Advanced social network analysis, assignment 2

Grotto Tommaso

1. Introduction

This analysis uses the EIES dataset, which captures a directed and valued network of 32 network scientists communicating through an Electronic Information Exchange System (EIES). The network data reflects the strength of relationships, ranging from 0 (no connection) to 4 (close friendship). To simplify the analysis, the network was dichotomized: ties with values of 3 and 4 were recoded as 1 (indicating a connection), while values of 0, 1, and 2 were recoded as 0 (indicating no connection). Additionally, gender attributes for the individuals in the network were included for further analysis.

Three Exponential Random Graph Models (ERGMs) were specified to analyze the network. Model 1: structural effects. This model focuses on structural properties of the network: edges, the baseline tendency for ties to form; mutual, propensity for reciprocity in ties; gwidegree and gwodegree, generalized effects for in-degrees and out-degrees; gwesp (OTP and ITP), geometrically weighted edgewise shared partners, capturing triadic closure and transitivity.

Model 2: dyadic independent with attributes. This model uses edges and mutual effects and it incorporates individual-level attributes and their effects on tie formation: nodeicov (Gender), influence of gender on receiving ties; nodeocov (Gender), influence of gender on sending ties; nodematch (Gender), tendency for ties to form between individuals of the same gender.

Model 3: full model (structural and attributes). The final model combines structural and attribute-based effects, integrating all terms from Models 1 and 2. This allows for a comprehensive examination of how both network structure and individual attributes interact to shape tie formation.

These models provide a layered approach to understanding the dynamics of the EIES network, from purely structural effects to the interplay between structure and attributes.

2. Models results

The results from the three Exponential Random Graph Models (ERGMs) provide insights into the structural and attribute-based dynamics of the EIES friendship network. Below, we focus on Model 3, which combines structural and attribute effects, while highlighting differences with Models 1 and 2 when relevant.

Table 1: ERGMs results for the three models

	Model 1			Model 2			Model 3		
	Est	SE	Sign	Est	SE	Sign	Est	SE	Sign
edges	-4.82	0.36	***	-2.61	0.16	***	-4.87	0.39	***
mutual	2.60	0.34	***	2.25	0.28	***	2.68	0.35	***
gwideg	2.19	1.13	.				2.18	1.26	.
gwodeg	4.37	2.79					9.19	4.47	*
gwesp.OTP	1.97	021	***				1.85	0.22	***
gwesp.ITP	-0.62	0.11	***				-0.71	0.11	***
nodeicov Gender				0.52	0.20	**	0.22	0.17	
nodeocov Gender				1.03	0.18	***	0.70	0.18	***
nodematch Gender				-0.06	0.14		-0.07	0.16	

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Model 3: Full Model (Structural and Attributes)

Edges
The negative and highly significant coefficient confirms that the network is sparse, meaning that the baseline probability of forming ties is low. This aligns with expectations for a directed friendship network, where connections are selective.

Mutual
Reciprocity remains a dominant feature of the network, with a positive and highly significant coefficient, this indicates a strong tendency for individuals to reciprocate ties.

GWIDEGREE (Generalized Weighted In-Degree)
The marginally significant coefficient suggests a tendency for certain individuals to attract many incoming ties, reflecting potential popularity effects. While this effect is not so pronounced, it remains relevant in the context of friendships, where some individuals may naturally draw more attention or connection requests.

GWODEGREE (Generalized Weighted Out-Degree)
This effect is significant, indicating a strong preference for nodes that actively send many ties. In a friendship network, this could represent highly sociable individuals who initiate connections with many peers. This result is different from Model 1, since in that case the effect was not significant.

GWESP (Geometrically Weighted Edgewise Shared Partners)

OTP (Outgoing Two-Path Closure)
The positive and significant coefficient suggests a preference for cohesive triadic structures, where individuals are likely to form ties that reinforce existing connections. This reflects the importance of mutual friends in fostering social cohesion.

ITP (Incoming Two-Path Closure)
The negative and significant coefficient indicates that transitive closure for incoming ties is less common. This suggests that while individuals may form outgoing connections to reinforce triads, they are less likely to receive ties that close such structures.

Nodeicov (Gender)
The nonsignificant coefficient indicates that gender does not significantly influence the likelihood of receiving ties when structural factors are accounted for. This contrasts with Model 2, where gender had a significant effect on incoming ties, suggesting that structural dynamics may mediate gender-based preferences.

Nodeocov (Gender)
The positive and significant coefficient shows that females are more active in sending ties compared to males.

Nodematch (Gender)
The nonsignificant coefficient confirms that there is no significant tendency for individuals of the same gender to form ties. This lack of gender homophily is consistent across all models, suggesting that friendships in this network are not strongly influenced by gender similarity.

3. Goodness of fit

Evaluating the goodness of fit (GOF) is essential to assess how well Model 3 replicates the observed characteristics of the EIES friendship network. GOF analysis compares key network statistics between the observed network and a set of networks simulated based on the model. This ensures that the model not only provides statistically significant estimates but also accurately represents the structural and attribute-based dynamics of the network.

Table 2: Goodness of fit results for Model 3

	obs	min	mean	max	MC p-value
edges	204.00	159.00	204.57	258.00	0.997
mutual	60.00	35.00	60.63	84.00	0.991
gwideg	41.75	37.10	41.82	43.12	0.787
gwodeg	42.57	40.89	42.57	43.14	0.876
gwesp.OTP	271.16	175.98	272.59	374.97	0.942
gwesp.ITP	217.35	98.97	220.19	333.33	0.916
nodeicov.Gender	96.00	57.00	96.51	130.00	0.982
nodeocov.Gender	108.00	65.00	108.26	148.00	1.000
nodematch.Gender	90.00	64.00	90.02	123.00	1.000

Looking at the table we can see these results:

Edges
The observed number of edges aligns closely with the mean of the simulated networks, indicating that the model accurately captures the network's overall sparsity. The high MC p-value suggests no significant difference between observed and simulated edge counts, confirming a good fit for network density.

Mutual
The model captures the tendency for reciprocity effectively, as the observed and simulated values are nearly identical. This strong fit reflects the importance of mutual ties in the friendship network, a critical structural feature.

GWIDEGREE (Generalized Weighted In-Degree)
The observed in-degree distribution closely matches the simulated mean, suggesting the

model represents popularity effects (nodes receiving many incoming ties) well, although the MC p-value is the lowest among the others.

GWODEGREE (Generalized Weighted Out-Degree)
The out-degree distribution shows a near-perfect match, reflecting the model's ability to capture activity effects (nodes sending many ties).

GWESP (Geometrically Weighted Edgewise Shared Partners)

OTP (Outgoing Two-Path Closure)
The positive alignment indicates the model captures clustering tendencies related to outgoing ties, reinforcing the importance of triadic cohesion in the network.

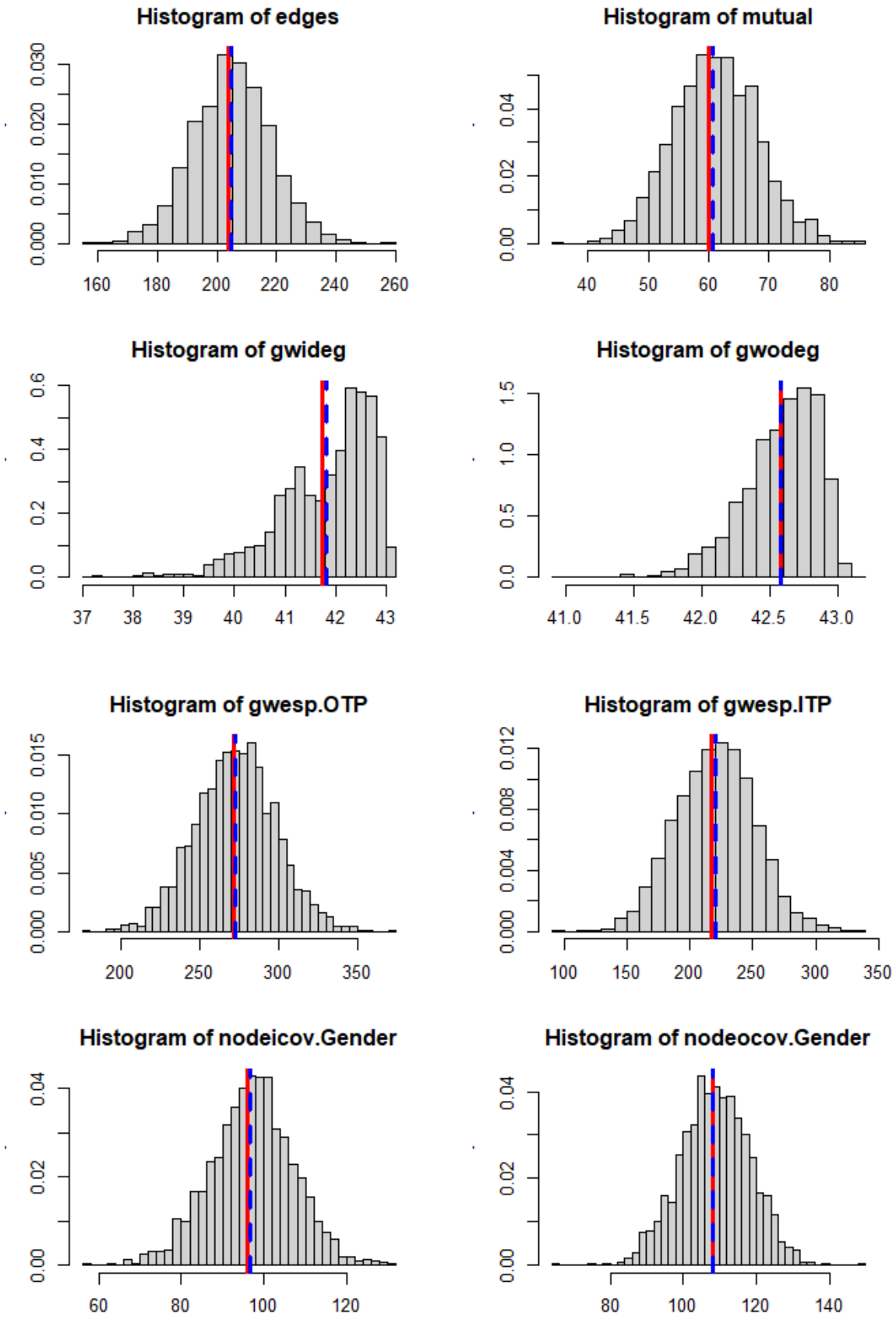
ITP (Incoming Two-Path Closure)
Similarly, the model effectively represents clustering for incoming ties, despite the observed tendency for less transitive closure in this context.

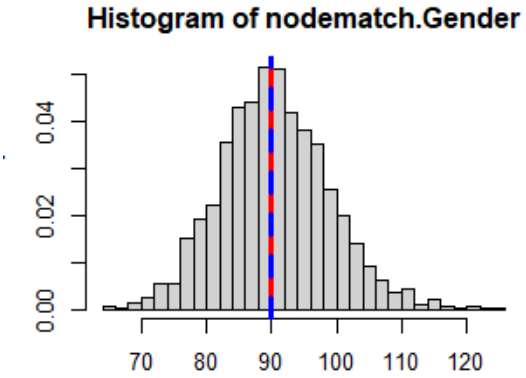
Nodeicov (Gender)
The observed and simulated values align closely, indicating that the model captures the influence of gender on receiving ties effectively.

Nodeocov (Gender)
The perfect fit reflects the model's accuracy in representing gender-based differences in outgoing ties.

Nodematch (Gender)
The model perfectly reproduces gender-based homophily, confirming that individuals are no more likely to form ties with others of the same gender.

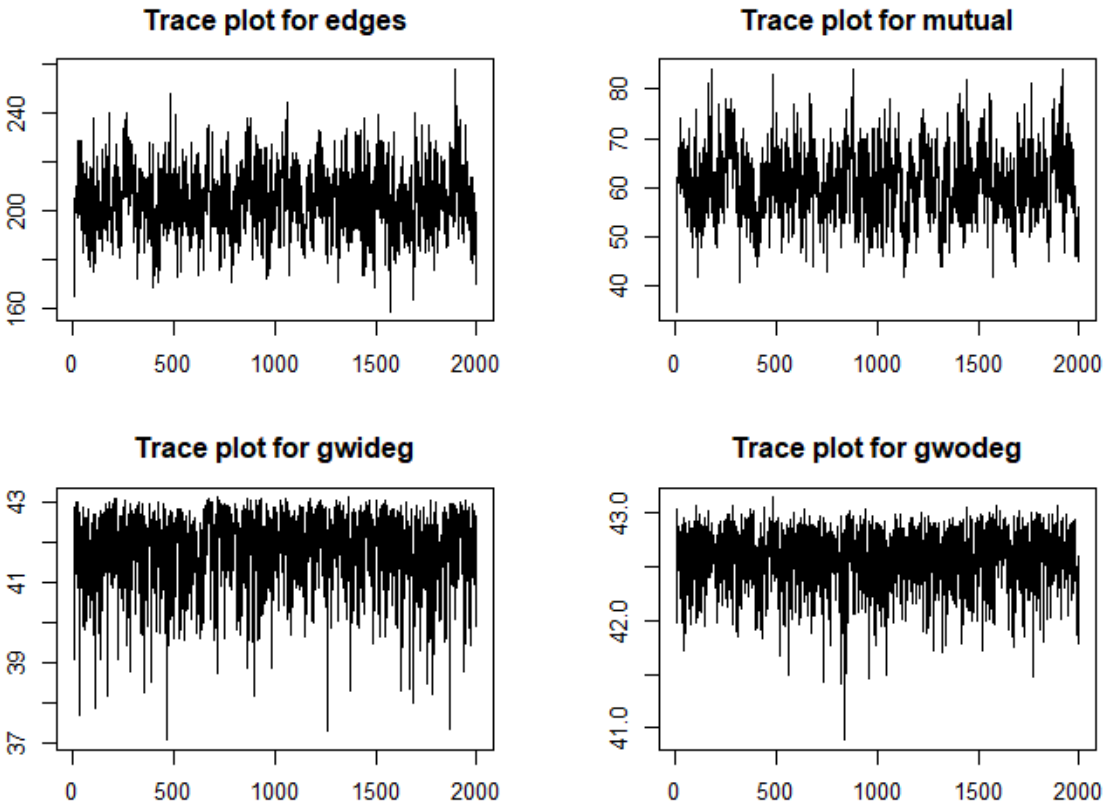
Graph 1: Histograms of the goodness of fit results for Model 3

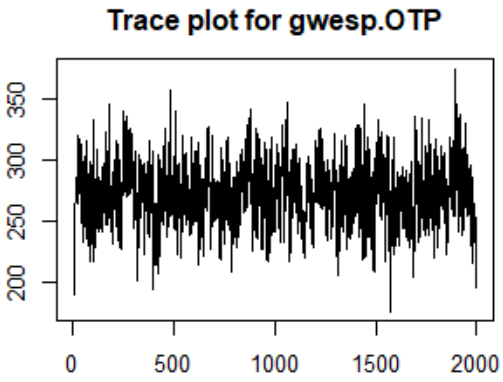


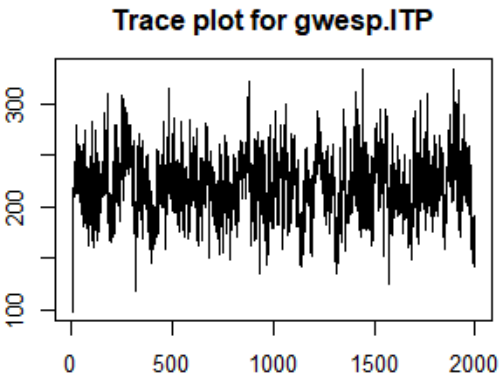


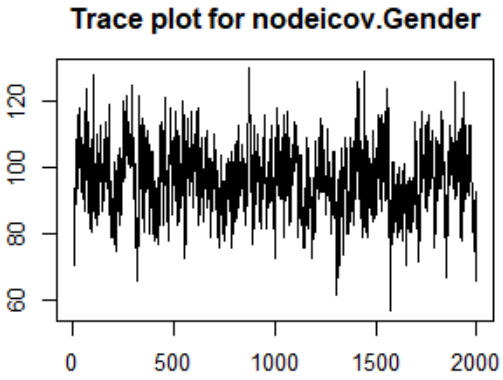
The observed statistics consistently fall within the distribution of simulated values, demonstrating the model's ability to replicate the network's structural and attribute-based features.

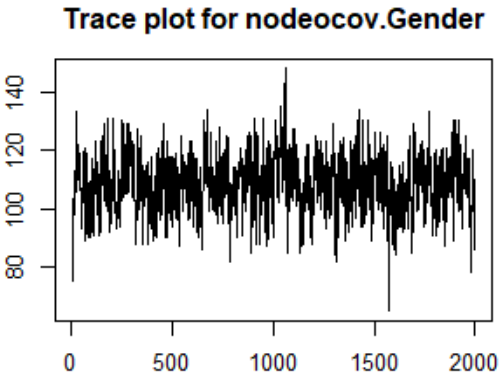
Graph 2: Trace plots of the goodness of fit results for Model 3

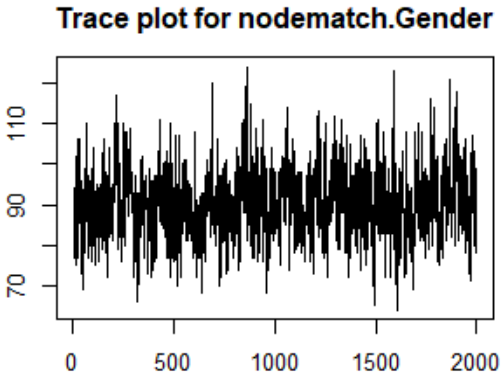






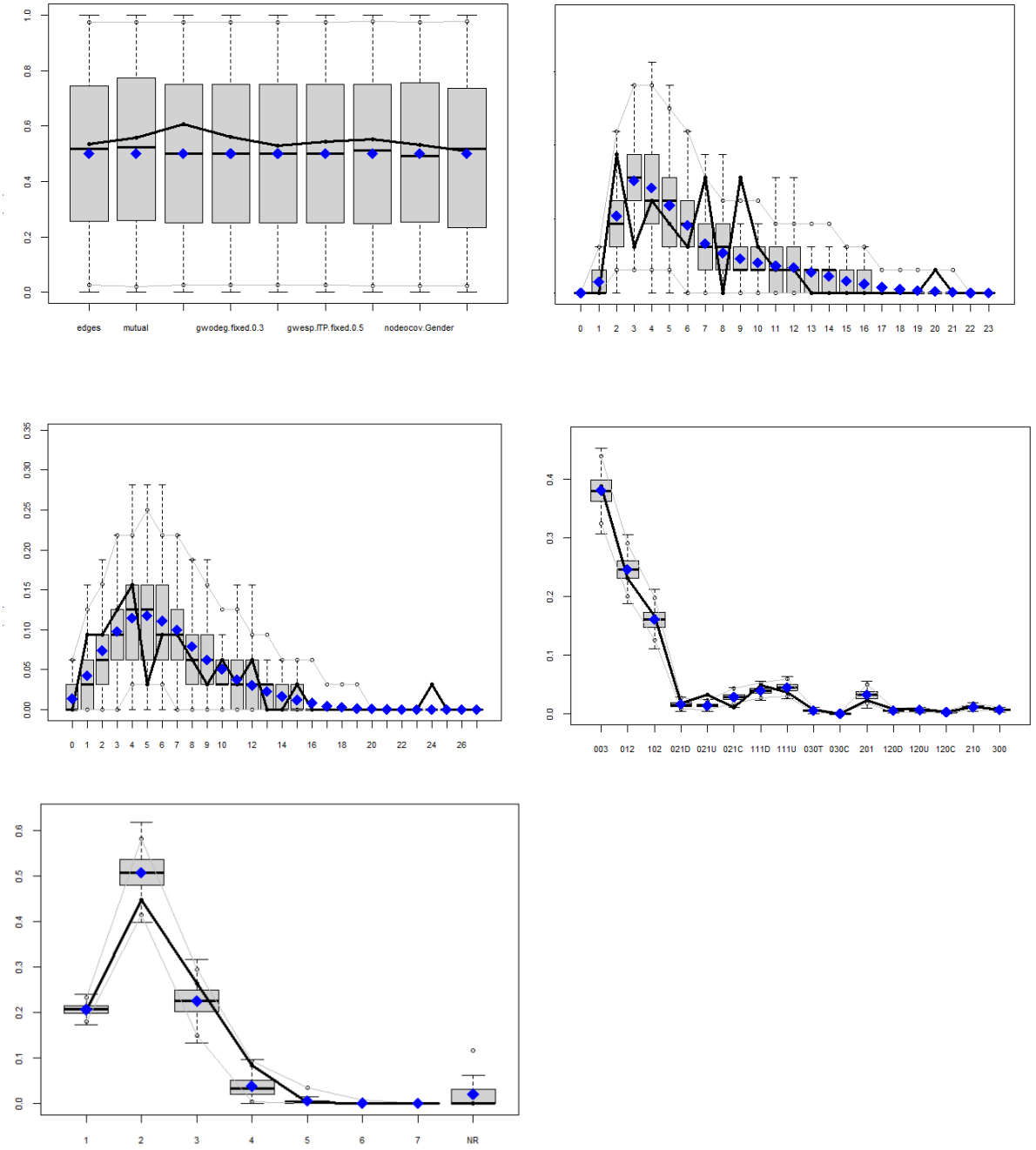






The trace plots show stable convergence of the simulated statistics around the observed values, indicating reliable model performance.

Graph 3: Boxplots of the goodness of fit results for Model 3



The boxplots illustrate minimal deviation between observed and simulated statistics, only on a few occasions they do not align, further supporting the model's goodness of fit.

The model effectively captures key network dynamics, including reciprocity, clustering tendencies, degree distributions, and the role of gender in tie formation and reception. The results demonstrate that the network is characterized by strong reciprocity and cohesive triadic structures, with individuals showing a preference for forming mutual and clustered ties. The inclusion of gender as a node attribute highlights its influence on outgoing ties,

though it does not significantly affect incoming ties or homophily. This suggests that while gender plays a role in network activity, it is not a primary driver of tie formation within the group.

The goodness-of-fit analysis confirms that Model 3 aligns closely with the observed network, with minimal discrepancies across all evaluated statistics. The high MC p-values indicate that the model replicates the observed network's features accurately, from overall density to gender-related effects.

Model 3 successfully balances structural and attribute-based effects, offering a robust explanation of the network's underlying dynamics. Its ability to replicate the observed network features, including density, reciprocity, clustering, and gender effects, suggests that it captures the essential mechanisms driving tie formation in this context.