Ecography

ECOG-02379
Wittemyer, G., Keating, L. M., Vollrath, F. and Douglas-Hamilton, I. 2016. Graph theory illustrates spatial and temporal features that structure elephant rest locations and reflect risk perception. – Ecography doi: 10.1111/ecog.02379

Supplementary material

APPENDIX 1

Definition of resting events and rest nodes

Resting events were defined as two consecutive GPS points (at most 4 hours apart to allow for missed GPS points) located less than 52.5 meters apart. This 52.5 meter threshold represents the median distance measured between successive GPS points collected on an elephant (including GPS error) during resting events defined from (1) activity focal sampling and (2) collar attached activity censor data showing no activity across consecutive GPS fixes. Location of the resting event was assigned as the mean location of the GPS locations comprising the resting event.

To define resting locations (nodes) comprising multiple resting events, individual rest events were clustered based on Euclidian distance. We assessed the descriptive power of single linkage, complete linkage, and average linkage clustering methods by comparing respective cophenetic clustering coefficients (Horvath 2011). Average-linkage hierarchical clustering was the best-fitting method (*i.e.* maximum cophenetic coefficient) for all of the elephants (Table A1). A threshold tree-cutting approach was applied to delineate communities of rest points, defined as resting nodes, for which the threshold was visually assigned as the asymptotic value in the cumulative number of nodes to distance relationship.

Network Metrics

Networks, a collection of points (nodes) that are joined by lines (edges) (Newman 2003), can be classified into categories based on the type of relationships that the edges represent or if only a certain type of edge is used. To derive metrics for analysis of graph theoretic properties of elephants rest site structure, we calculated static unweighted networks (edges representing the presence of an observed movement of the elephant between nodes), static weighted networks (where edge weights represented the count of movements between nodes), and temporal networks (including information on the timing of movements between nodes incorporate changes over time).

Using the static unweighted network, we calculated the *global clustering coefficient* (transitivity) defined as the ratio of triangles to unique connected triples in the network (Newman

2010). This metric measures the probability that two nodes are connected to each other given that they are both connected to another node (Newman 2010). More specifically, the global clustering coefficient *C* was calculated as:

C = (number of triangles) x 3 / (number of connected triples)

We computed the mean fraction of node strength from self loops for each undirected weighted static network as the strength from self loops divided by the total strength for each node, averaged across all nodes in the network. We used this mean fraction of strength from self loops to provide insight into the propensity to return to the same sleeping site consecutively given that both rests are in the same treatment group.

Finally, we computed the normalized number of repeated paths of size three as the number of time-ordered triplets that occur at least twice in the time-ordered network representation, divided by the number of unique time-ordered triplets in the network. A time-ordered triplet is a set of three ordered connected nodes (i.e. three consecutive and ordered rest event locations). We excluded time-ordered triplets with self loops from this analysis to avoid duplication with our previous self loop analysis. We used this normalized number of repeated paths to provide a measure of the repeatability of sequential node use.

Random Network Metrics

We investigated both the implications of the spatial distribution of the nodes on our metrics and the significance of our results using a randomization test. Specifically, we simulated a random walk on the nodes for each elephant using a gamma distribution to model the distance between consecutive rests (shape parameters obtained using moment matching with the distances in the corresponding resting network; the random walker moved to the node nearest to a randomly chosen distance from the previous node with each step). Each random walk contains the same number of rest events as the data set for the actual elephant. We generated 10,000 random networks for each elephant, from which distributions of network metrics (clustering coefficients, self loops and repeated motifs) and use distributions for each node were generated. These distributions were then compared with observed values. Results are presented in Table A2.

References

Horvath, S. 2011. Weighted Network Analyses: Applications in Genomics and Systems Biology. — Springer.

Newman, M. E. J. 2003. The structure and function of complex networks. — SIAM Rev. 45: 167-256.

Table A1. Cophenetic coefficients for rest event clusters for nine elephants using average linkage, single linkage, and complete linkage clustering methods. Bold values highlight the maximum cophenetic coefficient for each elephant.

Elephant ID	Cophenetic Coefficient		
	Average Linkage	Single Linkage	Complete Linkage
1	**0.81**	0.73	0.72
2	**0.71**	0.51	0.68
3	**0.85**	0.78	0.83
4	**0.85**	0.77	0.83
5	**0.95**	0.87	0.91
6	**0.87**	0.76	0.85
7	**0.81**	0.72	0.69
8	**0.89**	0.73	0.74
9	**0.73**	0.50	0.64

Table A2. Network metrics for the random networks generated for each of the nine elephant. Values presented are means across 10,000 randomizations. Error estimates represent standard deviation.

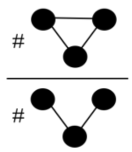
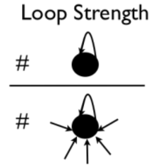
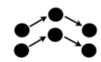
Elephant	Global Clustering Coefficient 	Mean Fraction Self-Loop Strength 	Repeated Paths (size 3)  # Paths Occured At Least Twice / # Paths Occured At Least Once
1	0.091 ± 0.006	0.015 ± 0.003	0.004 ± 0.001
2	0.037 ± 0.002	0.008 ± 0.002	0.001 ± 0.001
3	0.120 ± 0.010	0.103 ± 0.012	0.012 ± 0.005
4	0.103 ± 0.010	0.118 ± 0.011	0.008 ± 0.004
5	0.118 ± 0.008	0.112 ± 0.009	0.028 ± 0.007
6	0.078 ± 0.005	0.073 ± 0.006	0.006 ± 0.002
7	0.084 ± 0.006	0.021 ± 0.004	0.003 ± 0.002
8	0.090 ± 0.007	0.036 ± 0.004	0.005 ± 0.002
9	0.074 ± 0.005	0.020 ± 0.004	0.003 ± 0.002

Table A3. Model selection results for the top models (representing >95% of the AIC weights) of rest properties (restlessness, rest duration and number of rests per day), including number of parameters (k), Akaike Information Criterion (AIC) scores, delta AIC scores (ΔAIC), and AIC weights. Models were fitted to 228,720 rest events from 9 different elephants, with individual elephant treated as a random effect in the model (1 | elephant).

Response Variable	Covariates	*k*	*AIC*	Δ*AIC*	*AIC* Weights
Restlessness					
	protected * time + season * time + (1 | elephant)	7	14843.74	0.00	0.72
	protected * season + protected * time + season * time + (1 | elephant)	9	14845.61	1.87	0.28
Rest Duration[1]					
	season + protected * time + (1 | elephant)	5	61784.26	0.00	0.35
	protected * season + protected * time + (1 | elephant)	7	61784.29	0.03	0.35
	protected * time + season * time + (1 | elephant)	7	61785.82	1.57	0.16
	protected * season + protected * time + season * time + (1 | elephant)	10	61786.04	1.79	0.14
Number of Rests[2]					
	protected * time + season * time + (1 | elephant)	7	36947.11	0.00	0.54
	protected * season + protected * time + season * time + (1 | elephant)	10	36947.46	0.35	0.46

1 The top two models of rest duration were equivalent, but the 95% confidence intervals in the additional parameters in the second model overlapped 0. Therefore, results for the first model were presented.

2 The top two models of number of rests per day were equivalent, but the 95% confidence intervals in the additional parameters in the second model overlapped 0. Therefore, results for the first model were presented.

Table A4. Model selection results for the top models (representing >95% of the AIC weights) of rest network structure (degree, node preference, and repeated paths), including Akaike Information Criterion (AIC) scores, delta AIC scores (ΔAIC), and AIC weights. Model selection results for the top 5 models (i.e. those with ΔAIC < 2) and global model of Self Loops are presented, given the lack of model differentiation for this response variable.

<table>
<tr><th>Response Variable</th><th>Covariates</th><th>AIC</th><th>ΔAIC</th><th>AIC Weights</th></tr>
<tr><td>Degree[1]</td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td>%protected + node_area + %day + nearest neighbor + dist.water +%protected*dist.water (1 | elephant)</td><td>17026.6</td><td>0.00</td><td>0.50</td></tr>
<tr><td></td><td>%protected + node_area + %day + nearest neighbor + dominance_scaled + dist.water +%protected*dist.water + (1 | elephant)</td><td>17044.63</td><td>0.02</td><td>0.50</td></tr>
<tr><td>Preferred Nodes</td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td>node_area + %day + nearest neighbor + dist.water +%protected*dist.water + (1 | elephant)</td><td>1110.93</td><td>0.00</td><td>0.88</td></tr>
<tr><td></td><td>%protected + node_area + %day + nearest neighbor + dist.water +%protected*dist.water + (1 | elephant)</td><td>1115.17</td><td>4.24</td><td>0.11</td></tr>
<tr><td>Repeated Paths</td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td>%protected + %day+ dist.water +%protected*dist.water + (1 | elephant)</td><td>2109.48</td><td>0.00</td><td>0.72</td></tr>
<tr><td></td><td>%protected + %day + dominance_scaled + dist.water +%protected*dist.water + (1 | elephant)</td><td>2111.40</td><td>1.92</td><td>0.28</td></tr>
<tr><td>Self Loop Strength[2]</td><td></td><td></td><td></td><td></td></tr>
</table>

%protected + season + dist.water (1 | elephant)	1911.56	0.00	0.16
%protected + season + dist.water + dominance (1 | elephant)	1912.55	0.99	0.10
%protected + season + dist.water + dist.water*%protected + (1 | elephant)	1913.05	1.49	0.07
%protected + season + dist.water + nearest neighbor (1 | elephant)	1913.32	1.76	0.06
%protected + %day + season + dist.water (1 | elephant)	1913.55	1.99	0.06
%protected + %day + nearest neighbor + dominance + season + dist.water + dist.water*%protected + (1 | elephant)	1919.11	7.55	0.01

1 The top two models of degree were equivalent, but the 95% confidence intervals in the additional parameter (dominance-scaled) in the second model overlapped 0. Therefore, results for the top model were presented.

2 Top models of Self Loop Strength were not strongly differentiated, but the 95% confidence intervals in the additional parameters in secondary models overlapped 0. Therefore, results for the top model were presented.