Supplementary Material – Exponential Random Graph Models for Dynamic Signed Networks: An Application to International Relations

Cornelius Fritz*
School of Computer Science and Statistics, Trinity College Dublin
Marius Mehrl
School of Politics and International Studies, University of Leeds
Paul W. Thurner
Geschwister Scholl Institute of Political Science, LMU Munich
Göran Kauermann
Department of Statistics, LMU Munich

October 6, 2024

1 Technical Details

1.1 Gibbs Sampler

Before describing the estimation and model selection procedures in more detail, the Gibbs sampler used to obtain samples from the SERGM, as defined in equations (2) and (3) in the main article for static and dynamic networks, respectively, is described and visualized in algorithm 1. Since the temporal case is equivalent to the static case, we focus on the simpler case of static networks. The goal is to sample a network $\boldsymbol{y}$ following the SERGM distribution characterized by a set of sufficient statistics (possible choices are given in Section 2.2 of the main article), each weighted by the parameter $\boldsymbol{\theta}$. For the kth step out of K of the Gibbs sampler, we start by selecting a random dyad $(i^{(k)}, j^{(k)})$ from all possible edges. Then, we sample the value of $(i^{(k)}, j^{(k)})$ from a multinomial distribution with $n = 1$ and success probabilities

$$\frac{\exp\left\{\boldsymbol{\theta}^\top \boldsymbol{s}(\boldsymbol{y}^+_{ij,t}, \boldsymbol{y}_{t-1})\right\}}{\sum_{k\in\{+,-,0\}} \exp\left\{\boldsymbol{\theta}^\top \boldsymbol{s}(\boldsymbol{y}^k_{ij,t}, \boldsymbol{y}_{t-1})\right\}}, \frac{\exp\left\{\boldsymbol{\theta}^\top \boldsymbol{s}(\boldsymbol{y}^-_{ij,t}, \boldsymbol{y}_{t-1})\right\}}{\sum_{k\in\{+,-,0\}} \exp\left\{\boldsymbol{\theta}^\top \boldsymbol{s}(\boldsymbol{y}^k_{ij,t}, \boldsymbol{y}_{t-1})\right\}}, \frac{\exp\left\{\boldsymbol{\theta}^\top \boldsymbol{s}(\boldsymbol{y}^0_{ij,t}, \boldsymbol{y}_{t-1})\right\}}{\sum_{k\in\{+,-,0\}} \exp\left\{\boldsymbol{\theta}^\top \boldsymbol{s}(\boldsymbol{y}^k_{ij,t}, \boldsymbol{y}_{t-1})\right\}}.$$

*Corresponding Autor: fritzc@tcd.ie

Subsequently, the dyad $(i^{(k)}, j^{(k)})$ is set to be the result of this multinomial distribution in $\boldsymbol{y}^{(k)}$. As shown in the main manuscript in equation (4), this multinomial distribution corresponds to the conditional distribution of the SERGM. Therefore, continuing this process long enough will provide samples from the desired distribution. Since these types of MCMC methods suffer from autocorrelation between successive steps of the sampler (e.g., $\boldsymbol{y}^{(k)}$ and $\boldsymbol{y}^{(k+1)}$ differ at most in the entry $(i^{(k+1)}, j^{(k+1)})$), we use only every Mth result of the algorithm as a sample. This commonly used method is called *thinning*.

Result: $\boldsymbol{y}^{(1)}, ..., \boldsymbol{y}^{(K)}$
Set $\boldsymbol{y}^{(0)}, \boldsymbol{\theta}$ and K
Set $\boldsymbol{y} = \boldsymbol{y}^{(0)}$
for $k = 1, ..., K$ **do**
  Select dyad $(i^{(k)}, j^{(k)})$ randomly from all possible dyads
  Sample $Y_{i^{(k)},j^{(k)}} | \boldsymbol{Y}^{C}_{i^{(k)},j^{(k)}}$ from $\{1, -1, 0\}$ with probabilities given in equation (4) of the main article
  Set $\boldsymbol{y}^{(k)} = \boldsymbol{y}$
end

Algorithm 1: Gibbs sampler for signed networks.

1.2 Partial Stepping Algorithm

Following standard theory of exponential families, $\boldsymbol{\theta}$ maximizing the approximate likelihood detailed in (11) of the main article only exists if the observed sufficient statistics $\sum_{t=1}^{T} \boldsymbol{s}(\boldsymbol{y}_t, \boldsymbol{y}_{t-1})$ are inside the convex hull spanned by the sampled sufficient statistics $\left(\sum_{t=1}^{T} \boldsymbol{s}(\boldsymbol{y}_t^{(m)}, \boldsymbol{y}_{t-1}), ..., \sum_{t=1}^{T} \boldsymbol{s}(\boldsymbol{y}_t^{(M)}, \boldsymbol{y}_{t-1})\right)$ (Barndorff-Nielsen, 1978). Since this condition does not hold for arbitrary values of $\boldsymbol{\theta}_0$, we adapt the partial stepping algorithm introduced by Hummel et al. (2012) to find an adequate $\boldsymbol{\theta}_0$.

In the kth step of this iterative procedure, we substitute $\sum_{t=1}^{T} \boldsymbol{s}(\boldsymbol{y}_t, \boldsymbol{y}_{t-1})$ in (11) of the main article by

$$\boldsymbol{\xi}^{(k)} = \gamma^{(k)} \sum_{t=1}^{T} \boldsymbol{s}(\boldsymbol{y}_t, \boldsymbol{y}_{t-1}) + \left(1 - \gamma^{(k)}\right) \widehat{\boldsymbol{m}}^{(k)}, \tag{1}$$

where $\gamma^{(k)} \in (0, 1]$ and $\widehat{\boldsymbol{m}}^{(k)} = \frac{1}{M} \sum_{m=1}^{M} \sum_{t=1}^{T} \boldsymbol{s}(\boldsymbol{y}_t^{(m)}, \boldsymbol{y}_{t-1})$ is the estimated mean of the sufficient statistics of networks sampled under $\boldsymbol{\theta}^{(k)}$. We select the largest possible value of $\gamma^{(k)}$ in (1) such that even the point marginally closer to $\sum_{t=1}^{T} \boldsymbol{s}(\boldsymbol{y}_t, \boldsymbol{y}_{t-1})$, defined by $1.05\gamma^{(k)} \sum_{t=1}^{T} \boldsymbol{s}(\boldsymbol{y}_t, \boldsymbol{y}_{t-1}) + \left(1 - 1.05\gamma^{(k)}\right) \boldsymbol{m}^{(k)}$, is inside the convex hull spanned by the sampled statistics. One can test whether a point $\in \mathbb{R}^p$ lies in this convex hull via a linear programming problem (details can be found in Hummel et al., 2012 and Krivitsky et al., 2023).

To update $\boldsymbol{\theta}^{(k)}$ to $\boldsymbol{\theta}^{(k+1)}$ for a given $\gamma^{(k)}$, we thus optimize

$$\left(\boldsymbol{\theta}^{(k+1)} - \boldsymbol{\theta}^{(k)}\right)^{\top} \boldsymbol{\xi}^{(k)} - \log\left(\frac{1}{M} \sum_{m=1}^{M} \exp\left\{(\boldsymbol{\theta} - \boldsymbol{\theta}_0)^{\top} \left(\sum_{t=1}^{T} \boldsymbol{s}(\boldsymbol{y}_t^{(m)}, \boldsymbol{y}_{t-1})\right)\right\}\right), \tag{2}$$

with a Newton-Raphson algorithm and $\boldsymbol{y}_t^{(m)}\ \forall\ m = 1, ..., M$ and $t = 1, ..., T$ sampled from model (3) of the main article. To ease this step, we assume that $\sum_{t=1}^{T} \boldsymbol{s}(\boldsymbol{Y}_t, \boldsymbol{y}_{t-1})$ follows a p-variate Gauss distribution with mean $\boldsymbol{m}^{(k)}$ and covariance matrix $\boldsymbol{\Sigma}^{(k)}$, which is the covariance matrix of the sufficient statistics under $\boldsymbol{\theta}^{(k)}$. Both terms can be estimated with samples $\boldsymbol{Y}^{(1)}, ..., \boldsymbol{Y}^{(M)}$. Then we can state the optimal value of (2) in closed form:

$$\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} + \left(\widehat{\boldsymbol{\Sigma}}^{(k)}\right)^{-1} \left(\boldsymbol{\xi}^{(k)} - \widehat{\boldsymbol{m}}^{(k)}\right).$$

The algorithm terminates when we estimate $\boldsymbol{\gamma}^{(k)} = 1$ two iterations in a row, we then continue the procedure with $\boldsymbol{\xi}^{(k)} = \sum_{t=1}^{T} \boldsymbol{s}(\boldsymbol{y}_t, \boldsymbol{y}_{t-1})$ until the estimates stabilize.

1.3 Evaluation of the AIC

To decide between alternative specifications of the sufficient statistics, a common method is to select the model with the lowest AIC value. The AIC is defined as

$$\text{AIC}(\boldsymbol{M}) = 2p - 2\ell(\hat{\boldsymbol{\theta}}; \boldsymbol{y}), \tag{3}$$

where $\boldsymbol{M}$ is a SERGM for temporal networks with a particular specification of the sufficient statistics and estimated parameters $\hat{\boldsymbol{\theta}}$ and $\ell(\hat{\boldsymbol{\theta}}; \boldsymbol{y}) = \log\left(\prod_{t=1}^{T} \mathbb{P}_{\boldsymbol{\theta}}(\boldsymbol{Y}_t = \boldsymbol{y}_t | \boldsymbol{Y}_{t-1} = \boldsymbol{y}_{t-1})\right)$ is the log likelihood. To evaluate (3), we have to calculate the value of the intractable logarithmic likelihood at $\hat{\boldsymbol{\theta}}$, which we can restate by

$$\ell(\hat{\boldsymbol{\theta}}; \boldsymbol{y}) = r(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}}_{\text{Ind}}; \boldsymbol{y}) + \ell(\hat{\boldsymbol{\theta}}_{\text{Ind}}; \boldsymbol{y}), \tag{4}$$

where $\hat{\boldsymbol{\theta}}_{\text{Ind}} \in \mathbb{R}^p$ is the estimate of the sub-model including only the subset from the sufficient statistics that abide the conditional dependence assumption (the coefficients of all other (endogenous) statistics are fixed at 0). Due to this characteristic, $\ell(\hat{\boldsymbol{\theta}}_{\text{Ind}}; \boldsymbol{y})$ is equivalent to the log likelihood in a multinomial regression and can be computed in closed form. To evaluate $r(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}}_{\text{Ind}}; \boldsymbol{y})$, we follow Hunter and Handcock (2006) and apply path sampling (Gelman and Meng, 1998) to approximate

$$\log\left(\mathbb{E}_{\boldsymbol{\theta}_0}\left(\exp\left\{(\boldsymbol{\theta} - \boldsymbol{\theta}_0)^\top \left(\sum_{t=1}^{T} \boldsymbol{s}(\boldsymbol{Y}_t, \boldsymbol{y}_{t-1})\right)\right\}\right)\right) = \log\left(\frac{\boldsymbol{\kappa}(\boldsymbol{\theta})}{\boldsymbol{\kappa}(\boldsymbol{\theta}_0)}\right)$$

from (11) in the main article in a more precise manner. If we specify a smooth mapping $\boldsymbol{\theta} : [0, 1] \to \mathbb{R}^p$ with $\boldsymbol{\theta}(0) = \hat{\boldsymbol{\theta}}_{\text{Ind}}$ and $\boldsymbol{\theta}(1) = \hat{\boldsymbol{\theta}}$ and let $0 = u_0 < u_1 < ... < u_J = 1$ for $\boldsymbol{J} \in \mathbb{N}$ be a fixed grid of so-called bridges and its finite support, the following approximation holds:

$$\begin{aligned}
\log\left(\frac{\boldsymbol{\kappa}(\boldsymbol{\theta})}{\boldsymbol{\kappa}(\boldsymbol{\theta}_0)}\right) &= \sum_{j=1}^{J} \frac{1}{u_j - u_{j-1}} \mathbb{E}_{\boldsymbol{\theta}(u_j)}\left(\left(\frac{\mathrm{d}}{\mathrm{d}u}\boldsymbol{\theta}(u)\Big|_{u=u_j}\right)^\top \sum_{t=1}^{T} \boldsymbol{s}\left(\boldsymbol{Y}_t, \boldsymbol{y}_{t-1}\right)\right) \\
&\approx \frac{1}{\boldsymbol{M}} \sum_{j=1}^{J} \sum_{m=1}^{M} \frac{1}{u_j - u_{j-1}} \left(\left(\frac{\mathrm{d}}{\mathrm{d}u}\boldsymbol{\theta}(u)\Big|_{u=u_j}\right)^\top \sum_{t=1}^{T} \boldsymbol{s}\left(\boldsymbol{y}_t^{(j,m)}, \boldsymbol{y}_{t-1}\right)\right),
\end{aligned} \tag{5}$$

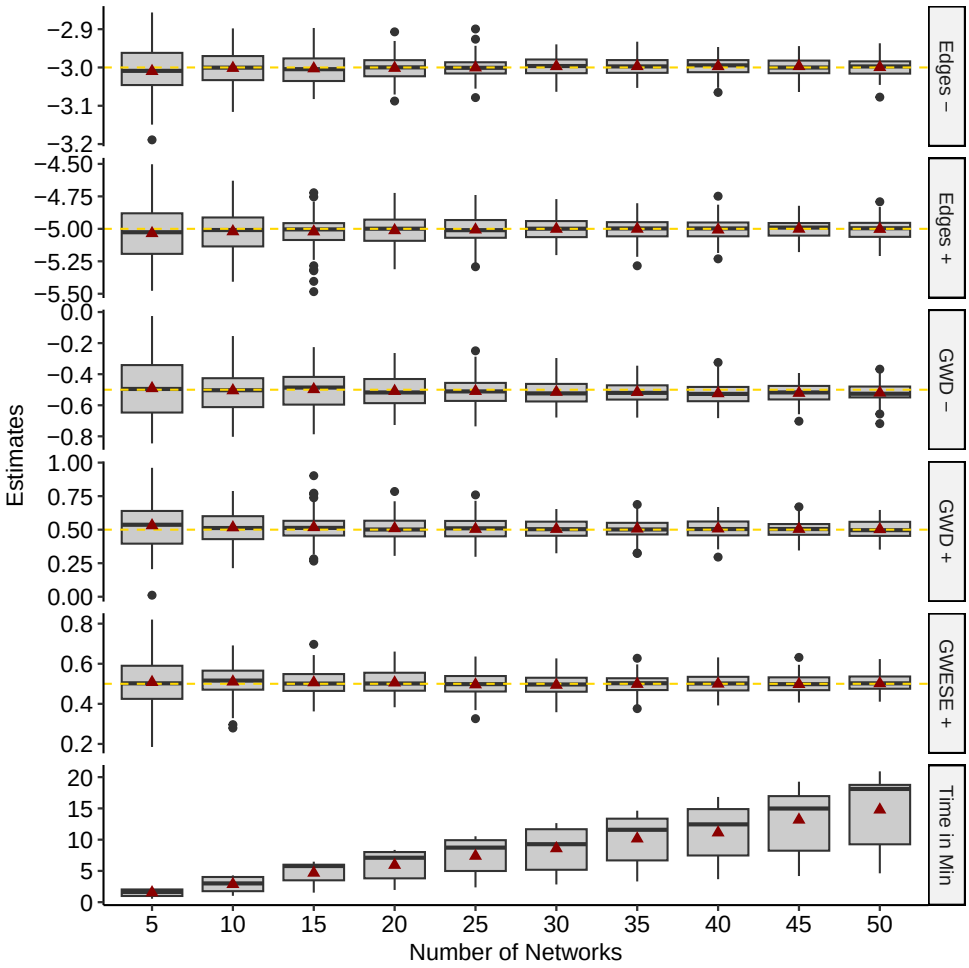


Figure 1: Simulation results in setting 1: The x-axis relates to different number of consecutively observed networks, while the y-axis represents the value of the coefficients (in the first five rows) and the needed time for estimation (in the last row). The results of all 100 simulations are summarized in box plots, the yellow dashed line relates to the true coefficient value, and the red triangles depict the arithmetic mean of the simulations.

where $\boldsymbol{y}_t^{(j,1)}, ..., \boldsymbol{y}_t^{(j,M)}$ are networks sampled conditional on $\boldsymbol{y}_{t-1}$ under $\boldsymbol{\theta}(u_j)\ \forall\ j = 1, ..., J$. For our implementation, we set $\boldsymbol{\theta}(u) = \hat{\boldsymbol{\theta}}_{\text{Ind}} + u(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}_{\text{Ind}})$, corresponding to a linear path from $\hat{\boldsymbol{\theta}}_{\text{Ind}}$ to $\hat{\boldsymbol{\theta}}$ and $\frac{\mathrm{d}}{\mathrm{d}u}\boldsymbol{\theta}(u) = \hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}_{\text{Ind}}$. Plugging (5) into the first row of (11) of the main article permits the computation of (3). For a more technical derivation of (5), we refer to Hunter and Handcock, 2006 or Gelman and Meng, 1998.

2 Simulation Study

To assess the quality of the proposed estimation algorithm, we carry out two separate Monte Carlo simulation studies. Generally, we simulate 100 networks with 100 actors under the SERGM introduced in the main article with the following five network effects and true coefficients:

		ML			MPL		
	Coef.	AVE	RMSE	CP	AVE	RMSE	CP
Edges +	−5.0	−5.012	0.156	0.99	−4.996	0.161	0.69
Edges −	−3.0	−3.003	0.038	0.93	−2.998	0.039	0.86
GWD +	0.5	0.508	0.134	0.99	0.499	0.141	0.61
GWD −	−0.5	−0.508	0.135	0.93	−0.509	0.141	0.74
GWESE +	0.5	0.490	0.082	0.98	0.487	0.087	0.66

Table 1: Simulation results in setting 2: The AVE (AVerage Estimate), RMSE (Root-Mean-Squared Error), and CP (Coverage Probability) are provided for the Maximum Likelihood (ML) and Maximum Pseudo-Likelihood (MPL) approach.

1. Number of positive edges (Edges +): −5
2. Number of negative edges (Edges −): −3
3. Geometrically weighted degree distribution of positive edges (GWD +): 0.5
4. Geometrically weighted degree distribution of positive edges (GWD −): - 0.5
5. Geometrically weighted positive edgewise-shared enemies distribution (GWESE +): 0.7

We apply the partial stepping algorithm detailed in Section 1.2 to find the maximum likelihood estimators (the computational settings of the estimation routine are reported in Section 5). With the first experiment, we study how the number of networks affects the accuracy of the parameter estimates and the needed time until convergence. Therefore, we vary T between 5 and 50. In the second study, we investigate the coverage probabilities of our estimators in the setting of repeated sampling. This analysis will allow us to assert the uncertainty quantification reported in our findings.

Setting 1 (Accuracy under growing T): The result of this simulation study are presented in Figure 1 and uncover three takeaways. First, it is visible that, on average, the correct parameter estimate is recovered in all scenarios, which can be seen as empirical proof of the correctness of the implementation and derivation of the algorithms from Section 1.2. Second, the accuracy of the estimates gets better with the number of observed networks since the width of the box plots becomes smaller with T increasing. Heuristically, this empirical observation follows from the consistency of the ML estimator with growing T. Third, the time needed for the estimation grows almost linearly with T and hits its maximal mean around 25 minutes if $T = 50$.

Setting 2 (Coverage probability): To emulate the core findings of structural balance theory, we add another endogenous effect calculating the geometrically weighted positive edgewise-shared friends distribution (GWESF +) with a coefficient of 1 to the second simulation setting. Now, we contrast the maximum likelihood (ML) estimator with the

maximum pseudo-likelihood (MPL) estimator. For this comparison, we investigate the average point estimate (AVE), the root-mean-squared error (RMSE), and the coverage probabilities (CP) in Table 1. The AVE of a specific coefficient θ_p is its average estimate over the 100 simulation runs:

$$\text{AVE}(\theta_p) = \frac{1}{S}\sum_{s=1}^{S}\hat{\theta}_{p,s},$$

where $\hat{\theta}_{p,s}$ is the estimate of the pth coefficient in the sth simulation run. To assess how volatile the estimation error is over the simulation runs, we report the RMSE of θ_p:

$$\text{RMSE}(\theta_p) = \sqrt{\frac{1}{S}\sum_{s=1}^{S}\left(\hat{\theta}_{p,s}-\theta_p\right)^{\top}\left(\hat{\theta}_{p,s}-\theta_p\right)},$$

where θ_p is the ground truth of the pth coefficient. In order to explore the adequacy of the uncertainty quantification, we compute the percentage of runs in which the true parameter lies within the confidence intervals provided by the respective estimation technique. This coverage probability should be around 95%. The results in Table 1 indicate that, on average, the coefficient estimates of the ML and MPL routines are correct. However, the MPL approach has almost consistently a higher RMSE than the ML approach. In line with the observation of van Duijn et al. (2009) for binary ERGMs, the coverage probabilities of the MPL are off. However, extending the techniques recently developed by Schmid and Hunter (2023) or Schmid and Desmarais (2017) could correct this phenomenon. In summary, the results underpin the usage of the MPLE as the starting value for the stepping algorithm.

3 Details on the Application to International Cooperation and Conflict

3.1 Data Visualization and Covariate Details

Here, we visualize the network and offer additional details regarding the data sources for the application of the SERGM to interstate relations presented in Section 4. As discussed in this section, we source data from Kinne (2020) and Palmer et al. (2022) to construct a network spanning the years 2000-2010, where positive ties represent Defense Cooperation Agreements (DCAs) and negative ties Militarized Interstate Disputes (MIDs) between states. A snapshot of the resulting network, as observed in 2005, is presented in Figure 2, and yearly descriptive summaries of the exogenous information are provided in Table 2.

For this application, we also use additional data to construct our exogenous covariates. The information underlying these variables, as well as the MID data, are sourced from the `peacesciencer` package (Miller, 2022), but the original data sources are as follows: We measure countries' absolute political difference (Abs. Polity Diff.) using their polity scores (Marshall et al., 2018), their relative military power by taking the ratio of their Composite Indicators of National Capabilities* (CINC Ratio; Singer et al., 1972), their difference in wealth via their absolute GDP difference (Abs. GDP Diff; Anders et al., 2020) and obtain

*We use the higher CINC value in the ratio's numerator.

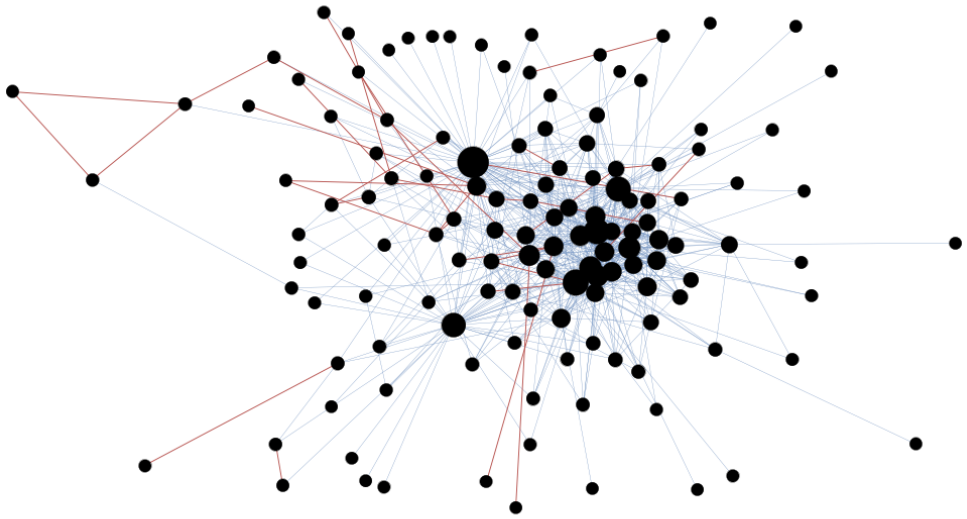


Figure 2: Network of MIDs (red) and DCAs (blue) in 2005. The size of each node relates to the degree (positive plus negative) of the respective country.

Table 2: Descriptive information about the exogenous dyadic covariates.

Year	Avr. Polity	Var. Polity	Avr. CINC	Var. CINC	Avr. GDP	Var. GDP
2001	2.925	42.659	0.006	<0.001	24.853	3.637
2002	3.082	42.829	0.006	<0.001	24.893	3.619
2003	3.184	43.356	0.006	<0.001	24.935	3.601
2004	3.116	43.487	0.006	<0.001	24.981	3.615
2005	3.211	44.332	0.006	<0.001	25.043	3.612
2006	3.429	42.986	0.006	<0.001	25.110	3.614
2007	3.510	42.923	0.006	<0.001	25.176	3.611
2008	3.497	42.183	0.006	<0.001	25.238	3.623
2009	3.592	41.860	0.006	<0.001	25.291	3.607
2010	3.687	40.641	0.006	<0.001	25.333	3.576

their geographical distance from Schvitz et al. (2022), log-transforming it before inclusion (Abs. Distance). For each covariate, we separately estimate effects on the propensity of a positive and negative edge. The only exception to this rule is the effect of the absolute distance, which is assumed to be equal for both types of edges. From 164 countries that participated in at least one MID or DCA between 2000 and 2010, we excluded 15 countries where covariate information was missing.

We now briefly discuss the estimation results for these covariates, as reported for Model 1 in Table 1 of the main article. These estimates are ceteris paribus, i.e., when accounting

for network dependencies via the endogenous terms. Regarding cooperation, countries are found to be more likely to formally work together via defense cooperation agreements if they are politically more similar, more comparable in their wealth, but also differ more in their material military capabilities. In particular, the first result is in line with previous research showing that similar regimes are more likely to ally (Lai and Reiter, 2000; Warren, 2016) while the second indicates that for DCAs, which regulate activities such as the joint research and development of military technology, countries' economic match also plays a role. The finding that countries are more likely to cooperate as their CINC ratio increases indicates, instead, that DCAs also follow a hierarchical structure where powerful states enter agreements with less powerful ones (Lake, 2009). In contrast, we find that states are more likely to fight when their CINCs, and thus military capabilities, are more similar, while their differences in regime type and wealth do not matter. Finally, the absolute distance between countries has a positive coefficient, suggesting, surprisingly, that the further away countries are from each other, the more likely they are to interact.

3.2 Goodness-of-Fit Assessment

3.2.1 Comparison between Model 1 and 2

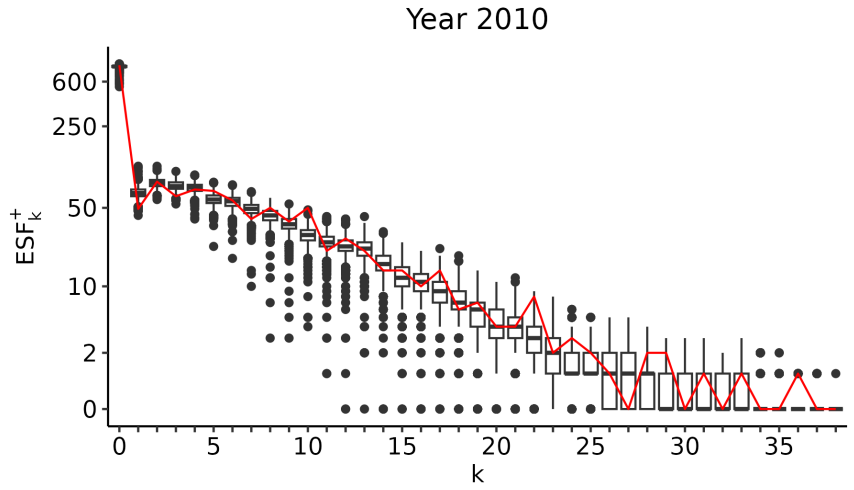
We also plot the results of the model assessment for Models 1 and 2 in the year 2010 in Figure 3. Since the degree distribution is captured in the same way in both models, we only focus on the distribution of edgewise-shared friends and enemies. According to Figure 3 (a) and (b), both models can capture the distribution of edgewise-shared friends. However, Figure 3 (c) and (d) indicates that the edgewise-shared enemies distribution is captured better in Model 1. This is in line with the comparison of AIC values reported in the main manuscript, which clearly indicates the superiority of Model 1 and thus of incorporating the number of common partners via simultaneous statistics.

3.2.2 Yearly Assessment

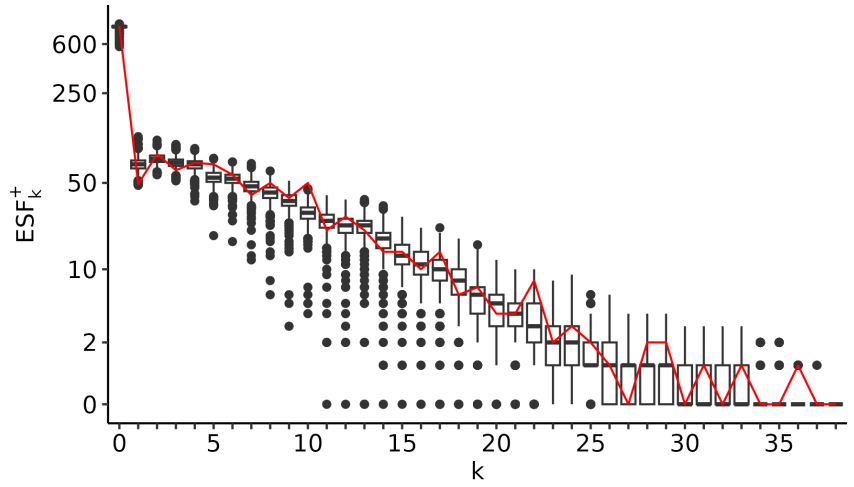
In the main article, the goodness-of-fit assessment was only shown for the year 2010. One can, however, carry out the same assessment for all years, as shown in Figures 4 to 12. These are substantively in line with the figure for the year 2010.

3.3 MCMC Diagnostics

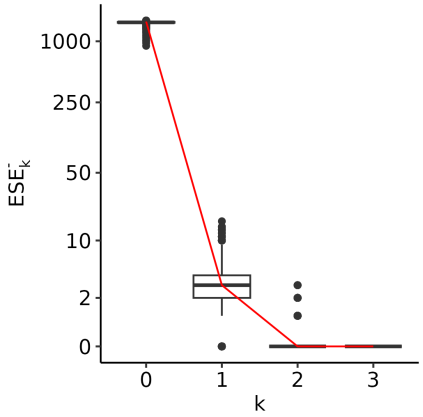
In figures 13–17, we present some diagnostic plots of the MCMC chain used in the final iteration of Model 1 in Section 4 of the main article. We average the Markov chain of each sufficient statistic around its observed value for better readability. Overall, one can observe that the model's estimates converged, are not degenerate, and are equal to the maximum likelihood estimates since the Markov chain oscillates around 0.



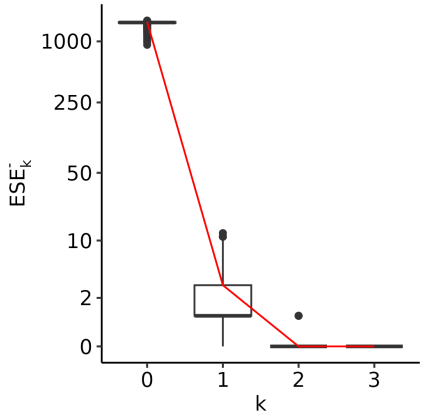
(a) Model 1



(b) Model 2



(c) Model 1



(d) Model 2

Figure 3: Goodness-of-fit assessment in the year 2010 for Models 1 and 2.

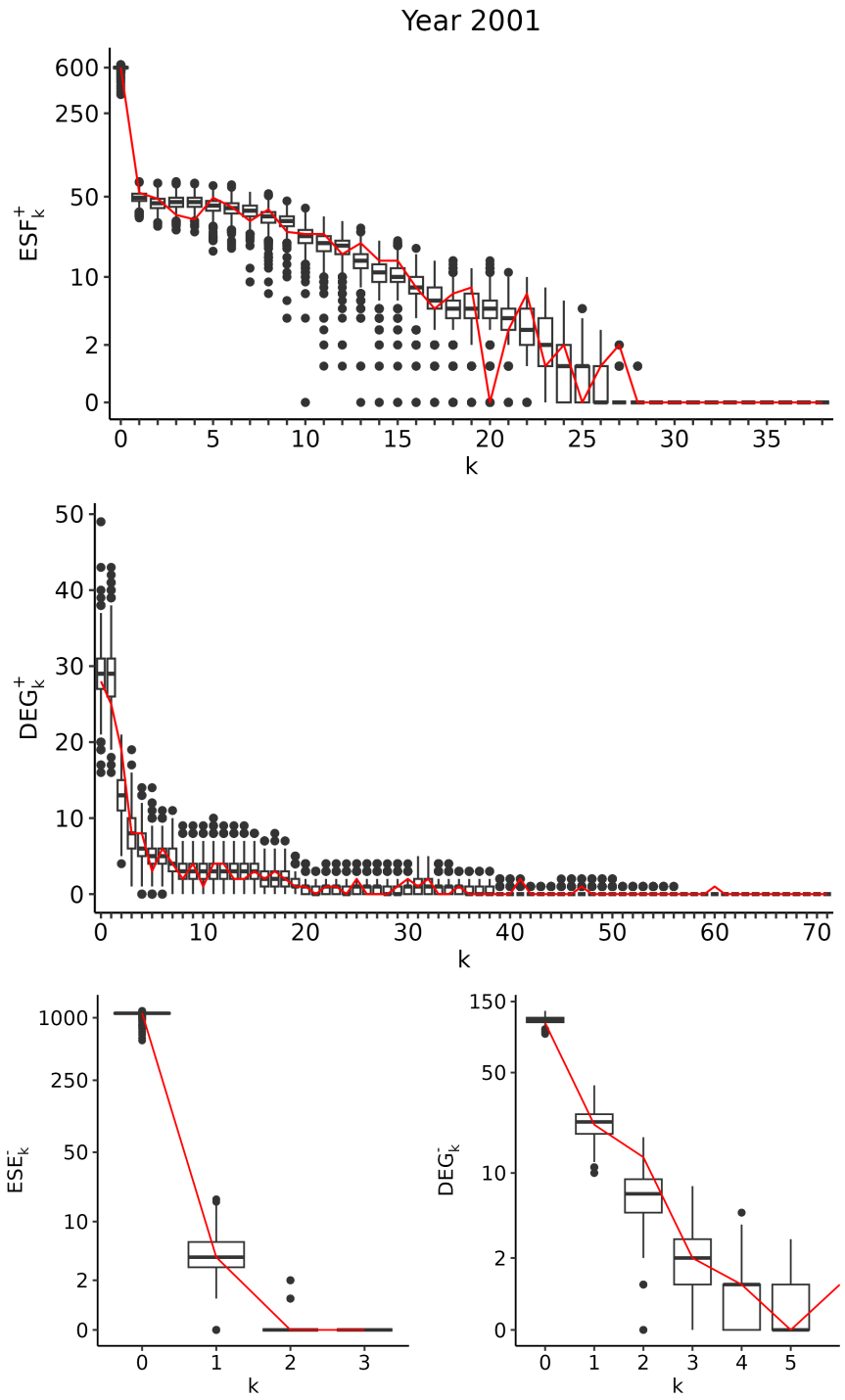


Figure 4: Goodness-of-fit assessment in the year 2001.

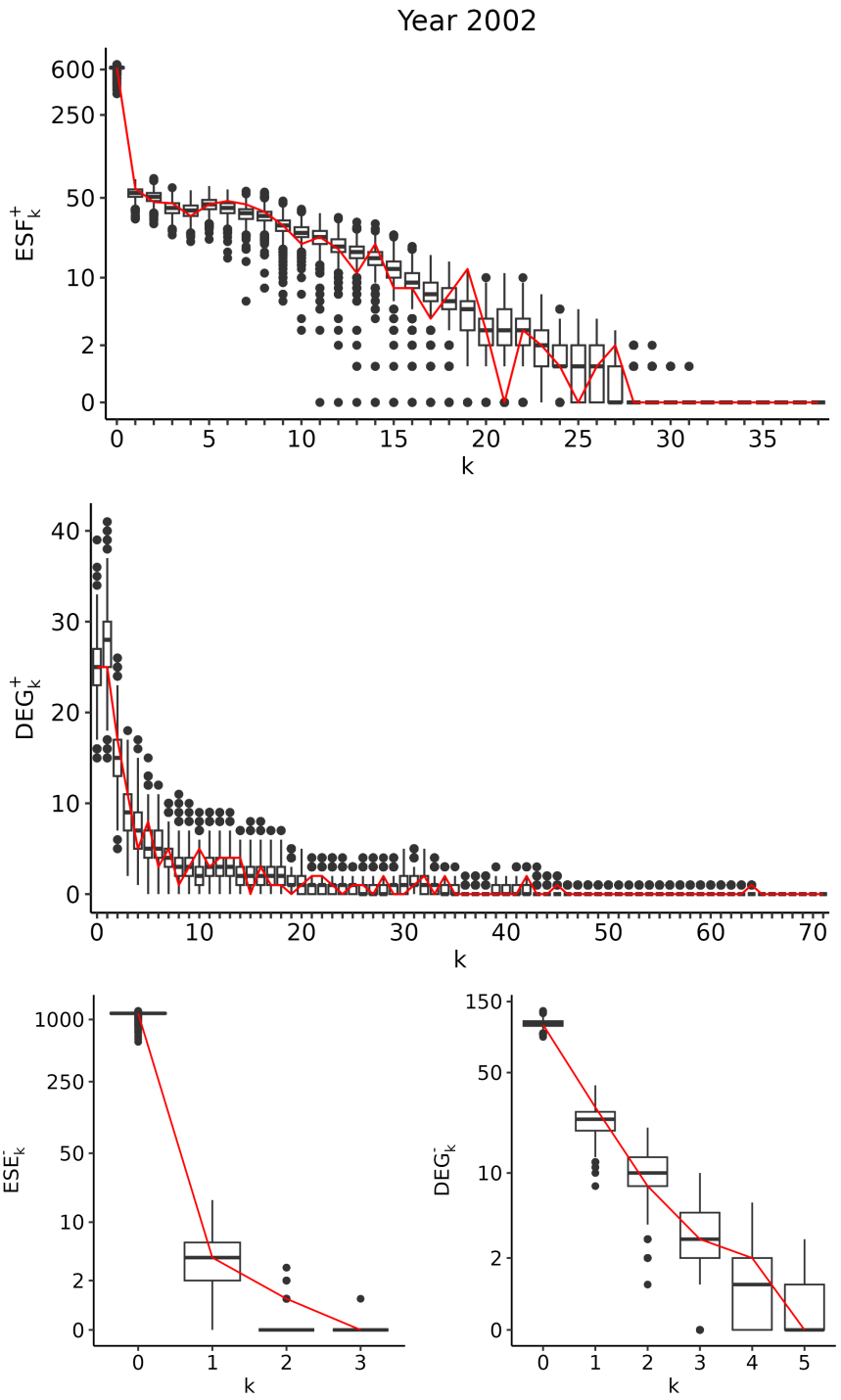


Figure 5: Goodness-of-fit assessment in the year 2002.

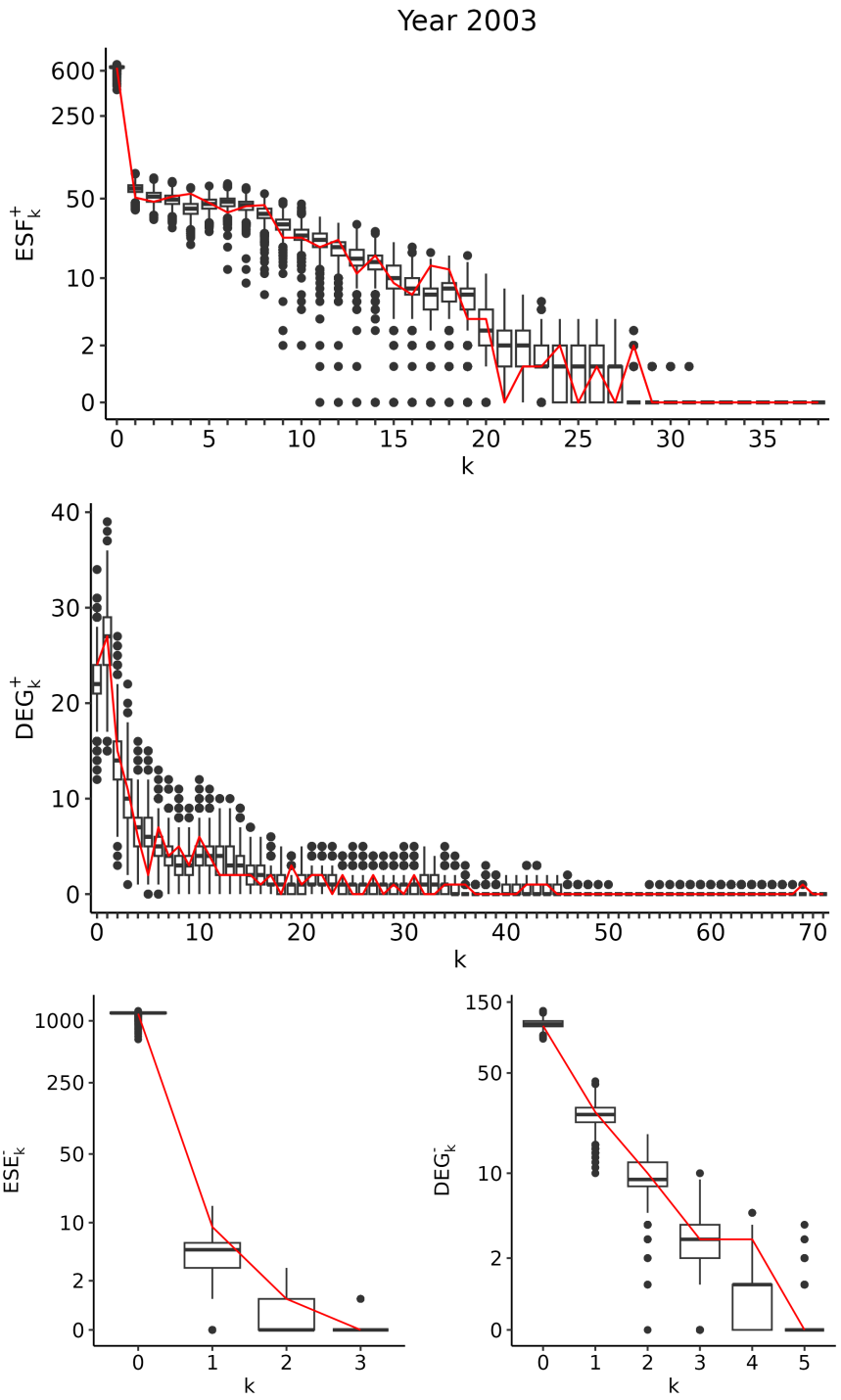


Figure 6: Goodness-of-fit assessment in the year 2003.

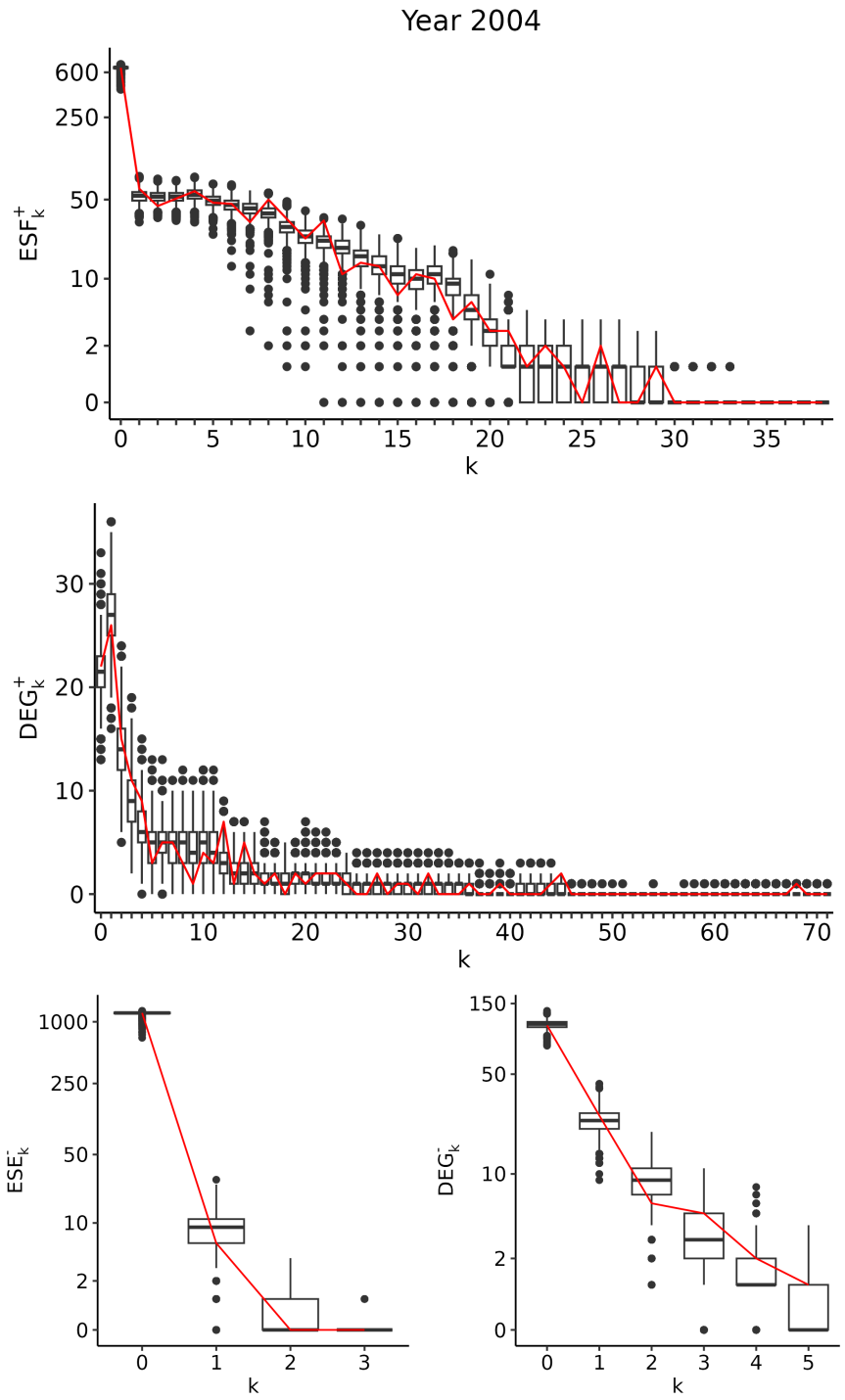


Figure 7: Goodness-of-fit assessment in the year 2004.

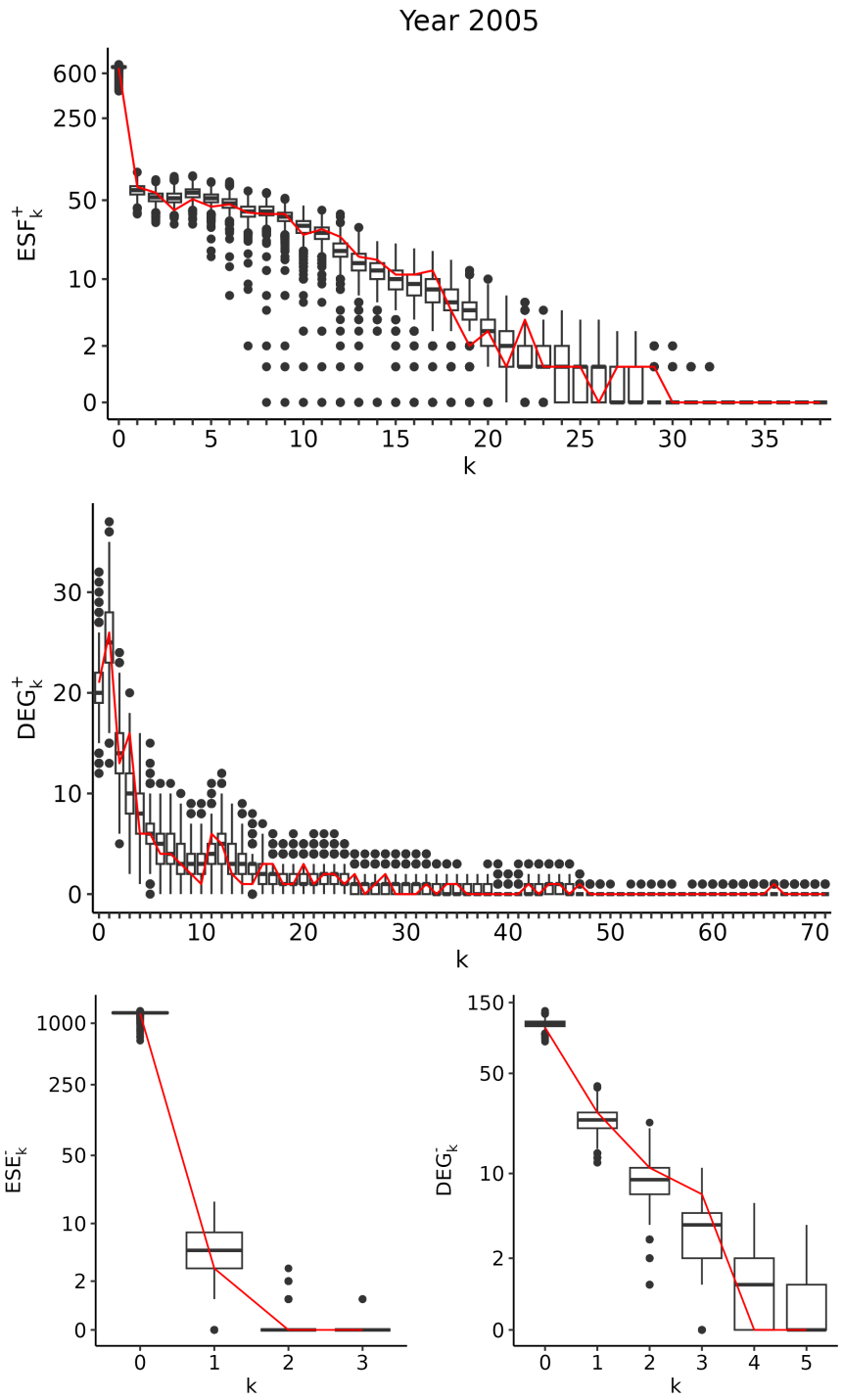


Figure 8: Goodness-of-fit assessment in the year 2005.

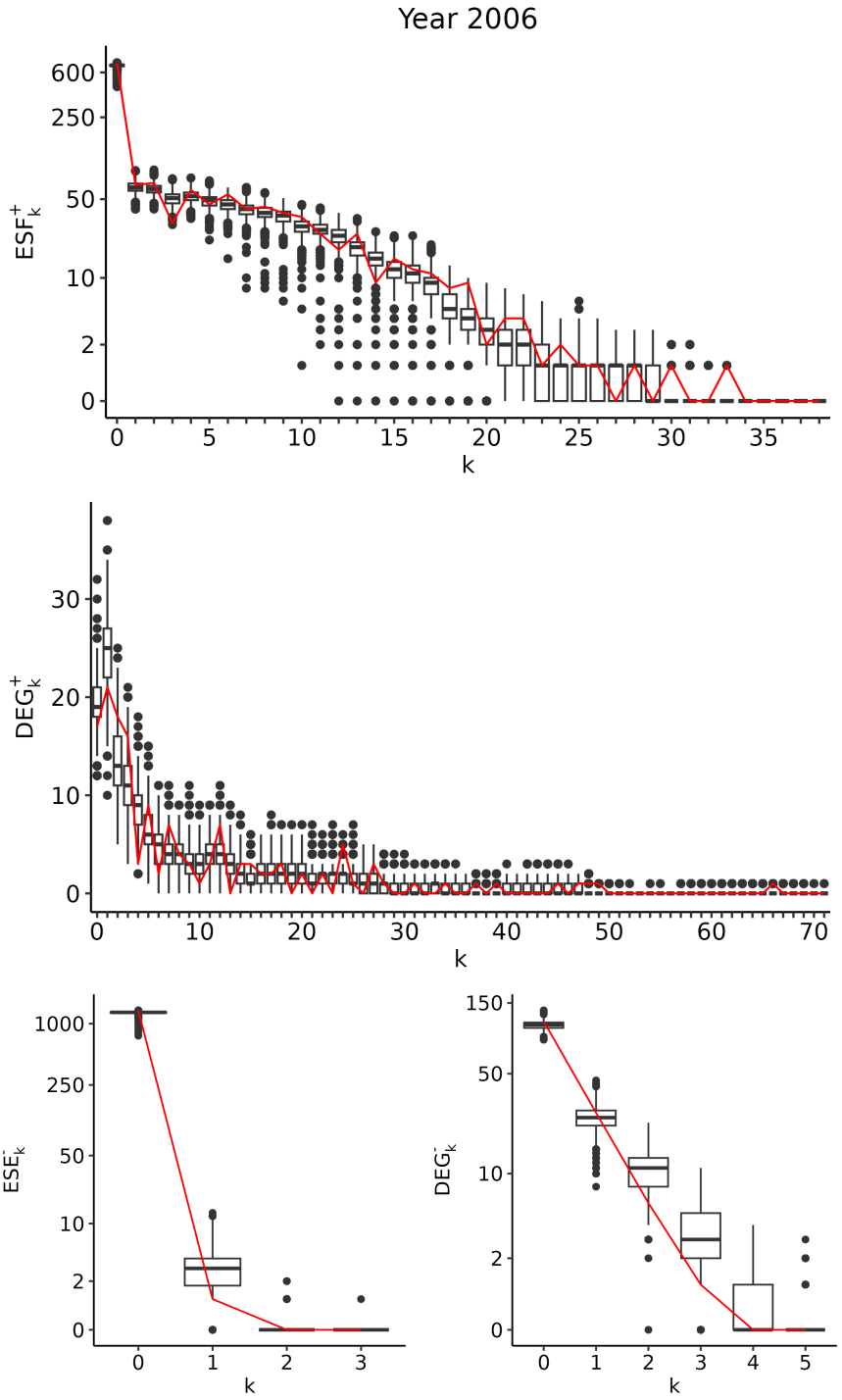


Figure 9: Goodness-of-fit assessment in the year 2006.

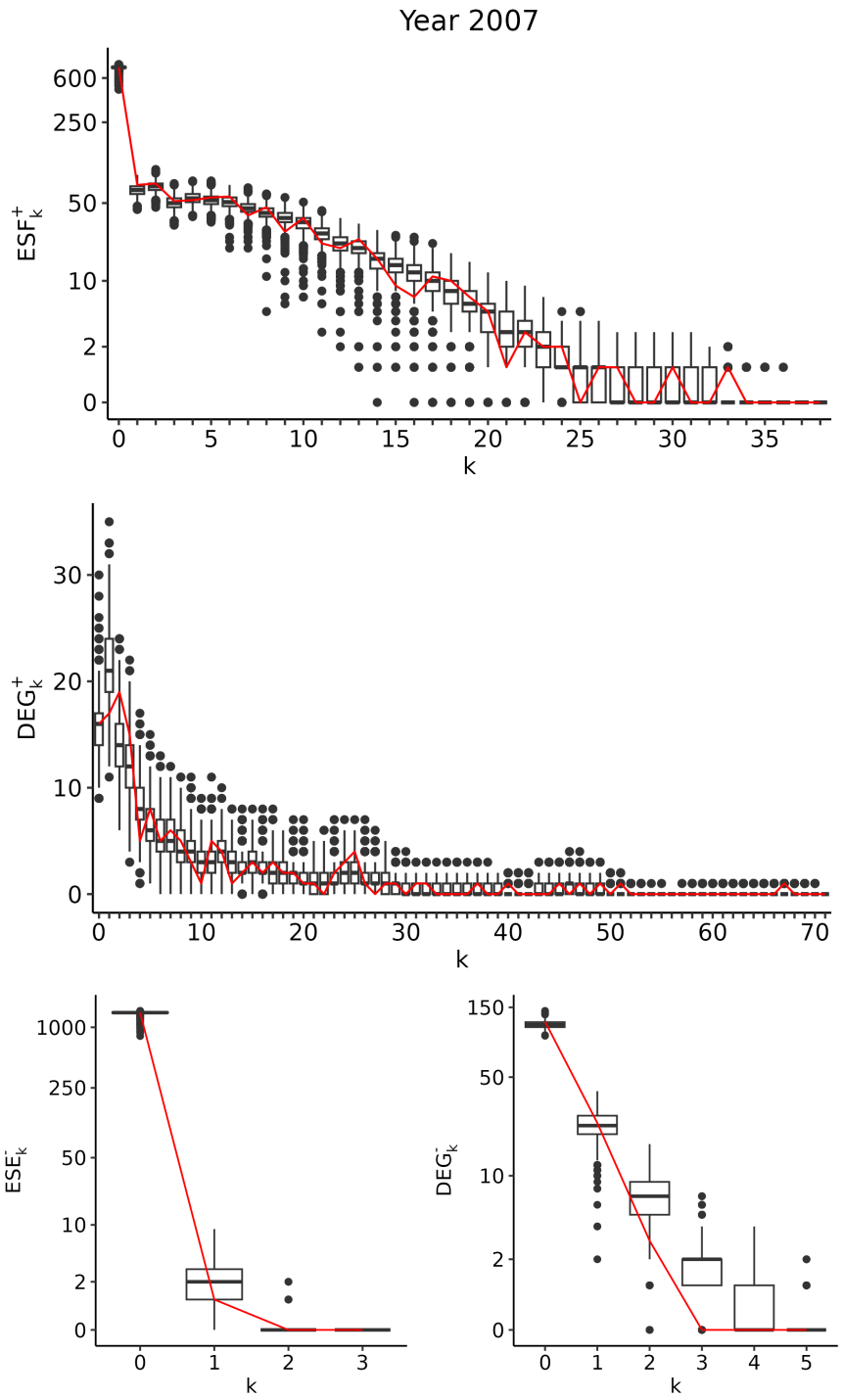


Figure 10: Goodness-of-fit assessment in the year 2007.

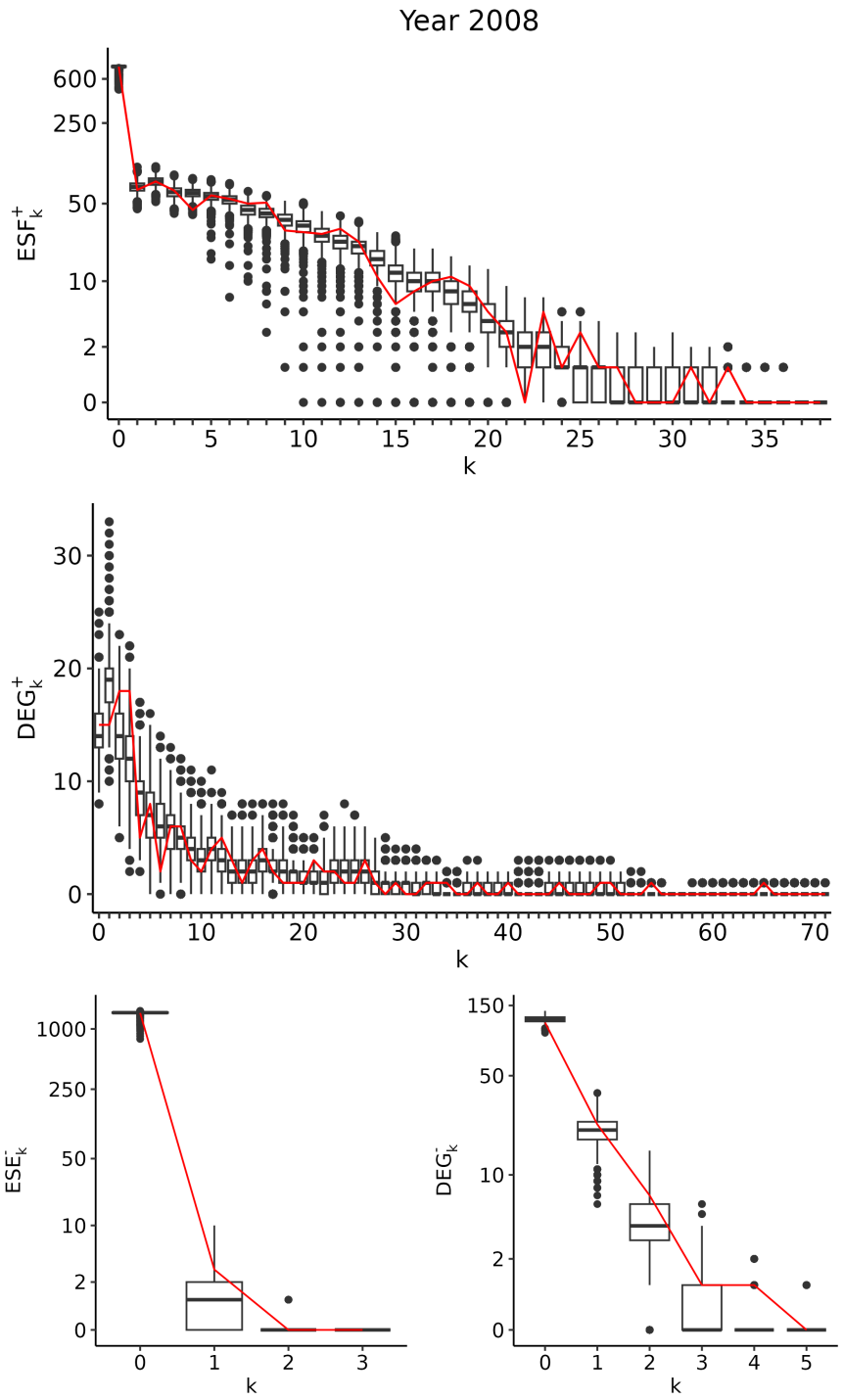


Figure 11: Goodness-of-fit assessment in the year 2008.

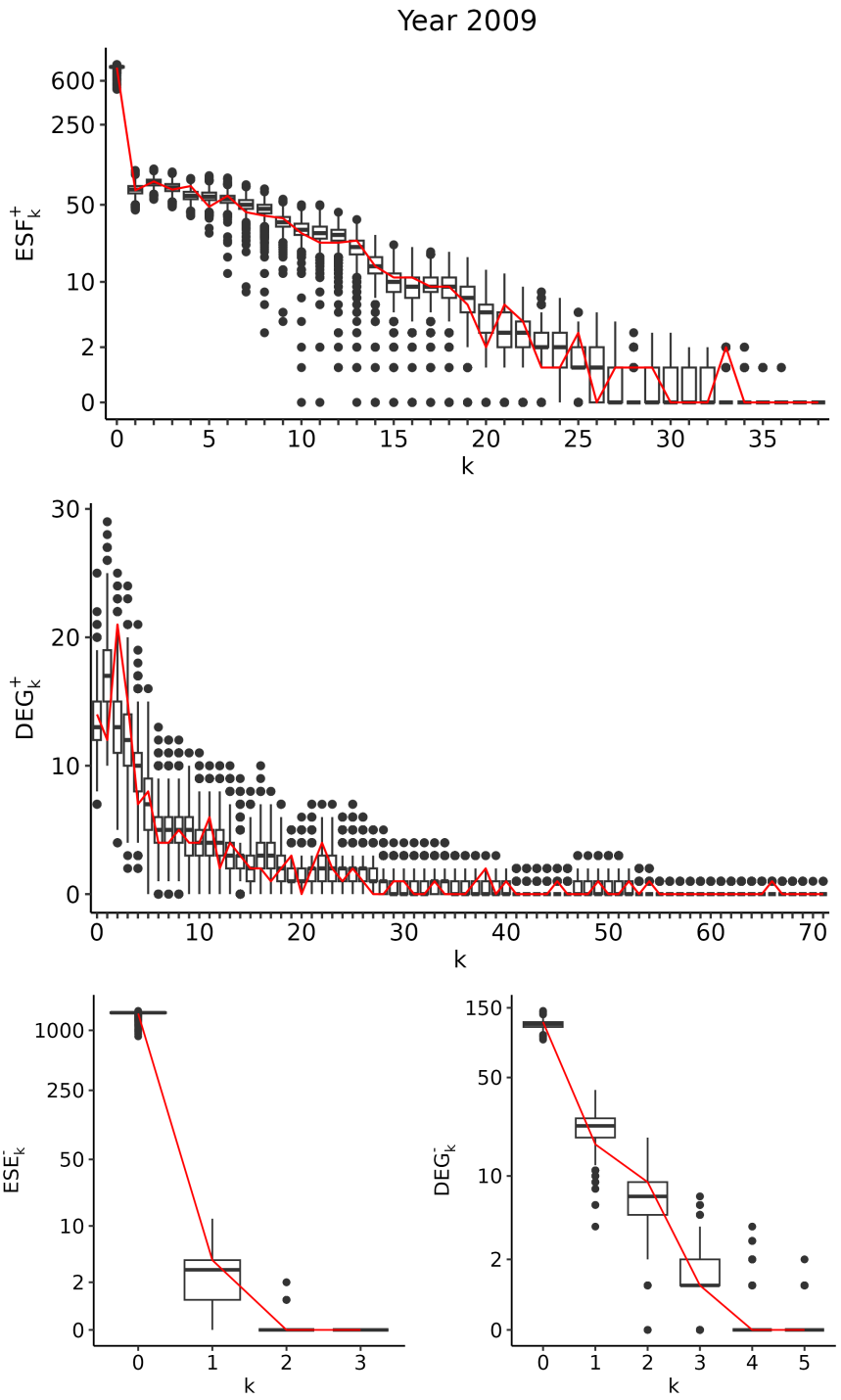


Figure 12: Goodness-of-fit assessment in the year 2009.

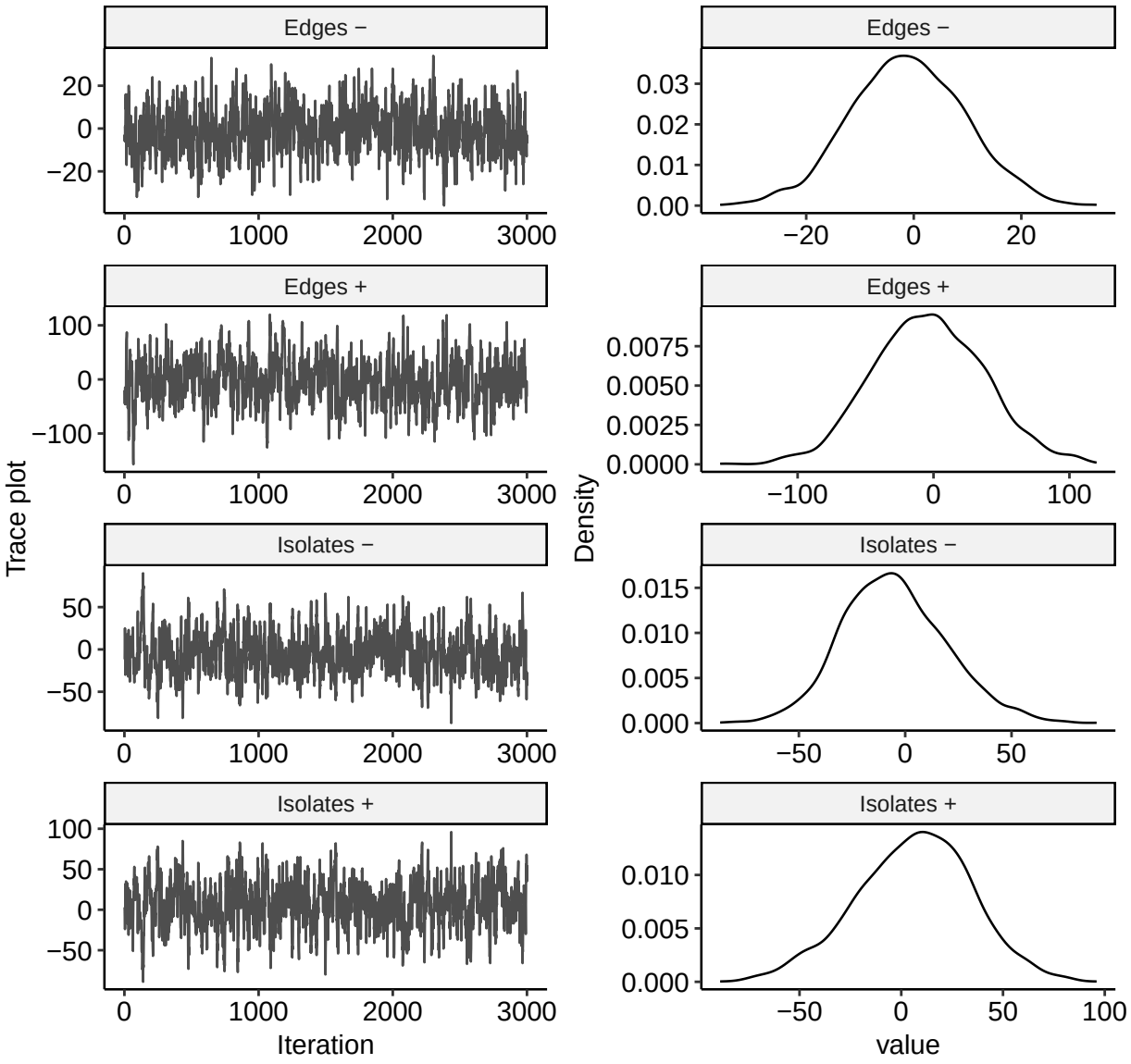


Figure 13: MCMC diagnostics of Model 1.

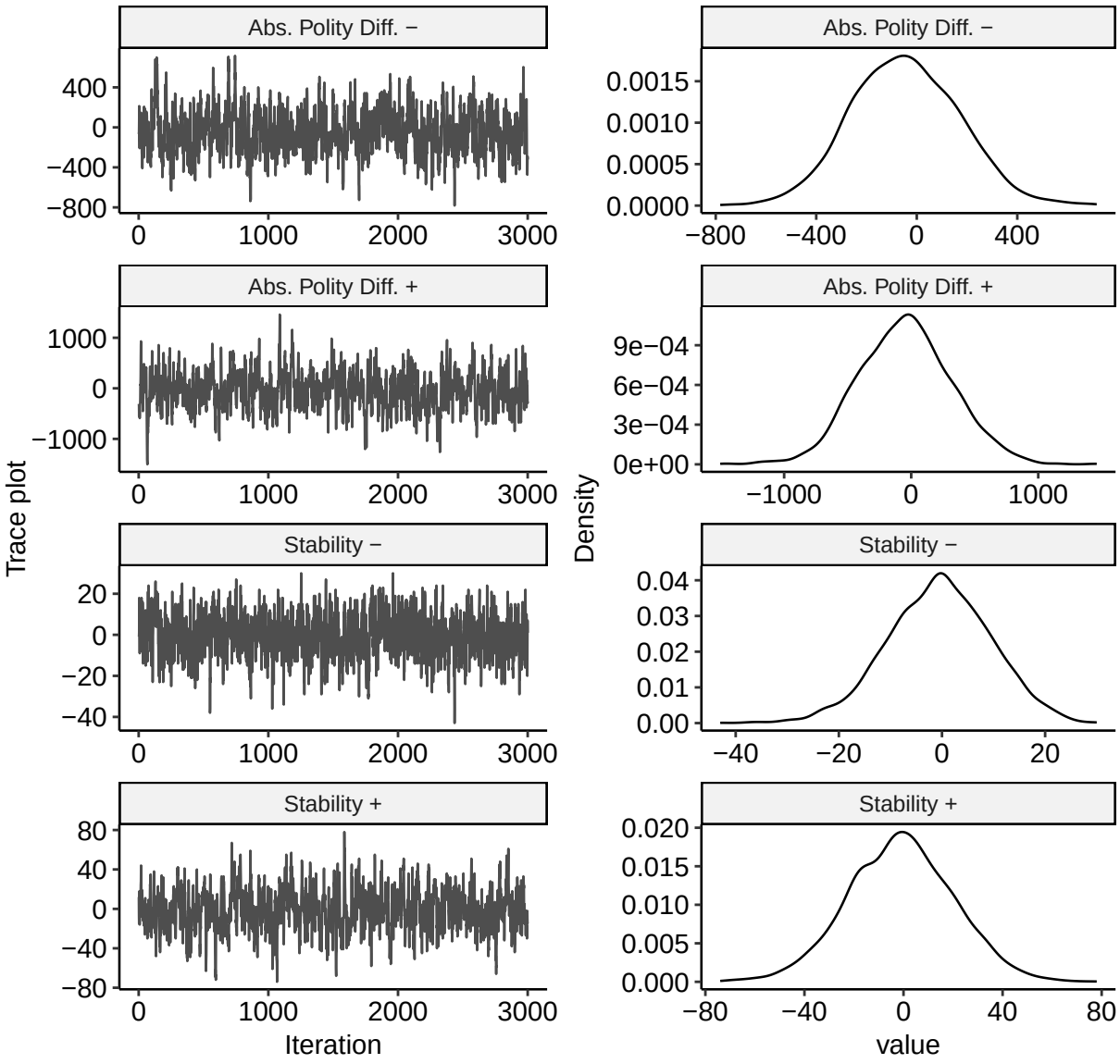


Figure 14: MCMC diagnostics of Model 1.

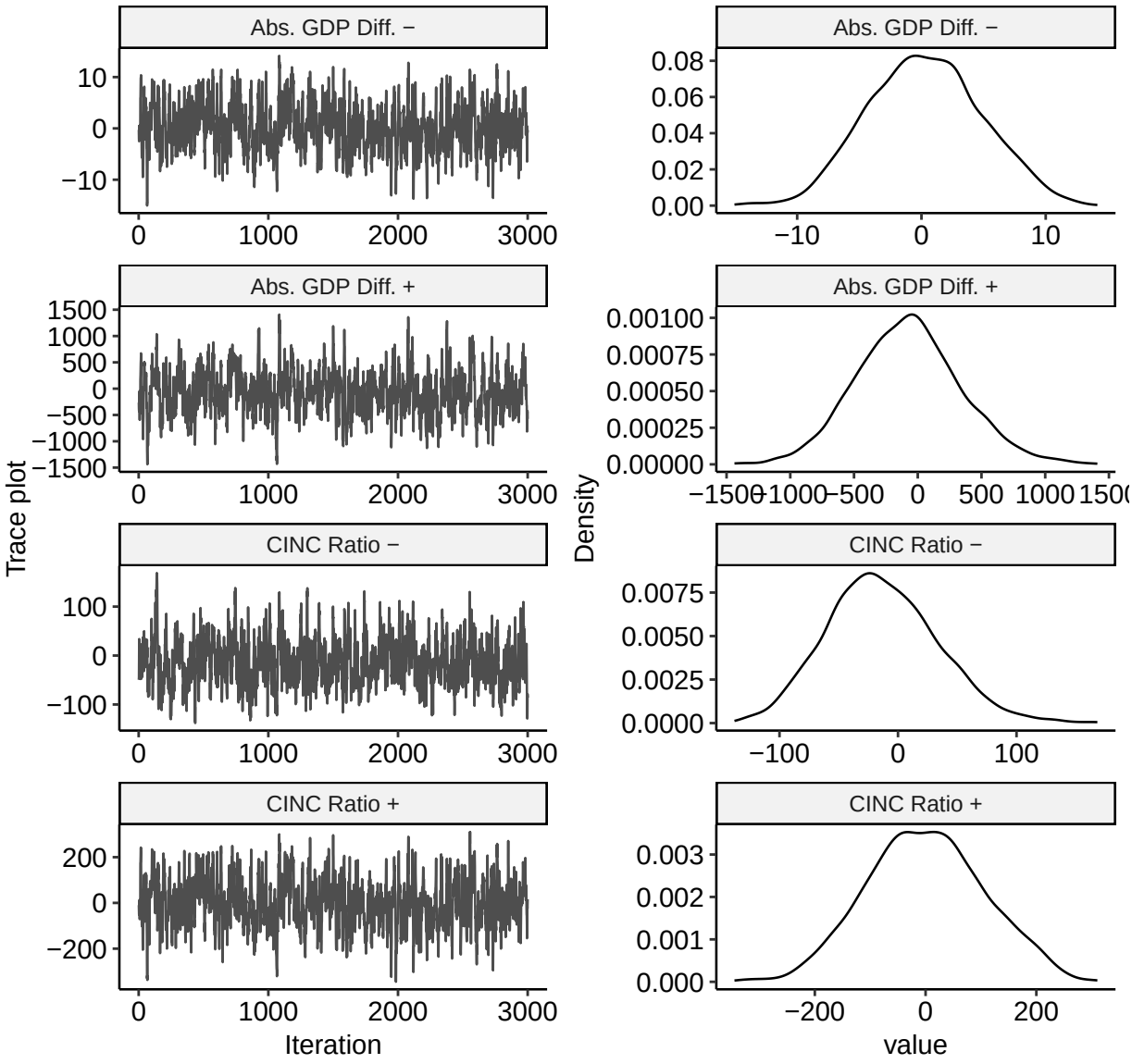


Figure 15: MCMC diagnostics of Model 1.

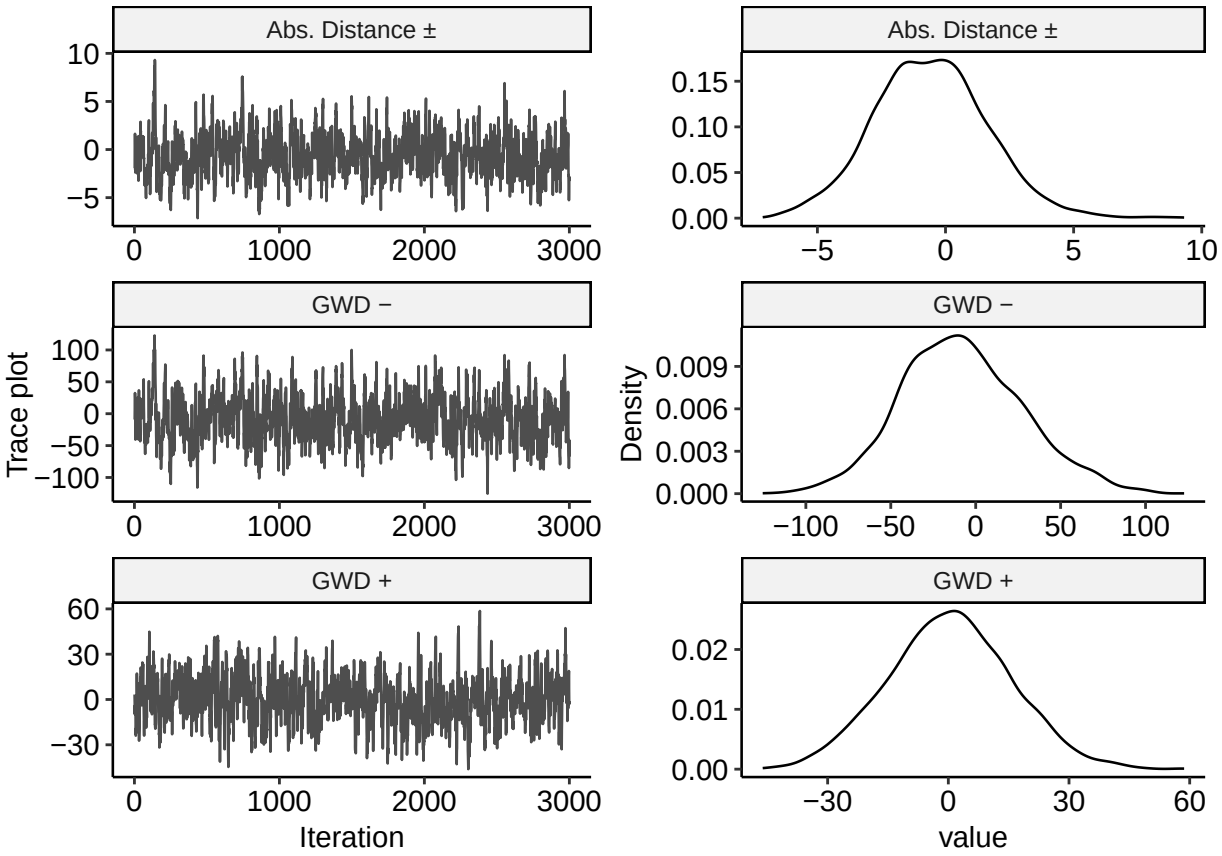


Figure 16: MCMC diagnostics of Model 1.

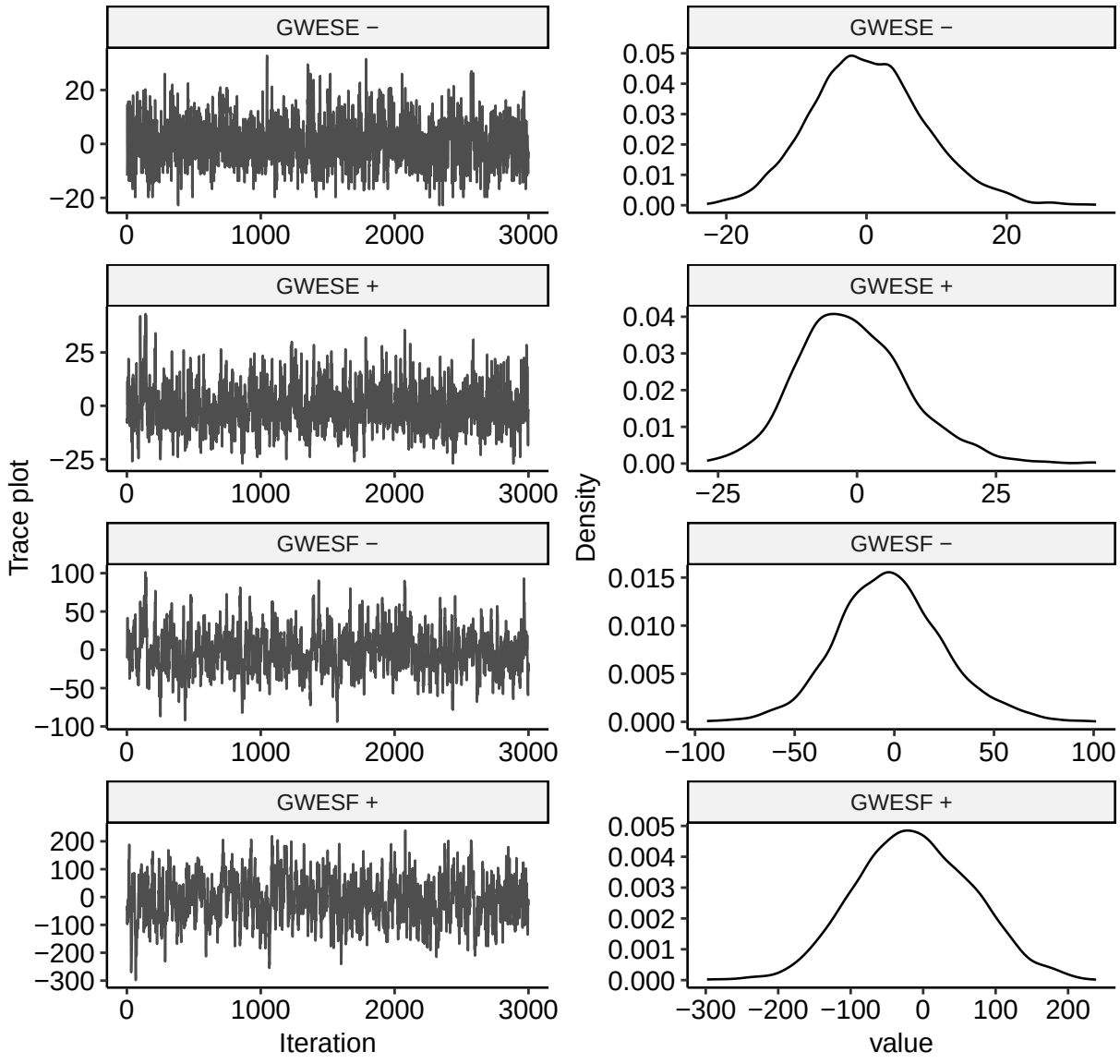


Figure 17: MCMC diagnostics of Model 1.

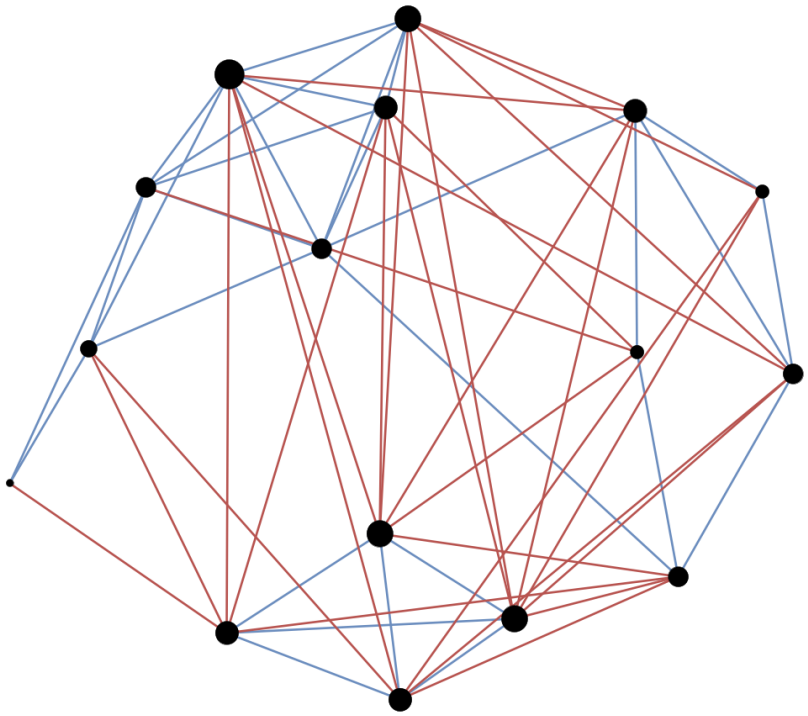


Figure 18: Network of enmity (red) and friendship (red) among New Guinean Highland tribes. The size of each node relates to the degree (positive plus negative) of the respective people.

4 Application to a static network: Enmity and Friendship among New Guinean Highland Tribes

4.1 Data Visualization and Sources

Next to dynamic networks one can also apply the SERGM to static networks. We demonstrate this with network data on interactions between the New Guinean Highland Tribes originally collected by Read (1954) and presented in Hage and Harary (1984). We source these data from the `R` package `signnet` (Schoch, 2020). The network covers relations of enmity and friendship among sixteen subtribes of the Gahuku-Gama, based on the anthropological work of Read (1954). Hage and Harary (1984) introduce it as an example of a network which is not perfectly balanced due to the existence of triads with zero or two positive ties but note that 82% of triads are balanced nonetheless. The full network is plotted in Figure 18. We now apply the SERGM to this static network to test whether structural balance effects can be recovered from it. The SERGM we specify includes edge terms, *GWESE*, *GWESF* as well as a degree statistic. To show the flexibility with which these statistics can be specified, we include the edge and *GWESF* terms separately for positive and negative ties, but the *GWESE* and *GWD* statistics only for positive ties. For

	Dependence		Independence	
	Coef.	CI	Coef.	CI
Edges +	-8.744	[-12.182,-5.306]	-0.76	[-1.13,-0.39]
Edges −	-1.647	[-2.766,-0.528]	-0.76	[-1.13,-0.39]
$GWESE^+$	0.45	[0.015,0.885]	-	
$GWESF^+$	0.068	[-0.228,0.364]	-	
$GWESF^-$	0.932	[0.616,1.248]	-	
GWD^+	5.492	[2.252,8.732]	-	
AIC	139.613		170.355	

Table 3: Results of the models applied to the New Guinean Highland Tribes.

the sake of comparison, we also estimate a model that drops all endogenous network terms and hence includes only the two edge terms. Results of both models are presented in Table 3.

4.2 Results and Model Assessment

In Table 3, it is apparent that the fully specified Dependence model has a lower AIC than the Independence model, which does not account for endogenous network terms, indicating that it is preferable in terms of performance. Table 3 also offers some evidence that structural balance drives tie formation among the Gahuku-Gama: *GWESF* has a positive effect on positive ties whose 95%-Confidence Intervals clearly exclude zero while its effect on negative ties is very close to zero. This implies that here, friends of friends are indeed more often friends but not less often enemies than one would expect by chance. $GWESE^+$ also exhibits a positive and statistically significant effect, suggesting that subtribes with a common enemy are more likely to share an alliance than in a random network of the same size. Finally, the effect of GWD^+ is also positive and statistically significant, meaning that a subtribe's probability of gaining a further positive tie increases with the number of such ties it already has.

Figure 19 offers a visual assessment of the goodness-of-fit of the Dependence model. Here, we can see that while the observed network lines up quite well with the simulated ones in terms of Edgewise-Shared Friends, model fit is more problematic for Edgewise-Shared Enemies where the observed network is regularly outside the interquartile range of the simulated networks. Similarly, the model does not do a good job of capturing the observed network's negative degree distribution. Based on these plots, one may consider re-running the Dependence model while specifying *GWESE* for both positive and negative ties and including GWD^-. Nonetheless, this application demonstrates the possibility of using the SERGM for the analysis of static networks. At the same time, Figure 20 visualizes the goodness-of-fit of the independence model from Table 3. For this example, one can clearly see that incorporating the additional terms inducing dependence augments the fit of the edgewise-shared enemies and friends distributions.

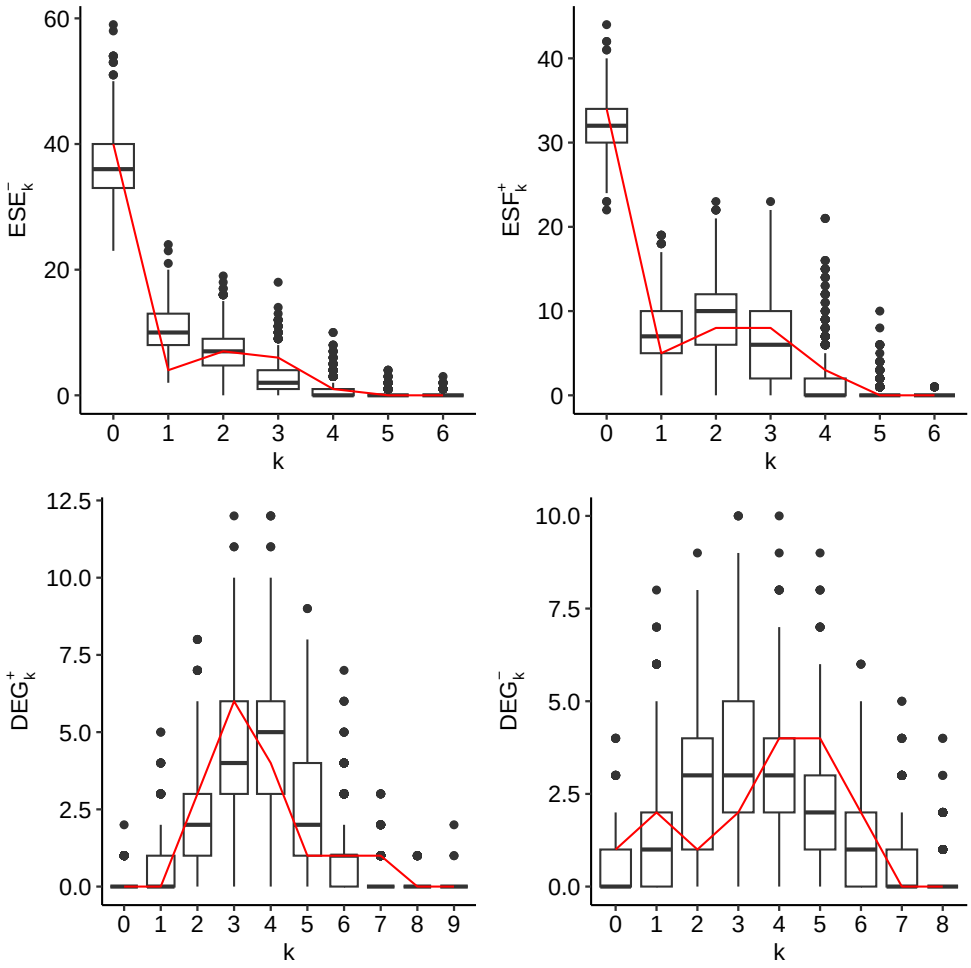


Figure 19: Model assessment of the dependence model of the New Guinean Highland Tribes.

4.3 MCMC Diagnostics

Finally, we also present the MCMC diagnostics for this additional application. Below are thus shown the MCMC trace plots for all its covariates (Figures 21–22). Overall, these MCMC diagnostics indicate a good convergence of the model.

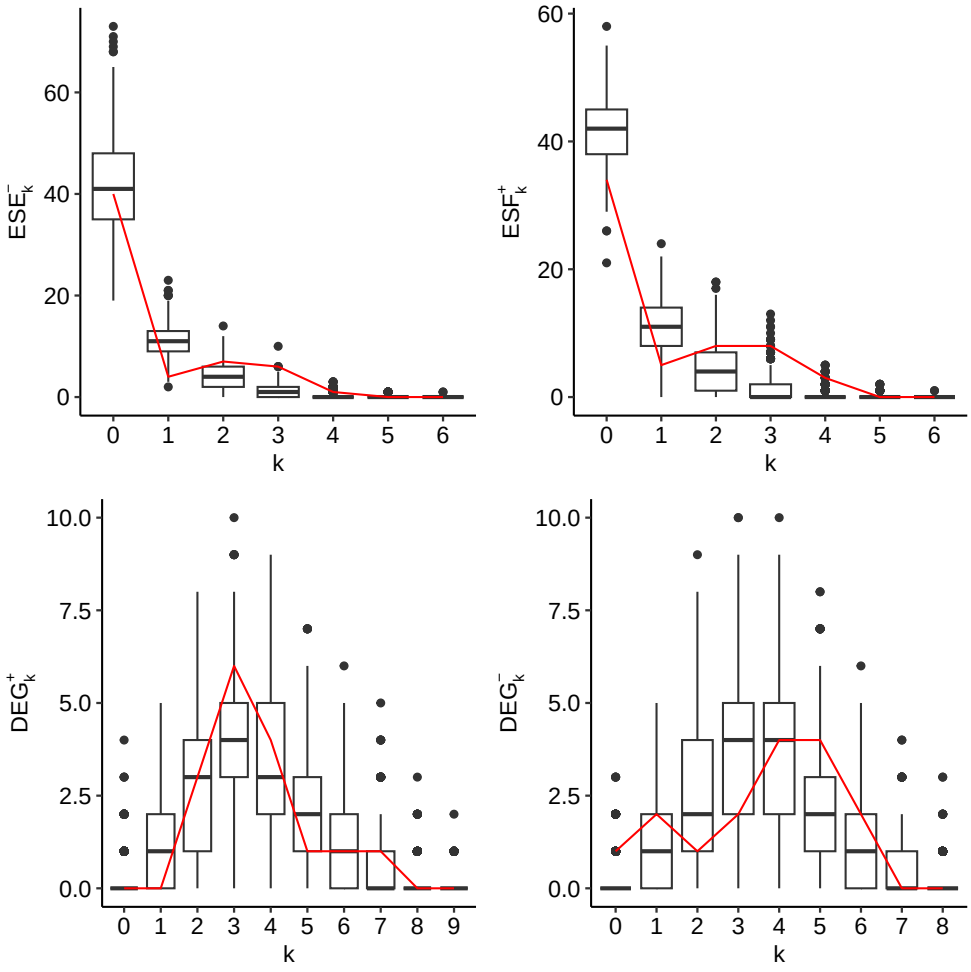


Figure 20: Model assessment of the independence model of the New Guinean Highland Tribes.

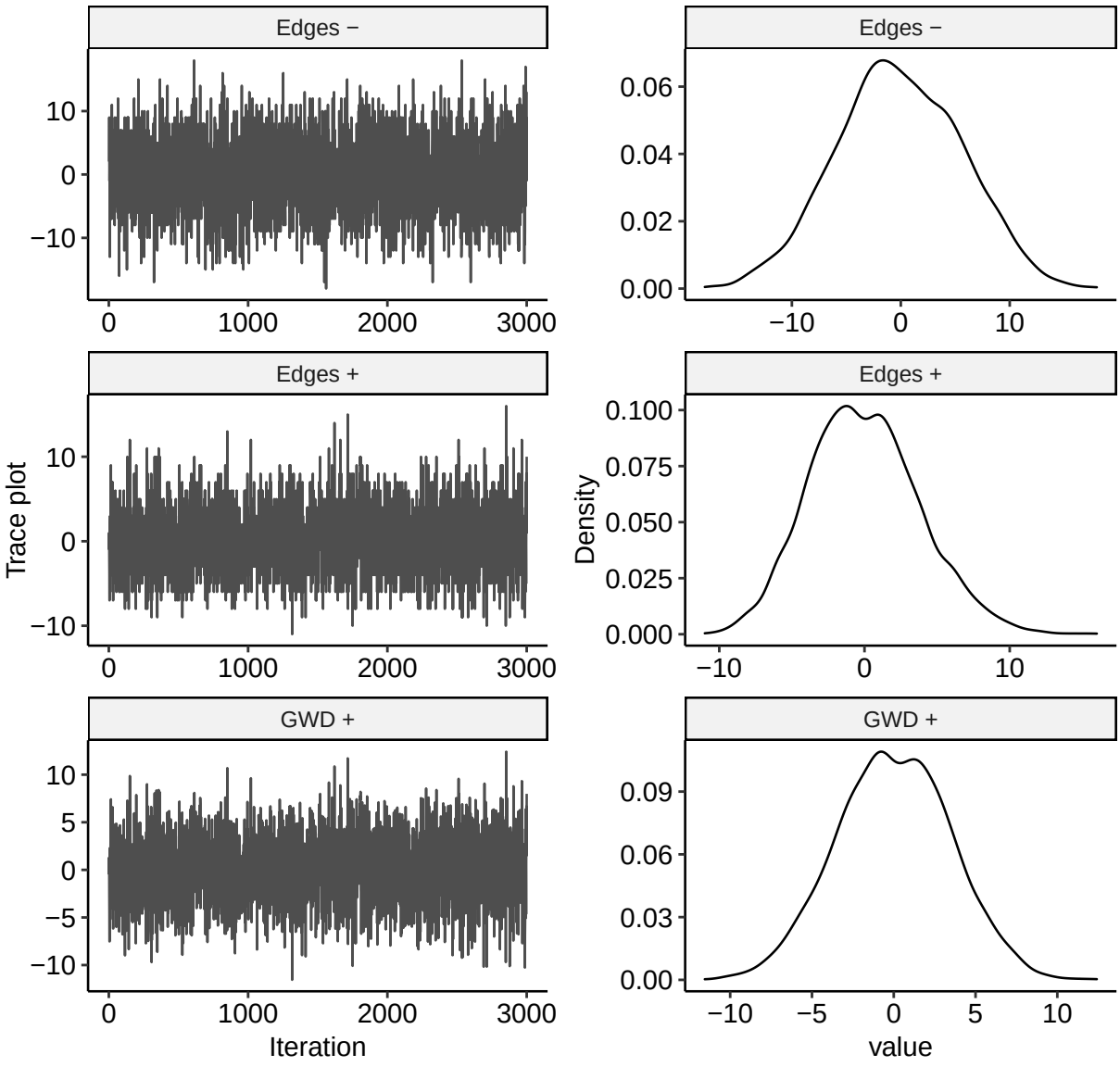


Figure 21: MCMC diagnostics of Dependence Model.

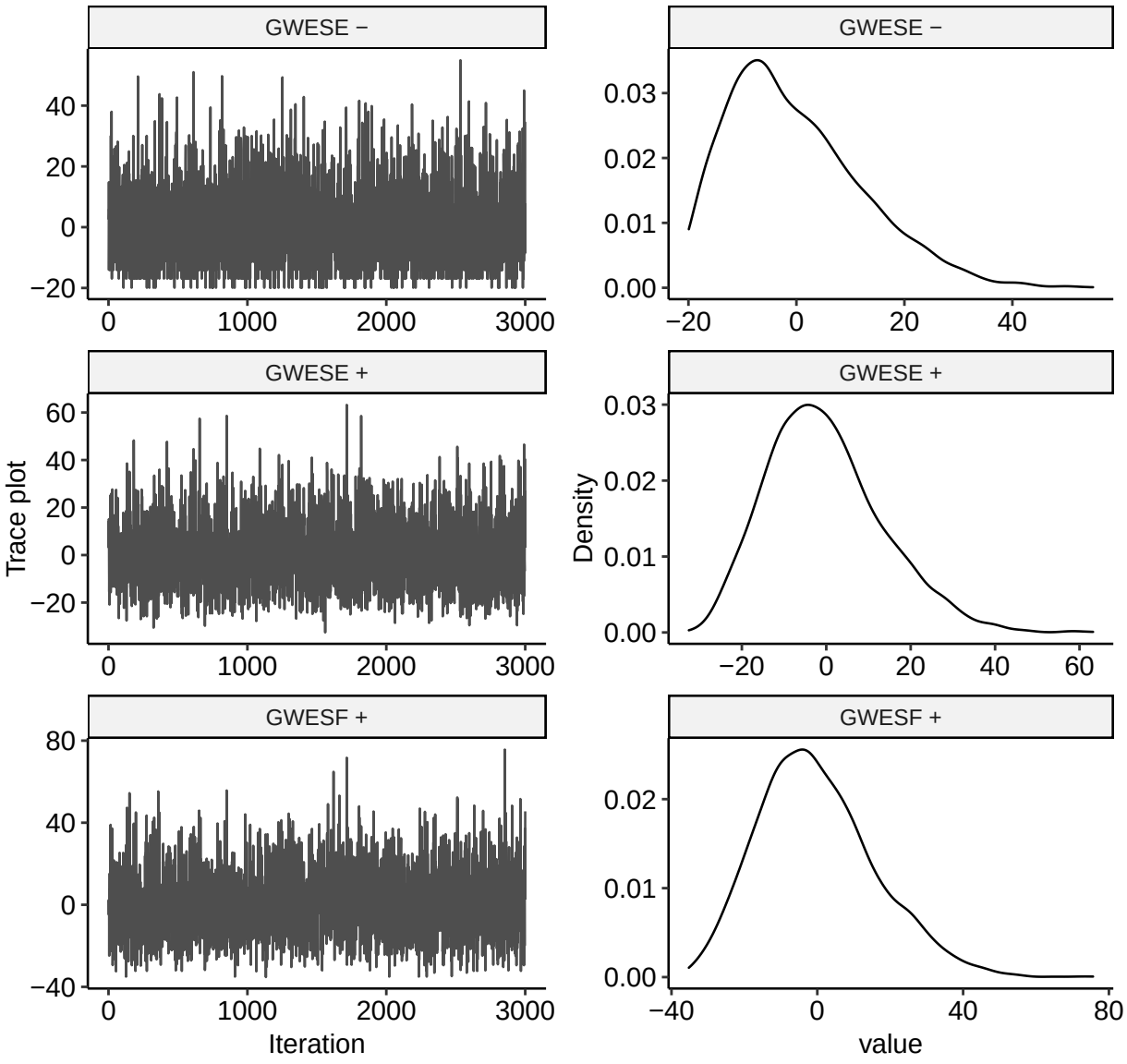


Figure 22: MCMC diagnostics of Dependence Model.

		Cooperation and Conflict	Tribes	Simulation
	Fixed α	log(2)	1.5	log(2)
	Grid size for γ	3,000	2,000	20
	Number of Bridges	16	16	–
Estimation	Burn-In	10,000	10,000	20,000
	MCMC Interval	1,000	1,000	1,000
	M	2,000	1,000	250
	Start Empty	False	False	True
Variance	Burn-In	100,000	10,000	20,000
	MCMC Interval	10,000	1,000	10,000
	M	3,000	3,000	500
	Start Empty	True	True	True
Bridge	Burn-In	100,000	10,000	–
	MCMC Interval	1,000	2,000	–
	M	2,000	1,000	–
	Start Empty	True	True	–

Table 4: Setting of the parameters of the fitting of the MCMC estimation procedure. One can define separate configurations for the Gibbs sampler used for the estimation of the parameters, the quantification of their variance, and the evaluation of the AIC via bridge sampling.

5 Computational Settings

We provide the set tuning parameters of the MCMC algorithm, the models for both applications, and the simulation study reported in Section 2 in Table 4. We performed sensitivity checks to ensure that the reported findings do not depend on the fixed parameters. In general, one should choose the values based on the density (the lower the density, the higher the burn-In and MCMC interval), size (the larger the size, the higher the burn-In and MCMC interval), and the strength of exogenous covariates (the stronger the influence of exogenous factors, the higher the burn-In and MCMC interval) of the network.

Supplementary Material References

Anders, T., C. J. Fariss, and J. N. Markowitz (2020). Bread before guns or butter: introducing surplus domestic product (sdp). *International Studies Quarterly 64*(2), 392–405.

Barndorff-Nielsen, O. (1978). *Information and exponential families in statistical theory.* New York: Wiley.

Gelman, A. and X. L. Meng (1998). Simulating normalizing constants: From importance sampling to bridge sampling to path sampling. *Statistical Science 13*(2), 163–185.

Hage, P. and F. Harary (1984). *Structural Models in Anthropology*. Cambridge: Cambridge University Press.

Hummel, R. M., D. R. Hunter, and M. S. Handcock (2012). Improving simulation-based algorithms for fitting ERGMs. *Journal of Computational and Graphical Statistics 21*(4), 920–939.

Hunter, D. R. and M. S. Handcock (2006). Inference in curved exponential family models for networks. *Journal of Computational and Graphical Statistics 15*(3), 565–583.

Kinne, B. J. (2020). The defense cooperation agreement dataset (dcad). *Journal of Conflict Resolution 64*(4), 729–755.

Krivitsky, P. N., A. R. Kuvelkar, and D. R. Hunter (2023). Likelihood-based Inference for Exponential-Family Random Graph Models via Linear Programming. *Electronic Journal of Statistics 17*(2), 3337–3356.

Lai, B. and D. Reiter (2000). Democracy, political similarity, and international alliances, 1816-1992. *Journal of Conflict Resolution 44*(2), 203–227.

Lake, D. A. (2009). *Hierarchy in International Relations*. Ithaca: Cornell University Press.

Marshall, M. G., T. R. Gurr, and K. Jaggers (2018). Polity iv project: Political regime characteristics and transitions, 1800–2017. *Center for systemic peace*.

Miller, S. V. (2022). {peacesciencer}: An r package for quantitative peace science research. *Conflict Management and Peace Science 39*(6), 755–779.

Palmer, G., R. W. McManus, V. D'Orazio, M. R. Kenwick, M. Karstens, C. Bloch, N. Dietrich, K. Kahn, K. Ritter, and M. J. Soules (2022). The MID5 dataset, 2011–2014: Procedures, coding rules, and description. *Conflict Management and Peace Science 39*(4), 470–482.

Read, K. E. (1954). Cultures of the central highlands, new guinea. *Southwestern Journal of Anthropology 10*(1), 1–43.

Schmid, C. S. and B. A. Desmarais (2017). Exponential random graph models with big networks: Maximum pseudolikelihood estimation and the parametric bootstrap. In *2017 IEEE International Conference on Big Data (Big Data)*, pp. 116–121.

Schmid, C. S. and D. R. Hunter (2023). Computing Pseudolikelihood Estimators for Exponential-Family Random Graph Models. *Journal of Data Science 21*(2), 295–309.

Schoch, D. (2020). *signnet: An R package to analyze signed networks*. https://github.com/schochastics/signnet.

Schvitz, G., L. Girardin, S. Rüegger, N. B. Weidmann, L.-E. Cederman, and K. S. Gleditsch (2022). Mapping the international system, 1886-2019: The cshapes 2.0 dataset. *Journal of Conflict Resolution 66*(1), 144–161.

Singer, J. D., S. Bremer, and J. Stuckey (1972). Capability distribution, uncertainty, and major power war, 1820-1965. In B. Russett (Ed.), *Peace, war, and numbers*, pp. 19–48. Beverly Hills: Sage.

van Duijn, M. A., K. J. Gile, and M. S. Handcock (2009). A framework for the comparison of maximum pseudo-likelihood and maximum likelihood estimation of exponential family random graph models. *Social Networks 31*(1), 52–62.

Warren, T. C. (2016). Modeling the coevolution of international and domestic institutions. *Journal of Peace Research 53*(3), 424–441.