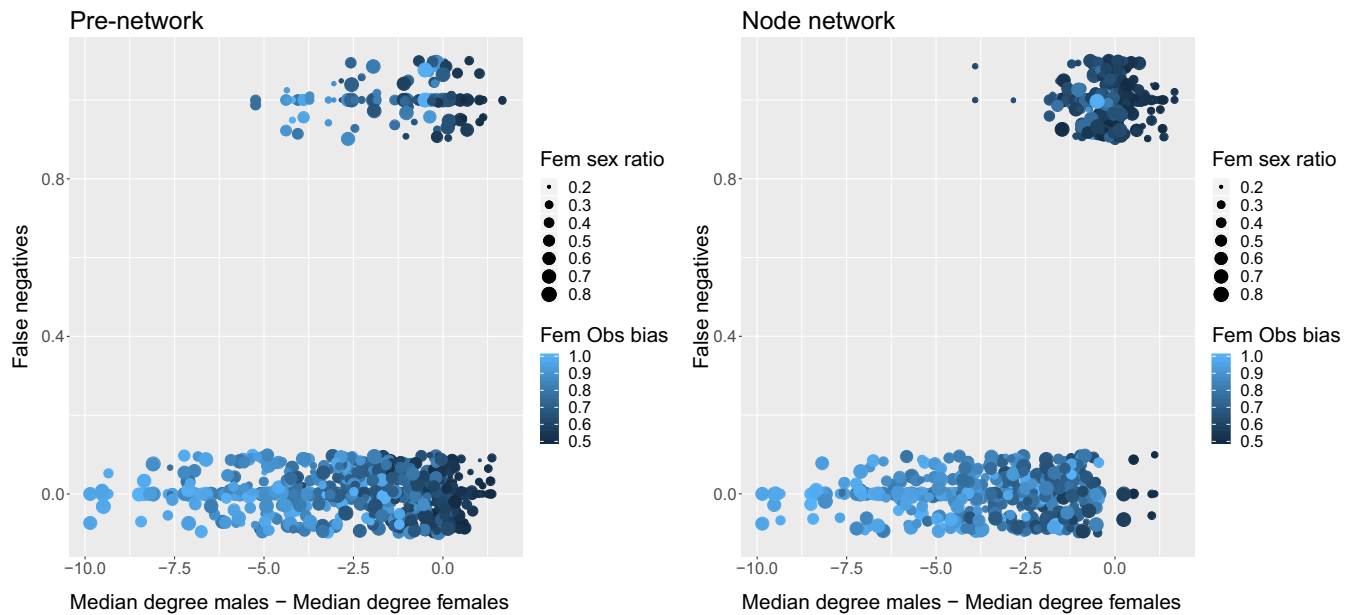
FIGURE 5 A four-dimensional plot of the presence (1) or absence (0) of false negatives (*y*-axis) according to the difference between males' weighted degree minus females' weighted degree (*x*-axis), female observation bias (different data point colours) and female sex ratio (different data point sizes). Left, Pre-network and right, Node network permutation procedure. Data points are jittered along the *y*-axis for optimal visualization

TABLE 6 Overview of the percentage of false positive/negatives according to scenarios with and without observation bias, different sex social phenotypes (SSP) and different permutation approaches (Pre-Net = pre-network or Node-net: node network)

	No observation bias[a]				Observation bias[b]			
	SSP equal False positives[c]		Fem SSP > Mal SSP False negatives[d]		SSP equal False positives[c]		Fem SSP > Mal SSP False negatives[d]	
Permutation approach	Pre-net	Node-net	Pre-net	Node-net	Pre-net	Node-net	Pre-net	Node-net
Percentage	37.0	5.6	9.4	3.2	35.6	60.8	12.6	36.6

[a]Females observed as often as males.

[b]Lower probability of observing females than males.

[c]Weighted degree of females is equal to that of males (no significant difference expected).

[d]Weighted degree of females is higher than that of males (significant difference expected).

differences in 9.88% (42/425) and 26.35% (112/425) of the remaining cases, respectively.

4 | DISCUSSION

Table 6 presents the summary of the rate of false positives/negatives according to the four scenarios. As expected from previous findings (Farine, 2017), the tests based on pre-network permutations performed better than those based on node network permutations in the scenario with observation bias and different sex social phenotype (SSP). Similarly, the pre-network permutation performed better in the scenario with observation bias but no difference in SSP; however, the rate of false positive was high for both permutation approaches in this case, with 35.6%, and 60.8% for the pre-network and node network permutation approaches, respectively. In both

scenarios with no observation bias, the node network permutation approach was more efficient (Table 6). However, whereas the rate of false negatives was low for both permutation procedures in the scenario with different SSP, the pre-network permutation procedure had a high rate of false positives, (37%) in the scenario with equal SSP. In sum, node network permutations were more efficient than pre-network permutations in scenarios with no observation bias. In scenarios with observation bias, however, node network permutations were outperformed by pre-network permutations. Pre-network permutations thus appeared more reliable in scenarios with observations bias but tended to detect spurious associations in the absence of observation bias.

It is difficult to understand why pre-network permutations are so sensitive to false positives. The logistic regression models did not provide a clear answer, since the percentage of variance explained by these models was very low [5.5%–6.8%] (Tables 2A and 4A). The

explanatory value of these factors is therefore low and should be taken with caution. Group size was the only constant factor across both scenarios, with the likelihood of false positives decreasing as the size of the group increased. However, even when group size was 100 individuals, the probability of false positives remained high ~20% (Figures S2 and S4). With regards to the likelihood of false negatives, the explanatory value of the logistic regression models is somewhat redundant because the rate of false negatives in these cases was low (Table 6).

In a previous study, Sundaresan et al. (2009) showed that the pre-network permutation approach might be prone to a high rate of false positives when the swapping of individuals among subgroups is carried out incorrectly. The authors simulated scenarios where sampling occasions consisted of the simultaneous sighting of two subgroups of individuals. As subgroups were formed at random, no structure was expected in the association of individuals. However, when individuals were swapped among subgroups without controlling for sampling occasion, the statistical test incorrectly identified non-random associations. Non-random associations were never detected when swaps were restricted to subgroups within the same sampling occasion. The spurious association was created because subgroups within the same sampling occasion were not independent; pairs of individuals seen in one subgroup could not be seen in the other subgroup. This suggests that in our case, the high rate of false positives may be a problem related to the swapping of individuals between subgroups. There are, however, several fundamental differences between our simulation approach and that carried out by Sundaresan et al. (2009). First, our simulations mimicked the focal sampling method, whereas those in the study by Sundaresan et al. (2009) mimicked 'group follows'. Second, we calculated the average number of associates (or interactions) per individual rather than the associations between pairs of individuals, so the statistics tested are different. Third, and perhaps most importantly, each sampling occasion in our study consisted of just one subgroup that was independently drawn from the population. Indeed, when Sundaresan et al. (2009) drew one subgroup from the entire population per sampling occasion, the rate of false positives dropped to 9%, a value close to what is expected by chance. Thus, the way individuals are swapped between subgroups seems an unlikely explanation in our case. An alternative explanation may be an insufficient number of permutations (n = 1,000) used to create the null distribution. We note, however, that our results remain qualitatively the same if we run 10,000 permutations. It therefore appears that there is insufficient variation in the null distribution, but the reasons behind this lack of variation remain unknown.

In contrast, when using the node network permutation approach, the explanatory value of the logistic regression models for the likelihood of false positive and negatives was high for all scenarios except the one with no observation bias and no difference in SSP (Table 2B). In this case, however, the rate of false positives was close to what was expected by chance (Table 6). In the scenario with no observation bias and different SSP between the sexes, the logistic regression model accounted for 70% of variance and showed that increasing values of group size and number of focal

samplings decreased the likelihood of false negatives (Table 3B; Figure S3). For this scenario, the rate of false negatives was very low (3.2%); the explanatory value of the model may thus seem redundant. Nevertheless, the pattern was evident: when group size was very small, there was a possibility that false negatives would occur (Figure 1 right; Figure S3). This was because the variation in subgroup sizes decreased with the size of the group, thus resulting in differences in weighted degree between the sexes that are closer to zero (Figure 1 right). Sampling effort could be used to control for the possibility of false negatives; however, a considerable sampling effort would be required for a small group size. For instance, for a group size of 10 individuals, 2,000 focal samples (200 per individual) will reduce the probability of false negatives to only ~50% (Figure S8). However, the probability of false negatives decreases rapidly as group size increases. For groups of 14 individuals, 2,000 samples reduce the probability to ~8%; and for groups of 20, 100 samples are enough to reduce the probability to 4% (Figure S8). In the other two scenarios (with observation bias), the explanatory value of the logistic regression models was also high; models accounted for 63.1% and 69.1% of variance in the scenarios with/without SSP differences, respectively (Tables 4B and 5B). In these cases, the degree of observation bias had the greatest effect on the likelihood of false positives or negatives (Figures 2 and 4). Indeed, the percentage of false positives decreased from 60.8% to 23% when the percentage of females being 'overlooked' was ≤20% and to 10.4% when it was ≤10% (observation bias equal to 0.8 and 0.9, respectively). Similarly, the percentage of false negatives reduced from 36.6% to 5.3% when the percentage of females being 'overlooked' was ≤20% and to 4.7% when it was ≤10% (observation bias equal to 0.8 and 0.9, respectively). Thus, a high certainty in the number of observations of individuals appears to guarantee the good performance of the node network permutation approach.

It is difficult to envision how controlling the other factors (group size, female sex ratio and sampling effort) could enhance the performance of the node permutation approach because these factors have opposite effects on the probability of false positives and negatives (Tables 4B and 5B). Increasing values of group size decrease the probability of false positives but increase the probability of false negatives, and the same effect occurs with values of female sex ratio and sampling effort (Figure S9). It thus seems redundant to control for these factors without having a clear expectation as to whether the sexes should (or not) differ in a given metric. An alternative solution may be to carry out meta-analyses. Voelkl, Vogt, Sena, and Wurbel (2018) showed that in preclinical animal research, standardization of studies (following the same protocol or group) is a major cause of poor reproducibility (predicting false positives or negatives), and that the inclusion of more representative study samples is required to improve the external validity and reproducibility in this domain. This means that groups with slightly different group sizes and sex ratios must be studied to assess whether the results are comparable and if the latter include false positives or false negatives.

So, which permutation approach is best to perform statistical analyses when using the focal sampling method? To answer this question, researchers should consider the sociality of their study species. Societies with fluid group membership (fission–fusion societies, troops with subgroup units or animal aggregations) may consist of dozens of individuals that are not individually recognized, where group size is estimated rather than known, and where groups frequently split into smaller ones. In this case, it may be necessary to control for demographic factors such as habitat selection or migration. Observation bias may be high in these cases due to the high number of individuals and/or because individuals are not individually recognized. Hence, the pre-network permutation approach should be preferred. Societies with stable group membership, on the other hand, consist of groups of known size that mainly remain cohesive, and are usually composed of individually recognized group members. In this case, it is fair to assume that preferential associations are due to the social motives of individuals (rather than demographic factors) and that observation bias is low (because individuals can be monitored and identified most of the time). In these cases, node network permutations should be preferred. Finally, there are cases in which data collection involves directional behavioural data (e.g. grooming that is given or received) and where researchers are interested in patterns of reciprocity and behaviour exchange (Puga-Gonzalez, 2017). The only possible approach for these cases is the node network permutation, as no study to date has adapted the pre-network approach to test this kind of behavioural patterns. Researchers should thus be cautious when using the node network permutation approach for this purpose and should be aware of the pitfalls highlighted here.

The pre-network and node network permutation approaches are two of the most commonly used statistical methods when dealing with the data dependency structure of socio-behavioural data. The pre-network permutation approach has been improved over the last decades, and has been shown to work well with association data. In contrast, when applied to behavioural data collected via focal sampling, the pre-network approach appears to be prone to false positives, while the statistical power of the node network approach seems limited when there is a high observation bias. The need for new permutation methods is thus clearly necessary for sampling methods collecting behavioural data and when high observation bias is suspected. Although recent efforts have tried to adapt the pre-network permutation approach to these type of sampling methods, this study shows that this method may not always work well. New methods should be thoroughly explored under different types of scenarios and conditions to assess their full efficiency. The use of simulations seems crucial for this purpose; by running simulations, researchers know beforehand what should or should not be identified by the statistical test. In this way, researchers can explore multiple scenarios and investigate when tests may fail. We hope that future studies will identify solutions to the problems highlighted here and thus facilitate the development of new methodologies to overcome them.

AUTHORS' CONTRIBUTIONS

I.P.G., S.S. and C.S. conceived the ideas and designed the methodology; I.P.G. rewrote part of the R scripts, ran the simulations, analysed the data, led the writing of the manuscript and created the figures and tables. All authors contributed critically to the drafts and gave final approval for publication.

DATA AVAILABILITY STATEMENT

The code to generate the simulated data, do the data analysis and create the figures of the MS has been uploaded to the DRYAD repository. It can be found under: Code and data of MS 'Null models for animal social network analysis and data collected via focal sampling: pre- or network permutation?' https://doi.org/10.5061/dryad.tmpg4f4w0 (Puga-Gonzalez, Sueur, & Sosa, 2020).

ORCID

Ivan Puga-Gonzalez https://orcid.org/0000-0003-2510-6760
Cédric Sueur https://orcid.org/0000-0001-8206-2739
Sebastian Sosa https://orcid.org/0000-0002-5087-9135

REFERENCES

Altmann, J. (1974). Observational study of behaviour: Sampling methods. *Behaviour*, 49, 227–267. https://doi.org/10.1163/156853974x00534

Aplin, L. M., Farine, D. R., Morand-Ferron, J., Cole, E. F., Cockburn, A., & Sheldon, B. C. (2013). Individual personalities predict social behaviour in wild networks of great tits (*Parus major*). *Ecology Letters*, 16, 1365–1372. https://doi.org/10.1111/ele.12181

Bejder, L., Fletcher, D., & Brager, S. (1998). A method for testing association patterns of social animals. *Animal Behaviour*, 56, 719–725. https://doi.org/10.1006/anbe.1998.0802

Cairns, S. J., & Schwager, S. J. (1987). A comparison of association indices. *Animal Behaviour*, 35, 1454–1469. https://doi.org/10.1016/s0003-3472(87)80018-0

Carnell, R. (2018). *lhs: Latin hypercube samples*. R package version 0.10. https://CRAN.R-project.org/package=lhs

Cranmer, S. J., Leifeld, P., McClurg, S. D., & Rolfe, M. (2017). Navigating the range of statistical tools for inferential network analysis. *American Journal of Political Science*, 61, 237–251. https://doi.org/10.1111/ajps.12263

Croft, D. P., Krause, J., Darden, S. K., Ramnarine, I. W., Faria, J. J., & James, R. (2009). Behavioural trait assortment in a social network: Patterns and implications. *Behavioral Ecology and Sociobiology*, 63, 1495–1503. https://doi.org/10.1007/s00265-009-0802-x

Croft, D. P., Madden, J. R., Franks, D. W., & James, R. (2011). Hypothesis testing in animal social networks. *Trends in Ecology & Evolution*, 26, 502–507. https://doi.org/10.1016/j.tree.2011.05.012

Farine, D. R. (2014). Measuring phenotypic assortment in animal social networks: Weighted associations are more robust than binary edges. *Animal Behaviour*, 89, 141–153. https://doi.org/10.1016/j.anbehav.2014.01.001

Farine, D. R. (2017). A guide to null models for animal social network analysis. *Methods in Ecology and Evolution*, 8, 1309–1320. https://doi.org/10.1111/2041-210x.12772

Farine, D. R., & Whitehead, H. (2015). Constructing, conducting and interpreting animal social network analysis. *Journal of Animal Ecology*, 84, 1144–1163. https://doi.org/10.1111/1365-2656.12418

Firth, J. A., Cole, E. F., Ioannou, C. C., Quinn, J. L., Aplin, L. M., Culina, A., … Sheldon, B. C. (2018). Personality shapes pair bonding in a wild bird social system. *Nature Ecology & Evolution*, 2, 1696–1699. https://doi.org/10.1038/s41559-018-0670-8

Firth, J. A., Sheldon, B. C., & Brent, L. J. N. (2017). Indirectly connected: Simple social differences can explain the causes and apparent consequences of complex social network positions. *Proceedings of the Royal Society B: Biological Sciences*, *284*. https://doi.org/10.1098/rspb.2017.1939

Fisher, D. N., & McAdam, A. G. (2017). Social traits, social networks and evolutionary biology. *Journal of Evolutionary Biology*, *30*, 2088–2103. https://doi.org/10.1111/jeb.13195

Formica, V. A., Wood, C. W., Larsen, W. B., Butterfield, R. E., Augat, M. E., Hougen, H. Y., & Brodie III, E. D. (2012). Fitness consequences of social network position in a wild population of forked fungus beetles (*Bolitotherus cornutus*). *Journal of Evolutionary Biology*, *25*, 130–137. https://doi.org/10.1111/j.1420-9101.2011.02411.x

Hemelrijk, C. K. (1990a). A matrix partial correlation test used in investigations of reciprocity and other social interaction patterns at group level. *Journal of Theoretical Biology*, *143*, 405–420. https://doi.org/10.1016/s0022-5193(05)80036-0

Hemelrijk, C. K. (1990b). Models of, and tests for, reciprocity, unidirectional and other social interaction patterns at a group level. *Animal Behaviour*, *39*, 1013–1029. https://doi.org/10.1016/s0003-3472(05)80775-4

Klaich, M. J., Kinas, P. G., Pedraza, S. N., Coscarella, M. A., & Crespo, E. A. (2011). Estimating dyad association probability under imperfect and heterogeneous detection. *Ecological Modelling*, *222*(15), 2642–2650. https://doi.org/10.1016/j.ecolmodel.2011.03.027

Manly, B. F. J. (1995). A note on the analysis of species cooccurrences. *Ecology*, *76*, 1109–1115.

Pinter-Wollman, N., Hobson, E. A., Smith, J. E., Edelman, A. J., Shizuka, D., de Silva, S., … McDonald, D. B. (2014). The dynamics of animal social networks: Analytical, conceptual, and theoretical advances. *Behavioral Ecology*, *25*, 242–255. https://doi.org/10.1093/beheco/art047

Puga-Gonzalez, I. (2017). Exchange behavior. In J. Vonk & T. Shackelford (Eds.), *Encyclopedia of animal cognition and behavior* (pp. 1–9). Cham, Switzerland: Springer International Publishing.

Puga-Gonzalez, I., Sueur, C., & Sosa, S. (2020). Code and data of MS "Null models for animal social network analysis and data collected via focal sampling: pre-network or node network permutation?, v3, Dryad, Dataset. https://doi.org/10.5061/dryad.tmpg4f4w0

R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Sih, A., Spiegel, O., Godfrey, S., Leu, S., & Bull, C. M. (2018). Integrating social networks, animal personalities, movement ecology and parasites: A framework with examples from a lizard. *Animal Behaviour*, *136*, 195–205. https://doi.org/10.1016/j.anbehav.2017.09.008

Silk, J. B. (2007). Social components of fitness in primate groups. *Science*, *317*, 1347–1351. https://doi.org/10.1126/science.1140734

Sosa, S. (2016). The influence of gender, age, matriline and hierarchical rank on individual social position, role and interactional patterns in *Macaca sylvanus* at 'la forêt des singes': A multilevel social network approach. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00529

Sosa, S. (2018). Social network analysis. In J. Vonk & T. Shackelford (Eds.), *Encyclopedia of animal cognition and behavior* (pp. 1–18). Cham, Switzerland: Springer International Publishing.

Stanton, M. A., & Mann, J. (2012). Early social networks predict survival in wild bottlenose dolphins. *PLoS ONE*, *7*, e47508. https://doi.org/10.1371/journal.pone.0047508

Stein, M. (1987). Large sample properties of simulations using latin hypercube sampling. *Technometrics*, *29*, 143–151. https://doi.org/10.1080/00401706.1987.10488205

Sueur, C., Romano, V., Sosa, S., & Puga-Gonzalez, I. (2019). Mechanisms of network evolution: A focus on socioecological factors, intermediary mechanisms, and selection pressures. *Primates*, *60*, 167–181. https://doi.org/10.1007/s10329-018-0682-7

Sundaresan, S. R., Fischhoff, I. R., & Dushoff, J. (2009). Avoiding spurious findings of nonrandom social structure in association data. *Animal Behaviour*, *77*, 1381–1385. https://doi.org/10.1016/j.anbehav.2009.01.021

Voelkl, B., Vogt, L., Sena, E. S., & Wurbel, H. (2018). Reproducibility of preclinical animal research improves with heterogeneity of study samples. *PLoS Biology*, *16*, e2003693. https://doi.org/10.1371/journal.pbio.2003693

Whitehead, H. (1999). Testing association patterns of social animals. *Animal Behavior*, *57*, F26–F29. https://doi.org/10.1006/anbe.1999.1099

Whitehead, H., Bejder, L., & Ottensmeyer, C. A. (2005). Testing associaiton patterns: Issues arising and extension. *Animal Behavior*, *69*. https://doi.org/10.1016/j.anbehav.2004.11.004

Whitehead, H., & Dufault, S. (1999). Techniques for analyzing vertebrate social structure using identified individuals: Review and recommendations. *Advances in the Study of Behavior*, *28*(28), 33–74. https://doi.org/10.1016/s0065-3454(08)60215-6

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Puga-Gonzalez I, Sueur C, Sosa S. Null models for animal social network analysis and data collected via focal sampling: Pre-network or node network permutation? *Methods Ecol Evol*. 2020;00:1–11. https://doi.org/10.1111/2041-210X.13400