Supplementary Materials- A Bayesian approach to burial networks to infer social changes in northeastern Taiwan during the European colonization period

Li-Ying Wang

Ben Marwick

21 January, 2021

# 1 Introduction to network analysis and archaeological applications

A social network is generally visualized as a graph consisting of a set of socially-relevant nodes/actors, connected by edges/ties representing one or more relations, such as friendship, collaborations, information flow, trade ties, or any other forms of connection of interest (Marin and Wellman, 2011; Wasserman et al., 1994). The ties can be classified into four major types, including similarities, social relations, interactions, and flows (Borgatti et al., 2009). In archaeology, actors can be people, groups, objects, places, or events, with ties built on similarity, proximity, or co-presence of material culture to create patterns reflecting influence, geographical distance, or affiliations in social groups (Brughmans and Peeples, 2018; Mills, 2017; Peeples, 2019). For example, past trade can be conceptualized as a network of individual entities connected by shared similarities, the flow of goods, to represent their interactions (Collar et al., 2015, p. 4). Similar concepts can be applied to a wide range of archaeological data with relational assumptions, such as burial contexts. Burial goods, especially high value goods, can reflect social practices in broader cultural contexts to represent personal wealth or social status from which we can infer social differentiation or complexity (Gamble and Zepeda, 2002; Janes, 2013). Burials with the same prestige goods could indicate some underlying social relations where individuals share similar access to trade, exchange, and gifting networks according to their status (eg. Coward, 2013, p. 252). This enables the exploration of the structure of the past social organization through the identification of the relationships among burials.

Network analysis has been increasingly applied by archaeologists in recent years to deal with past interactions and explore the underlying mechanisms. There are two common approaches to characterize network properties at two distinct scales: node/edge level and graph level (Peeples, 2019). Node level focuses on the role of node in a network, such as centrality, representing the individual influence or social prominence in a group, while graph level assesses the whole network attributes, such as density, clustering in a network, to generalize relationship patterns (Mills, 2017; Peeples, 2019). By quantifying those network properties, archaeologists can answer a wide range of research questions. Examples includes exploring the political centralization in the Kofun period in Japan through the hierarchical communication network constructed by prestige goods (Mizoguchi, 2013), the investigation of long term inter-site relationships from the Epipalaeolithic to the early Neolithic in the Near East according to trade items (Coward, 2013). Regarding burials, Sosna et al. (2013) examined spatial pattern of burials from the Early Bronze Age in Rebesovice with two hypothesized networks constructed according to cultural and chronological similarity between burials. Recently, complex network modeling evaluates networks at both node and graph level through simulations of particular processes and statistically testing the formation of network properties (Brughmans, 2013; Brughmans and Peeples, 2018; Freeman, 2004; Salvini, 2010). Such application includes simulations and testing food exchange modes for Ancestral Pueblos on the aggregation of households in the American Southwest simulation (Crabtree, 2015), or exploring the diffusion of fired bricks across Hellenistic Europe by comparing similarity networks of sites with random networks. Other examples show the assessment of hunter-gatherer exchange networks structure across the Kuril Islands using bootstrap simulation based on ceramic composition (Gjesfjeld, 2015).

Bayesian approaches to Exponential Random Graph Models (ERGMs) are effective ways to examine relationships by incorporating prior information into networks to better understand dependencies of variables in complex networks and improve computational issues in ERGMs (Caimo et al., 2017; Lehmann et al., 2020). Current approaches to network analysis used by archaeologists are mostly restricted to a single rational structure without consideration of interaction between network variables. Our use of Bayesian inference on ERGMs is the first application to archaeological data that can bring new insights to understand past social structures by characterizing network properties as a whole. ERGMs are an important family of statistical models for networks that allows direct modeling for the formation of edges, or ties, between nodes (Robins et al., 2007). The assumption is that possible ties in a network are random variables and dependent on actor variables or the presence or absence of other ties (Robins et al., 2007). In other words, networks in ERGMs are viewed as dependent variables, where network dependencies and the attributes of nodes/edges can influence the formation of a tie (Snijders, 2011). For example, nodes with similar attributes are more likely to form a relationship, such as friendship between people with the same hobby. Ties form a small structure in a network called a graph configuration, that describes the form of dependence, such as reciprocity (relationship between two actors), transitivity or clustering (relationship between two actors through a shared third actor), homophily (relationship between actors with a similar attribute), and popularity (actors have many relationships with others) (Morris et al., 2008; Robins et al., 2007; Snijders et al., 2006). Those configurations represent the structure or the property of a network and can be expressed by network statistics. By modeling those network statistics as direct functions of ties by specifying the forms of configurations, we can generate a distribution of random networks that represent our hypothesis-based model (Morris et al., 2008; Robins et al., 2007). Such a distribution consists of a large number of possible networks that enables statistical inference and comparison with an observed network (Robins et al., 2007). ERGMs helps us understand whether an observed network shows significantly more or less of a property of interest than the random networks generated from our model assumptions.

# 2 Kiwulan in the 17th northern Taiwan

The involvement in long-distance exchange is often associated with changes in political-economic strategies to a network mode, where ambitious individuals are able to build their local power base and personal influence through the distribution of high-value goods (Blanton et al., 1996; Feinman, 2000; Klehm, 2017). Northeastern Taiwan was involved in a regional exchange in East Asia since the 14th century and long-distance trade during the European presence in the 17th century (Chen, 2005; Wang and Liu, 2007). In 1626, the Spanish built their forts in Heping dao and Tamsui, northern Taiwan, and later were defeated and taken over by the Dutch in 1642 (Andrade, 2007). Indigenous people in northeastern Taiwan first encountered colonial power in 1632 when they were attacked by the Spanish and later in 1647 by the Dutch, who also asked them to pay annual tributes in a flexible way (Borao Mateo, 2009; Kang, 2016). The Dutch would give local Indigenous leaders symbolic items or trade goods to assert their political authority and consolidate their relationship with Indigenous societies.

# 3 Method

For trade beads with substantial differences in quantities between burials, we described each burial as having one of four levels according to their distributions across all burials (Figure 3.1). Gold-foil beads are in levels of high (>3), upper-middle (3), lower-middle (2), and low (1); carnelian beads are in levels of high (>6), upper-middle (4-6), lower-middle (3), and low (1-2); glass beads are in levels of high (>6), upper-middle (3-6), lower-middle (2), and low (1). If burial 1 and burial 2 both have high quantities of agate beads, then there will be a tie connecting them.

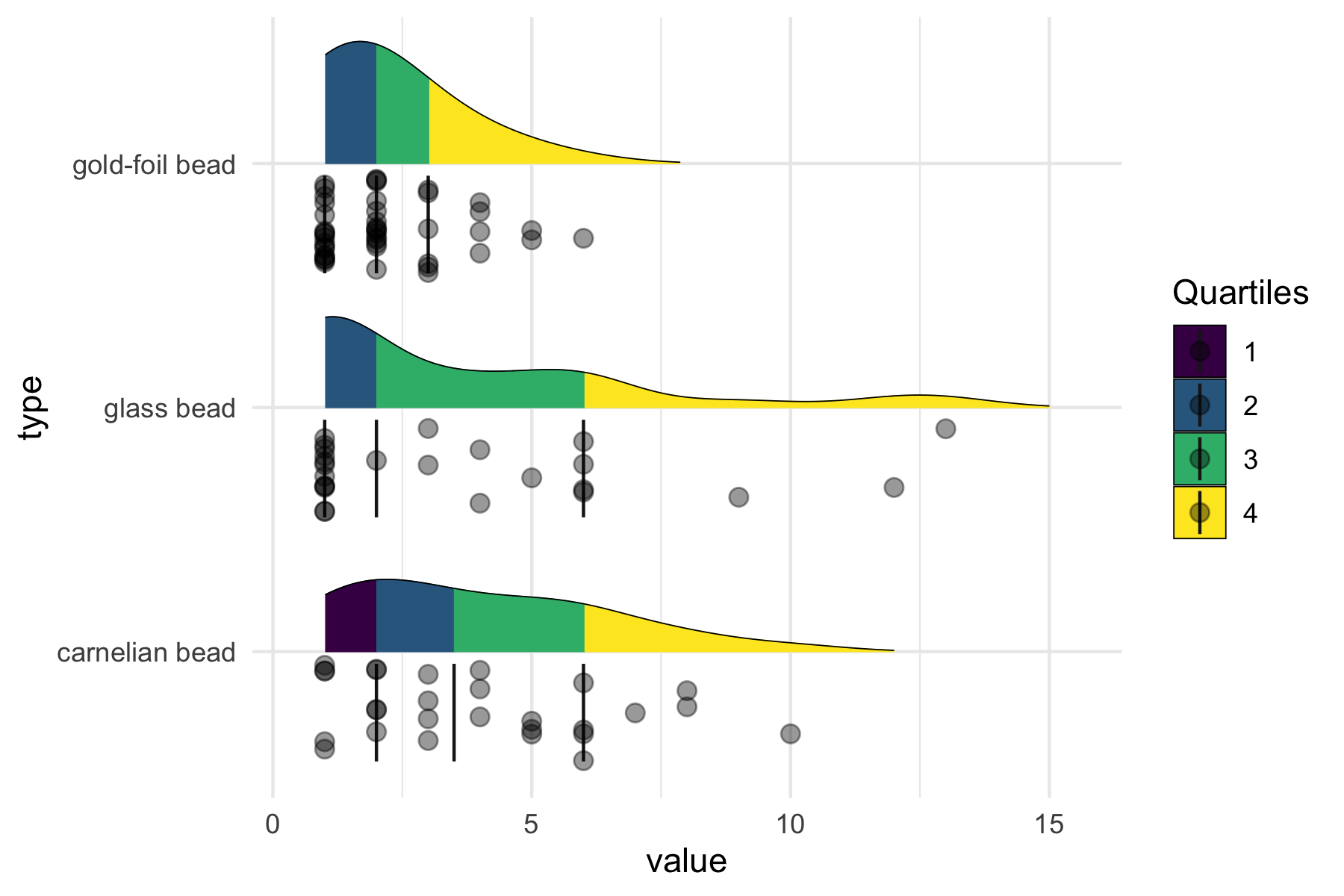


Figure 3.1: The distributions of different types of trade beads across burials analyzed in this paper.

After we set our model parameters, we simulated networks in a Bayesian framework using a Markov chain Monte Carlo (MCMC) algorithm. MCMC algorithms allow estimation of posterior distributions through direct random sampling the posterior without assuming the prior comes from any specific distribution (Hamra et al., 2013). We can obtain a posterior distribution by constructing a Markov chain that describes a sequence of moves from current state to the next state following probabilistic rules based on an algorithm. This enables a random or stochastic simulation in a long run where each move does not depend on the previous move. More chains ensures a more desirable posterior distribution that is close to the target distribution under study, or convergence. In Bayesian ERGMs, MCMC first selects a set of edges (or a set of empty pair of actors) with equal probability, and then switches to a pair of actors at random within the chosen set (Caimo and Friel, 2011). In our case, we set the number of chains to six. For each chain, the number of burn-in iterations was 100 and the number of iterations after the burn-in was 1000. We set the number of iterations used to simulate a network y’ at each iteration to 4000.

As an informal way to diagnose model convergence, we first evaluated three diagnostic visual summaries of our MCMC output, including density plots, trace plots, and autocorrelation plots for both models (Figure 3.2 and Figure 3.3) (Hamra et al., 2013). In general, models with informative priors have a better convergence of the MCMC. We can see that the diagnostic plots show stationary distributions and a trend of decreased autocorrelation. Then we summarize output from our two models by goodness-of-fit (GOF) diagnostics in the Bayesian framework, where the observed network is compared with the set of networks simulated from the estimated posterior distributions of the parameters of each model (Caimo and Friel, 2011; Caimo et al., 2017). We set network graphs simulated from the estimated posterior distribution in ERGMs. Our Bayesian GOF diagnostics summarized three distributions, including degree, minimum geodesic distance, and edgewise shared partner distributions. This provides a statistical approach to check how well the estimated posterior parameter distribution, based on our hypotheses, can reproduce networks with similar general structural features of the observed networks. We then compared the distribution of our observed networks, the networks before and after the arrival of Europeans, with the distribution of our hypothesized models. We expect to see the models fit with our hypotheses, indicative of a increased social inequality after the foreign contact. In addition, the covariates can give some more clues for the underlying mechanisms for the formations of each network, such as the relative importance of age and sex.

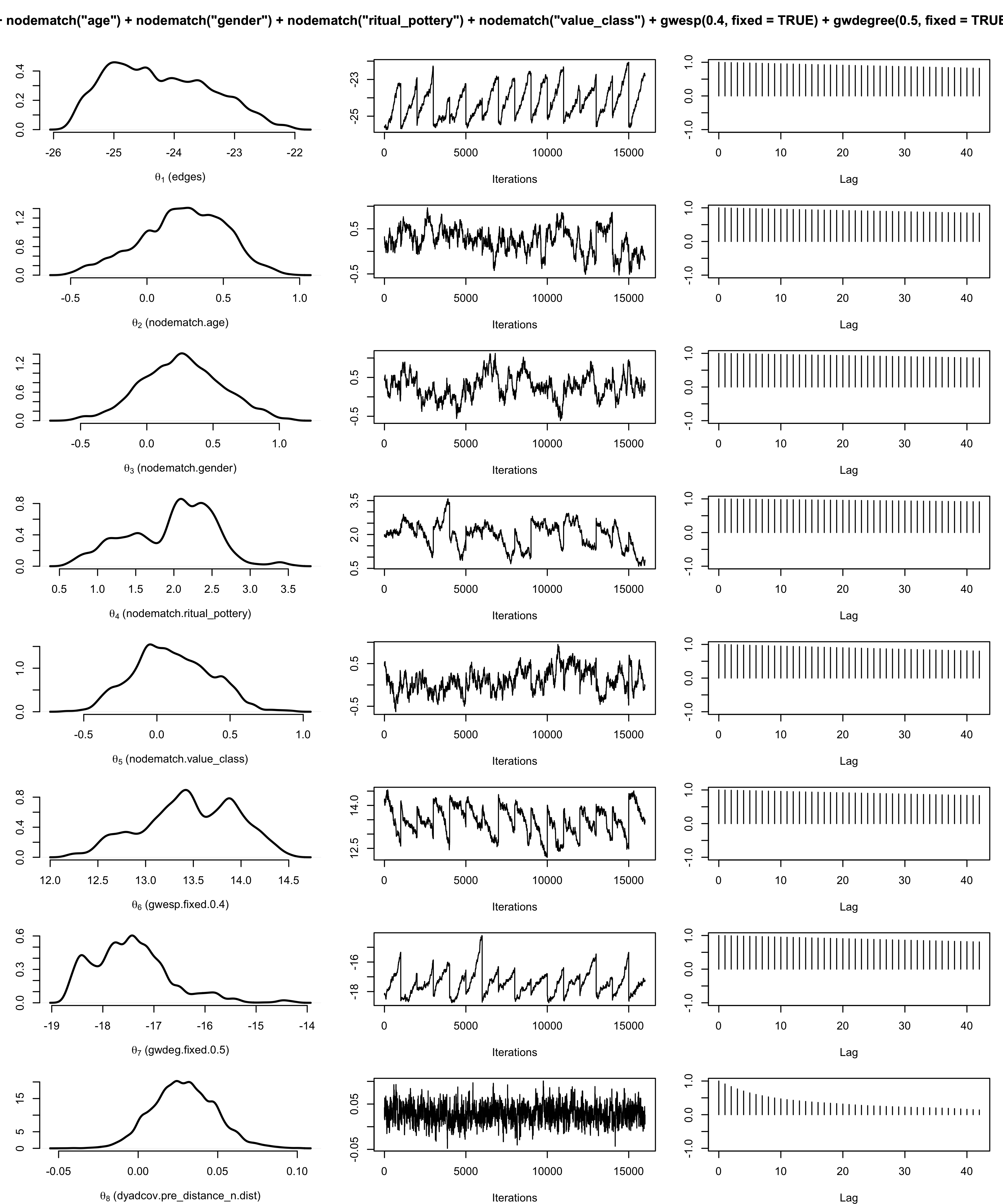


Figure 3.2: MCMC output of the pre-European model. In the order of edges, nodematch-age, nodematch-sex, nodematch-ritual, nodematch-value, gwesp, gwdeg, and dyadcov.distance.

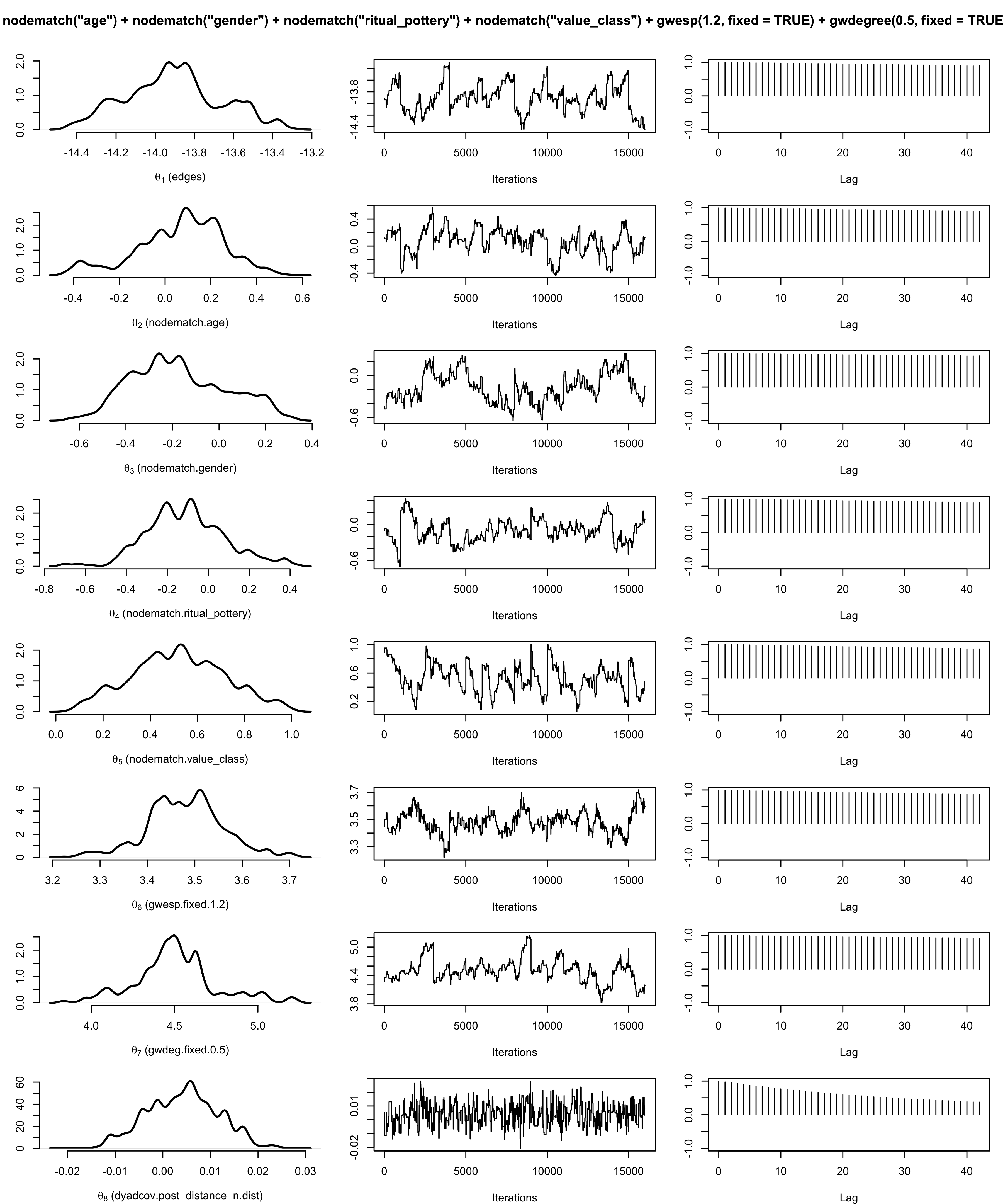


Figure 3.3: MCMC output of the post-European model. In the order of edges, nodematch-age, nodematch-sex, nodematch-ritual, nodematch-value, gwesp, gwdeg, and dyadcov.distance.

# 4 References

Andrade, T., 2007. How Taiwan became Chinese: Dutch, Spanish, and Han colonization in the seventeenth century. Columbia University Press, New York.

Blanton, R.E., Feinman, G.M., Kowalewski, S.A., Peregrine, P.N., 1996. A dual-processual theory for the evolution of mesoamerican civilization. Current anthropology 37, 1–14.

Borao Mateo, J.E., 2009. The spanish experience in Taiwan, 1626-1642 the baroque ending of a renaissance endeavor. Hong Kong University Press, Hong Kong.

Borgatti, S.P., Mehra, A., Brass, D.J., Labianca, G., 2009. Network analysis in the social sciences. science 323, 892–895.

Brughmans, T., 2013. Thinking through networks: A review of formal network methods in archaeology. Journal of Archaeological Method and Theory 20, 623–662.

Brughmans, T., Peeples, M.A., 2018. Network science. The Encyclopedia of Archaeological Sciences 1–4.

Caimo, A., Friel, N., 2011. Bayesian inference for exponential random graph models. Social Networks 33, 41–55.

Caimo, A., Pallotti, F., Lomi, A., 2017. Bayesian exponential random graph modelling of interhospital patient referral networks. Statistics in medicine 36, 2902–2920.

Chen, T.-j., 2005. Keelung shan yu Danshui yang: Dongya haiyu yu Taiwan zaoqishi yanjiu, 1400-1700 [Mount Keelung and Danshui Sea: A study of East Asian Seas and the early history of Taiwan from 1400 to 1700]. Lian jing, Taipei.

Collar, A., Coward, F., Brughmans, T., Mills, B.J., 2015. Networks in archaeology: Phenomena, abstraction, representation. Journal of Archaeological Method and Theory 22, 1–32.

Coward, F., 2013. Grounding the net: Social networks, material culture and geography in the epipalaeolithic and early neolithic of the near east (  21,000–6,000 cal bce), in: Knappett, C. (Ed.), Network Analysis in Archaeology: New Regional Approaches to Interaction. Oxford University Press, Oxford, pp. 247–280.

Crabtree, S.A., 2015. Inferring ancestral pueblo social networks from simulation in the central mesa verde. Journal of Archaeological Method and Theory 22, 144–181.

Feinman, G.M., 2000. Corporate/network: New perspectives on models of political action and the Puebloan Southwest, in: Schiffer, M.B. (Ed.), Social Theory in Archaeology. University of Utah Press, pp. 31–51.

Freeman, L., 2004. The development of social network analysis. A Study in the Sociology of Science.

Gamble, L.H., Zepeda, I.C., 2002. Social differentiation and exchange among the kumeyaay indians during the historic period in california. Historical Archaeology 71–91.

Gjesfjeld, E., 2015. Network analysis of archaeological data from hunter-gatherers: Methodological problems and potential solutions. Journal of Archaeological Method and Theory 22, 182–205.

Hamra, G., MacLehose, R., Richardson, D., 2013. Markov chain monte carlo: An introduction for epidemiologists. International journal of epidemiology 42, 627–634.

Janes, S., 2013. Death and burial in the age of the cypriot city-kingdoms: Social complexity based on the mortuary evidence. Bulletin of the American Schools of Oriental Research 145–168.

Kang, P., 2016. Zhimin xiangxiang yu difang liubian: Helan dongyindu gongsi yu Taiwan yuanzhumin [Colonial imagination and local variations: The Dutch East India Company and the Formosan Austronesians]. Lian-jing.

Klehm, C.E., 2017. Local dynamics and the emergence of social inequality in iron age botswana. Current Anthropology 58, 604–633.

Lehmann, B.C., Henson, R.N., Geerligs, L., White, S.R., others, 2020. Characterising group-level brain connectivity: A framework using bayesian exponential random graph models. bioRxiv 665398.

Marin, A., Wellman, B., 2011. Social network analysis: An introduction. The SAGE handbook of social network analysis 11.

Mills, B.J., 2017. Social network analysis in archaeology. Annual review of anthropology 46, 379–397.

Mizoguchi, K., 2013. Evolution of prestige good systems: An application of network analysis to the transformation of communication systems and their media, in: Knappett, C. (Ed.), Network Analysis in Archaeology: New Regional Approaches to Interaction. Oxford University Press, Oxford, pp. 151–180.

Morris, M., Handcock, M.S., Hunter, D.R., 2008. Specification of exponential-family random graph models: Terms and computational aspects. Journal of statistical software 24, 1548.

Peeples, M.A., 2019. Finding a place for networks in archaeology. Journal of Archaeological Research 27, 451–499.

Robins, G., Pattison, P., Kalish, Y., Lusher, D., 2007. An introduction to exponential random graph (p\*) models for social networks. Social networks 29, 173–191.

Salvini, A., 2010. Symbolic interactionism and social network analysis: An uncertain encounter. Symbolic Interaction 33, 364–388.

Snijders, T.A., 2011. Statistical models for social networks. Annual Review of Sociology 37.

Snijders, T.A., Pattison, P.E., Robins, G.L., Handcock, M.S., 2006. New specifications for exponential random graph models. Sociological methodology 36, 99–153.

Sosna, D., Galeta, P., Šmejda, L., Sladek, V., Bruzek, J., 2013. Burials and graphs: Relational approach to mortuary analysis. Social Science Computer Review 31, 56–70.

Wang, S.-C., Liu, Y.-C., 2007. Shiqi shiji qianhou Taiwan boli zhushi yu yancao, yandou de shuru wangluo -yige xin de jiaohuan jieduan [The import networks of tobacco, tobacco pipes, and glass bead ornaments into Taiwan circa the seventeenth century: A new phase of exchange]. Taida Journal of Art History 22, 51–83.

Wasserman, S., Faust, K., others, 1994. Social network analysis: Methods and applications. Cambridge university press.

##### 4.0.0.0.1 pagebreak

### 4.0.1 Colophon

This report was generated on 2021-01-21 23:00:27 using the following computational environment and dependencies:

#> ─ Session info ───────────────────────────────────────────────────────────────  
#> setting value   
#> version R version 4.0.2 (2020-06-22)  
#> os macOS Catalina 10.15.6   
#> system x86\_64, darwin17.0   
#> ui X11   
#> language (EN)   
#> collate en\_US.UTF-8   
#> ctype en\_US.UTF-8   
#> tz America/Los\_Angeles   
#> date 2021-01-21   
#>   
#> ─ Packages ───────────────────────────────────────────────────────────────────  
#> ! package \* version date lib source   
#> P assertthat 0.2.1 2019-03-21 [?] CRAN (R 4.0.2)  
#> P backports 1.2.1 2020-12-09 [?] CRAN (R 4.0.2)  
#> P bookdown 0.21 2020-10-13 [?] CRAN (R 4.0.2)  
#> P broom 0.7.3 2020-12-16 [?] CRAN (R 4.0.2)  
#> P callr 3.5.1 2020-10-13 [?] CRAN (R 4.0.2)  
#> P cellranger 1.1.0 2016-07-27 [?] CRAN (R 4.0.2)  
#> P cli 2.2.0 2020-11-20 [?] CRAN (R 4.0.2)  
#> P coda 0.19-4 2020-09-30 [?] CRAN (R 4.0.2)  
#> P colorspace 2.0-0 2020-11-11 [?] CRAN (R 4.0.2)  
#> P crayon 1.3.4 2017-09-16 [?] CRAN (R 4.0.2)  
#> P DBI 1.1.1 2021-01-15 [?] CRAN (R 4.0.2)  
#> P dbplyr 2.0.0 2020-11-03 [?] CRAN (R 4.0.2)  
#> P desc 1.2.0 2018-05-01 [?] CRAN (R 4.0.2)  
#> P devtools 2.3.2 2020-09-18 [?] CRAN (R 4.0.2)  
#> P digest 0.6.27 2020-10-24 [?] CRAN (R 4.0.2)  
#> P dplyr \* 1.0.3 2021-01-15 [?] CRAN (R 4.0.2)  
#> P ellipsis 0.3.1 2020-05-15 [?] CRAN (R 4.0.2)  
#> P evaluate 0.14 2019-05-28 [?] CRAN (R 4.0.1)  
#> P fansi 0.4.2 2021-01-15 [?] CRAN (R 4.0.2)  
#> P forcats \* 0.5.0 2020-03-01 [?] CRAN (R 4.0.2)  
#> P fs 1.5.0 2020-07-31 [?] CRAN (R 4.0.2)  
#> P generics 0.1.0 2020-10-31 [?] CRAN (R 4.0.2)  
#> P ggplot2 \* 3.3.3 2020-12-30 [?] CRAN (R 4.0.2)  
#> P ggridges \* 0.5.3 2021-01-08 [?] CRAN (R 4.0.2)  
#> P glue 1.4.2 2020-08-27 [?] CRAN (R 4.0.2)  
#> P gtable 0.3.0 2019-03-25 [?] CRAN (R 4.0.2)  
#> P haven 2.3.1 2020-06-01 [?] CRAN (R 4.0.2)  
#> P here 1.0.1 2020-12-13 [?] CRAN (R 4.0.2)  
#> P highr 0.8 2019-03-20 [?] CRAN (R 4.0.2)  
#> P hms 1.0.0 2021-01-13 [?] CRAN (R 4.0.2)  
#> P htmltools 0.5.1 2021-01-12 [?] CRAN (R 4.0.2)  
#> P httr 1.4.2 2020-07-20 [?] CRAN (R 4.0.2)  
#> P jsonlite 1.7.2 2020-12-09 [?] CRAN (R 4.0.2)  
#> P knitr 1.30 2020-09-22 [?] CRAN (R 4.0.2)  
#> P lattice 0.20-41 2020-04-02 [3] CRAN (R 4.0.2)  
#> P lifecycle 0.2.0 2020-03-06 [?] CRAN (R 4.0.2)  
#> P lubridate 1.7.9.2 2020-11-13 [?] CRAN (R 4.0.2)  
#> P magrittr 2.0.1 2020-11-17 [?] CRAN (R 4.0.2)  
#> P memoise 1.1.0 2017-04-21 [?] CRAN (R 4.0.2)  
#> P modelr 0.1.8 2020-05-19 [?] CRAN (R 4.0.2)  
#> P munsell 0.5.0 2018-06-12 [?] CRAN (R 4.0.2)  
#> P pillar 1.4.7 2020-11-20 [?] CRAN (R 4.0.2)  
#> P pkgbuild 1.2.0 2020-12-15 [?] CRAN (R 4.0.2)  
#> P pkgconfig 2.0.3 2019-09-22 [?] CRAN (R 4.0.2)  
#> P pkgload 1.1.0 2020-05-29 [?] CRAN (R 4.0.2)  
#> P plyr 1.8.6 2020-03-03 [?] CRAN (R 4.0.2)  
#> png 0.1-7 2013-12-03 [3] CRAN (R 4.0.0)  
#> P prettyunits 1.1.1 2020-01-24 [?] CRAN (R 4.0.2)  
#> P processx 3.4.5 2020-11-30 [?] CRAN (R 4.0.2)  
#> P ps 1.5.0 2020-12-05 [?] CRAN (R 4.0.2)  
#> P purrr \* 0.3.4 2020-04-17 [?] CRAN (R 4.0.2)  
#> P R6 2.5.0 2020-10-28 [?] CRAN (R 4.0.2)  
#> P Rcpp 1.0.6 2021-01-15 [?] CRAN (R 4.0.2)  
#> P readr \* 1.4.0 2020-10-05 [?] CRAN (R 4.0.2)  
#> P readxl 1.3.1 2019-03-13 [?] CRAN (R 4.0.2)  
#> P remotes 2.2.0 2020-07-21 [?] CRAN (R 4.0.2)  
#> P reprex 0.3.0 2019-05-16 [?] CRAN (R 4.0.2)  
#> P rlang 0.4.10 2020-12-30 [?] CRAN (R 4.0.2)  
#> P rmarkdown 2.6 2020-12-14 [?] CRAN (R 4.0.2)  
#> P rprojroot 2.0.2 2020-11-15 [?] CRAN (R 4.0.2)  
#> P rstudioapi 0.13 2020-11-12 [?] CRAN (R 4.0.2)  
#> P rvest 0.3.6 2020-07-25 [?] CRAN (R 4.0.2)  
#> P scales 1.1.1 2020-05-11 [?] CRAN (R 4.0.2)  
#> P sessioninfo 1.1.1 2018-11-05 [?] CRAN (R 4.0.2)  
#> P stringi 1.5.3 2020-09-09 [?] CRAN (R 4.0.2)  
#> P stringr \* 1.4.0 2019-02-10 [?] CRAN (R 4.0.2)  
#> P testthat 3.0.1 2020-12-17 [?] CRAN (R 4.0.2)  
#> P tibble \* 3.0.5 2021-01-15 [?] CRAN (R 4.0.2)  
#> P tidyr \* 1.1.2 2020-08-27 [?] CRAN (R 4.0.2)  
#> P tidyselect 1.1.0 2020-05-11 [?] CRAN (R 4.0.2)  
#> P tidyverse \* 1.3.0 2019-11-21 [?] CRAN (R 4.0.2)  
#> P usethis 2.0.0 2020-12-10 [?] CRAN (R 4.0.2)  
#> P vctrs 0.3.6 2020-12-17 [?] CRAN (R 4.0.2)  
#> P withr 2.4.0 2021-01-16 [?] CRAN (R 4.0.2)  
#> P xfun 0.20 2021-01-06 [?] CRAN (R 4.0.2)  
#> P xml2 1.3.2 2020-04-23 [?] CRAN (R 4.0.2)  
#> P yaml 2.2.1 2020-02-01 [?] CRAN (R 4.0.2)  
#>   
#> [1] /Users/EmilyWang/Desktop/School document/LW-Papers/kwl-burials-2020/renv/library/R-4.0/x86\_64-apple-darwin17.0  
#> [2] /private/var/folders/gd/flx6648926xgtkrpp55bl7200000gn/T/RtmpzTKNhM/renv-system-library  
#> [3] /Library/Frameworks/R.framework/Versions/4.0/Resources/library  
#>   
#> P ── Loaded and on-disk path mismatch.

The current Git commit details are:

#> Local: master /Users/EmilyWang/Desktop/School document/LW-Papers/kwl-burials-2020  
#> Remote: master @ origin (https://github.com/LiYingWang/kwlburials.git)  
#> Head: [cd10b3f] 2021-01-21: try code to save another set of plot

Word count: 1573