Research on the influencing factors of evolution of oil trade network based on TERGM model

Mengli Zhang¹, a, Yaoqi Guo², b
¹ Foundational Courses Department, Wuhan Donghu University, Wuhan 430212, PR China
² School of mathematics and statistics, Central South University, Changsha 410083, PR China

Abstract. This paper constructs a trade network using global oil trade data from 2001 to 2021, and uses TERGM model to explore the factors influencing the evolution of oil trade network. We found that the structures mutual, transitive closure, 3-outstars and 3-instars all boost the evolution of the oil trade network while cycle closure structure impede the evolution of the oil trade network. In addition, the level of national economic development inhibits the change of oil trade relations. Finally, geographical distance has little effect on the evolution of oil trade network, and the national border and common language have positive effect on the evolution of oil trade network.

Keywords: oil trade; complex network; Temporal Exponential Random Graph Models; influencing factors.

1. Introduction

With the development and progress of the social economy, the demand for oil consumption is increasing gradually. By the end of 2020, the world’s proven oil reserves are 244.4 billion tons, of which only the Middle East has 113.2 billion tons, accounting for 46.31%. The percentages of Central and South, North, CIS, Africa and Asia-Pacific are 20.80%, 14.78%, 12.11%, 6.79% and 2.4% respectively. The share of oil consumption in 2021 ranked Asia Pacific (38.65%), North America (22.6%), Europe (15%), the Middle East (8.8%), Central America (6.1%) and the Commonwealth of Independent States (4.6%). The uneven distribution of oil production and consumption in the world promotes the trans-regional flow of international oil trade, thus forming an oil trade network. In recent years, trade frictions and the instability of international relations make the international oil trade network more and more complex and interdependent, social factors such as economic development level and geographical distance cannot explain the evolution of oil trade pattern. Therefore, this paper studies the influence factors of oil trade evolution under the frame of network endogenous structure and external factors.

Complex network is an important tool to study network relationship. Gravity model is a common method used to study the influence factors of complex network. The gravity model was proposed by Salette & Tinbergen, and then many scholars used the gravity model to study the influence factors of trade network. In the gravity model, trade relations between two countries is taken as an explanatory variable, while GDP, geographical distance, similarity of official languages and trade policy are taken as explanatory variables. The gravity model does not take into account the correlation between the samples. In addition, the trade network is not only decided by bilateral trade relations, but also influenced by other trading countries. Therefore, it is reasonable to introduce the structure of network into the analytical framework. QAP is another model often used to study network relationships. Krackhardt proposed a new method QAP, which study the relationship between relational data and was proved the parameters was relatively unbiased. Although QAP can disrupt the correlation among samples to solve the problem of multiple collinearity, it does not include the endogenous structure of the network, there is still can not to explain how network endogenous structural dependence affects the formation of network relationships.

The exponential random graph model proposed by Wasserman & Pattison and Anderson & Wasserman is an advanced network statistical inference tool, which simulates real network by random networks. Zhang et al. exploring the status and structure of population movements among
China’s 328 large and medium-sized cities through Tencent big data, then using ERGM to analyze the factors that restrict and promote population movements\cite{16}. Furthermore, the temporal exponential graph model (TERGM) was proposed by Hanneke et al., which considers the effect of the last period network on the current period network based on the exponential random graph model. It can not only explore the spatial structure dependence of network, but also analyze the time dependence\cite{17}. He et al. construct a patent exchange network with 5 provinces, and use TERGM to explore the influence of the endogenous structural effects and exogenous factors\cite{18}.

The existing literatures provide the theoretical basis for the influence factors of trade from the perspective of binary relation, but they can not introduce the endogenous structure of network into the model to explain the formation of network relation. This paper analyzes the factors influencing the evolution of international oil trade network from the perspective of network dependence. The contribution of this paper as follow: (1) Describes the impact of dependence network structure on the formation of oil trade relations. (2) Taking the time factor into account, which reveals the influence factors of oil trade from the dynamic perspective.

2. Methodology and Data
2.1 Methodology

Complex network is a method to study the relationship. The global oil trade can be seen as a link between export country and import country which can establish a complex network. In this process we abstract each country as the node  $v$  of the network. If there is a trade relationship between two countries, a link is formed between the two nodes, with the arrow pointing from the export country to import country. In accordance with the above method, we constructed the oil trade network, and the global oil trade network for 2005, 2010, 2015, and 2020 four years are shown in Fig. 1.

Exponential random graph model is a statistical model used to study network structure and explain the formation mechanism of network relationship\cite{19}. The basic idea of the ERGM is the probability of the occurrence of a link in network when other link are known. It can consider both exogenous and endogenous variables. Based on ERGM model, TERGM model considers the time and the probability of forming the relation at the level of previous period. The network at time  $t$  is written as  $y_t$, and according Markov’s hypothesis, the network at period  $t$  is only related to the network pattern of the before  $k$  periods. The k-order Markov dependent TERGM model can be expressed by a formula as follows:

$$ P(Y_t = y_t|Y_{t-1}, Y_{t-2}, ..., Y_{t-K}, \theta) = \frac{1}{Z(\theta, y_{t-1}, y_{t-2}, ..., y_{t-K})} \exp \left( \sum_H \theta_H g_H(y_{t-1}, y_{t-2}, ..., y_{t-K}) \right) $$.  (1)

In above formula,  $P(\cdot)$  is the probability observed network formation,  $Z(\theta, y_{t-1}, y_{t-2}, ..., y_{t-K})$  is a normalized constant,  $H$  is a vector that the factor affecting the formation of network relationship,  $\theta_H$  is the coefficient vector of the influence factor.  $g_H$  represents the statistic that corresponds to  $H$  which including endogenous network structural effect, actor-relation effect, exogenous network effect and a lag variable that can reflect time-dependent features.

With the increase of nodes in network, the number of random networks simulated will increase exponentially. This makes it very difficult to take maximum likelihood parameters to estimate model parameters\cite{20}. Therefore, the maximum likelihood function is generally approximated using pseudo-maximum likelihood (MPLE)  \cite{21}  and Markoff chain Monte Carlo maximum likelihood (MCMC MLE)\cite{22}.
2.2 Data and preliminary

This paper uses the oil trade data from 2001 to 2021 to study the influence of factor of evolution of global oil trade network. Oil trade data is derived from the UN trade database(https://comtrade.un.org/data), HS code of which is 270900. The data of geographical distance between countries, border between countries and common language these three exogenous network variables are all from the CEPII database(http://www.cepii.fr/CEPII).

The Data pre-processing process: (1) For repeated trade records, we merged the record and the trade volume is averaged over the records. (2) In the study sample, only inter-country transactions were retained, and inter-state and regional trade records were deleted. (3) In order to highlight the important oil trade relations and weaken the influence of the unimportant relations on the network structure, this paper takes the trade relations whose annual oil trade volume is the first 70% into account.

3. Results analysis

In order to explore the factors that influence the evolution of oil trade network, we use temporal exponential random graph model to analyze the factors that influence the evolution of oil trade network structure.
### 3.1 Variable selection and description

In order to study the influencing factors of the structure evolution of oil trade network, this paper contribute TERGM model with endogenous structure effect, actor-relation effect and exogenous network effect step by step. The description of those variables are shown in Table 1. The circles in the configure represent nodes and arrows in the network represent edges. The black circles represent nodes with certain attributes. The dotted arrows indicate covariate networks which indicated the degree of coexistence between oil trade networks and covariate networks. In the statistics, \( y_{ij} \) denotes that if there is a relationship between node \( i \) and node \( j \) on the time \( t \), then \( y_{ij} = 1 \), others \( y_{ij} = 0 \).

#### Table 1 Description of primary variable types of TERGM

<table>
<thead>
<tr>
<th>Variables</th>
<th>parameter</th>
<th>configure</th>
<th>statistic</th>
<th>explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>edges</td>
<td></td>
<td></td>
<td>( 2\sum_{i&lt;j}y_{ij} )</td>
<td>This term is the basic network relation, which is equivalent to the intercept term in the regression model</td>
</tr>
<tr>
<td>mutual</td>
<td></td>
<td></td>
<td>( 2\sum_{i&lt;j}y_{ij}^r y_{ji}^r )</td>
<td>This term tests that whether there is significant reciprocity in the oil trade network</td>
</tr>
<tr>
<td>cycle closure</td>
<td></td>
<td></td>
<td>( \sum_{i,j,k}y_{ij}^r y_{jk}^r y_{ki}^r )</td>
<td>This term tests that whether there is significant cycle closure structure in the oil trade network</td>
</tr>
<tr>
<td>transitive closure</td>
<td></td>
<td></td>
<td>( \sum_{i,j,k}y_{ij}^r y_{jk}^r y_{ki}^r )</td>
<td>This term tests that whether there is significant transitive closure structure in the oil trade network</td>
</tr>
<tr>
<td>Out-3star</td>
<td></td>
<td></td>
<td>( \sum_i \sum_{j&lt;k&lt;l} y_{ij} y_{ik} y_{li} )</td>
<td>This term tests whether there is a significant divergence effect in the oil network</td>
</tr>
<tr>
<td>in-3star</td>
<td></td>
<td></td>
<td>( \sum_i \sum_{j&lt;k&lt;l} y_{ji} y_{ki} y_{li} )</td>
<td>This term tests whether there is a significant convergence effect in the network</td>
</tr>
<tr>
<td>Actor-relation effect</td>
<td>Main effect</td>
<td></td>
<td>( \sum_{i,j} y_{ij}^r \delta_i^r \delta_j^r )</td>
<td>This term tests whether certain attributes of the nodes contribute to the formation of network edges</td>
</tr>
<tr>
<td>Exogenous network</td>
<td>Covariable network</td>
<td></td>
<td>( \sum_{i,j} y_{ij}^r x_{ij} )</td>
<td>This term tests whether relationships in other networks</td>
</tr>
<tr>
<td>effect</td>
<td>Delayed Reciprocity</td>
<td>( \Sigma_{i,j} y_{ij} y_{j</td>
<td>i}^{-1} )</td>
<td>This term tests whether there is reciprocity between the last and current period data</td>
</tr>
<tr>
<td>Time effect</td>
<td>Stability</td>
<td>( \Sigma_{i.e} y_{ij} y_{j</td>
<td>i}^{-1} + \frac{1 - y_{ij}^{t}}{1 - y_{ij}^{t-1}} )</td>
<td>This term tests the stability of the network by counting the number of edges of the state not changed between the last and current period networks</td>
</tr>
</tbody>
</table>

The endogenous structural effect reflects the specific model of self-organization in the network, which does not involve the political, economic, cultural or other exogenous factors of the two countries, but only comes from the internal process of the network system. In the process of network formation, the nodes of the network can be regarded as actors with subjective will, and the attributes of the nodes determine the possibility of forming connections with other nodes. In the exponential random graph model, we refer to the actor-relationship effect as the factor that attributes of nodes affect the tendency of forming network relations. Exogenous network effect is the effect of other exogenous situational factors on relationship formation. Exogenous situational factors can be regarded as a network and are usually regarded as binary relationship covariates.

In this paper, Edge, reciprocity, delayed reciprocity, transmission cycle, closed cycle, convergence effect, divergence effect and stability are selected as endogenous structural variables. The Edge has no specific meaning and can be regarded as intercept term. Mutual means that there are both import and export trade relations between two countries. Triangular structure is the most basic structure of network, the analysis of triangular structure is helpful to understand the basic structural characteristics of network. The triangular structure can be divided into cycle closure and transitive closure. Transitive closure refers to the export/import from two of the three countries to the third country. From the structure of transitive closure, we can judge whether there is a hierarchical effect in the oil trade network. The divergence effect in oil trade is country export oil to many countries, while the convergence effect is a country import oil from many countries. In this paper, the parameter of convergence effect and divergence effect is set to 3, that is to say, if a country exports to three countries at the same time, the country is regard as has divergence effect, a country that imports from three countries at once is considered to have a convergent effect. If the coefficient of endogenous structure is greater than 0, the probability of the structure appearing in the trade network is greater than that in the random network, and vice versa. The time effect occurs when the current network is taken into account for the next period of the network. Delayed reciprocity means that the current one-way oil trade relationship will promote the next phase of the two countries to form a reciprocal oil trade relationship.

Node attributes can affect the formation of relationships in the network, that is, there are actor-relationship effects in the network. Studies have shown that the level of economic development or the gap between the level of economic development is an important factor affecting the relationship between trade networks. Oil is an important raw material for industrial production. Economic development will increase the consumption of oil and promote the formation of oil import relationship. Therefore, this article chooses the national economic development level (GDP) as the node attribute characteristics. When the attribute of actor-relationship variable is a continuous variable, the parameter estimation result of the variable is positive indicating that the attribute variable changes a unit, the probability of forming the edge of the network is greater than that of the random network.
In addition to the attributes of countries themselves, the characteristics of relations between countries also affect the formation of oil trade, that is, the external network effect. Geographical distance\textsuperscript{[24]}, national boundaries \textsuperscript{[25]}, common language\textsuperscript{[25]}, these have been shown to play a significant role in the impact of trade network factors. The geographical distance represents the geographical distance between the capitals of the two countries, the border between the two countries indicates the boundary line, and the common language indicates that the official languages of the two countries are the same. In this paper, geographical distance, national border and common language are added into the model as binary covariates. The positive and negative parameters of the external network effect explain that the corresponding variables can promote or inhibit the formation of trade network relations.

### 3.2 Results analysis of TERGM

MCMC-MLE and MPL are two methods of parameter estimation in TERGM model. MCMC-MLE estimates parameters by simulate real graph through random network, which results is more accurate than MPL. The MPL method replaces the estimated maximum likelihood function with the product of the conditional probability of each relation for the rest of the network. The advantage of this method is that the calculation is fast, but it is not possible to determine whether the sample size can satisfy the consistency. Therefore, the MCMC-MLE method is used to estimate the parameters of TERGM. We add the endogenous structure effect in Model 1, the actor-relationship effect in Model 2, and the endogenous structure effect in Model 3. The results of parameter estimation of TERGM are given in Table 2.

In Model 1, the endogenous effect of the network is included. In this model, all results in this model are significant at the level of 0.001. It shows that there are obvious characteristics of reciprocity, cycle closure, transitive closure, divergence and convergence in oil trade network. The coefficient of reciprocal structure is positive, which indicates that reciprocal trade relation promotes the formation of oil trade network. The coefficient of cycle closure structure is negative indicates that the number of cycle closure tends to decrease, and the cycle closure structure is less possible to form in the oil network. The coefficient of transitive closure is positive, which indicates that the number of transitive closure structures tends to increase in oil trade network and this structure is more possible to form. The coefficient of cycle closure is positive and transitive closure is negative, which indicates that there is a certain hierarchical effect in trading of oil.

The structural coefficient of circular closure is positive and the structural coefficient of transitive closure is negative, which indicates that there is a certain hierarchical effect in oil trade. That is to say, if there is a trade relation such as A→B→C, according to the positive effect of transitive structure and the negative effect of circularity, the probability of forming the trade relation of A→C is higher than that of forming C→A.

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>edges</td>
<td>-3.7974*** (0.0161)</td>
<td>-3.264*** (0.0205)</td>
<td>-2.0602*** (0.0363)</td>
</tr>
<tr>
<td>mutual</td>
<td>0.6929*** (0.0612)</td>
<td>0.4547*** (0.0581)</td>
<td>0.2421** (0.1067)</td>
</tr>
<tr>
<td>transitive closure</td>
<td>-0.2617*** (0.019)</td>
<td>-0.216*** (0.0191)</td>
<td>-0.144*** (0.0301)</td>
</tr>
<tr>
<td>cycle closure</td>
<td>0.1164*** (0.0032)</td>
<td>0.0994*** (0.0033)</td>
<td>0.0731*** (0.0041)</td>
</tr>
<tr>
<td>divergence</td>
<td>0.0053*** (0.0002)</td>
<td>0.0055*** (0.0002)</td>
<td>0.0036*** (0.0002)</td>
</tr>
<tr>
<td>convergence</td>
<td>0.0047*** (0.0001)</td>
<td>0.0052*** (0.0001)</td>
<td>0.0034*** (0.0001)</td>
</tr>
<tr>
<td>GDP</td>
<td>-0.0495*** (0.0124)</td>
<td>-0.0229* (0.0126)</td>
<td></td>
</tr>
</tbody>
</table>
The actor-relationship effect and the external network effect were added to Model 2, and the endogenous structure effect is consistent with the result of Model 1, and the change was small. All variables in the model were significant at the 0.1% level. The coefficient of economic development level is negative, which shows that the probability of forming new trade relations will be reduced by every unit of GDP increase. The coefficient of the distance between countries is negative, indicating that the distance between two countries has a negative effect on the oil trade relation build between two countries. The greater the distance between two countries, the less possible they are to form trade links. The coefficient is close to 0, which indicate geographical distance has little effect on the form of oil trade relation. A positive coefficient of the state border means that the possibility of trade links between two countries with a common border is increased.

In model 3, two variables delay reciprocity and stability with time effect are added to this model. The results can be seen in the estimation symbols in this model are constant with Model 1 and Model 2, and the level of significance remained almost same, with all variables being significant at the level of 0.001.

The coefficient of reciprocity is 0.2421, which shows that the formation probability of reciprocity edge is 1.2739($e^{0.2421}$) times that of random network. The coefficients of divergence effect and convergence effect are 0.0036 and 0.0034 respectively, which shows that the divergence structure of oil trade network is 1.0036($e^{0.0036}$) times of random network and convergence structure of oil trade network is 1.0034($e^{0.0034}$) times of random network. Every unit increase in a country’s economic development level, the probability of forming a new oil trade relation will be 0.9973($e^{-0.001}$) times that of the original one. In the framework of exponential random graph analysis, geographical distance has little influence on the formation of oil network. The coefficient of covariate network country border is 0.3783, and the probability of oil trade relationship between two countries with common boundary is 1.4598($e^{0.3783}$) times that of countries without common boundary. The coefficient of the covariate network common language is 0.1345, which means that the probability of a trade link between two countries using the same common language is 4.2207($e^{1.440}$) times higher than that between countries that do not share the same language. The use of a common language can reduce barriers to communication, not only facilitating trade, but even the recognition of cultural values that result from linguistic similarity, so countries with the same official language tend to build trade link. The significance level of delayed reciprocity was greater than 0.1, so it can not be concluded from the results that there is delayed reciprocity in the oil trade network. That is to say, one-way trade relations in the current period may not be able to determine the reciprocal structure in the next phase of the oil trade network. The stability coefficient is 2.9998 greater than 0. The effect of historical period relations on current period relations is studied by stability, the result is significantly positive and the coefficient is large, and the existing
trade relations have a tendency to keep stable with the change of time, it shows that the global oil trade network evolves steadily with time.

3.3 Goodness of fitting

In traditional regression models, AIC and BIC are usually used to measure the goodness of fitting. But these methods are suitable for the assumption of independence of dyad. When the covariate network is added to the model, the simulation based goodness of fitting evaluation method is needed\cite{26}. This method generates a series of random network graphs using the parameters estimated by the model at first, and then compares the network characteristics of the random network with those of the real network. Finally, based on the statistical index of network structure of simulation network and real network, the box-line graph of the statistical indexes of random network graph is drawn. If the real network falls in the middle of the box-line graph of the simulated network index value, then the fitting effect of the model is good\cite{27}. According to the parameter estimation results of TERGM model, 100 random networks are simulated every year, and GOF curves are drawn, in which the black solid line represents the statistical characteristics of the observation network and the box-line plot represents the statistical characteristics of the simulated network. In this paper, we compare the difference between the random network and the real network by using the following five statistical terms: dyad-wise shared partners, edge-wise shared partner, degree, triad census and geodesic distance.

![Graph showing the goodness of fitting](image)

**Fig. 2 Results of goodness of fitting**

The first five subgraphs in Fig. 2 show simulated versus actual values for the network’s five statistics terms. It can be found that the statistical features of real networks can all fall into the box-line diagram of the simulated network, which shows that the fitting effect of TERGM model is good. The final subgraph is the ROC/PR curve. In this figure the red curve is the ROC curve(receiver operating characteristic curve) and the blue curve is the PR curve(precision recall curves). These curves are derived from a confusion matrix whether there is an oil trade relationship between the two countries. The horizontal coordinate of the ROC curve is the false positive rate, indicating the proportion of relationships in the oil trade network that do not have trade links but that are predicted to have trade links. The vertical coordinate is the true positive rate, which indicates that there is a trade link in the forecast of the relationship and there is a trade link in the real network. The closer the ROC curve is to (0,1), the better the model is. From Fig. 1, the exponential random graph model in this paper has a good effect and can be used to explain the influence factors in the evolution of oil trade network.
4. Conclusion and recommendations

Oil is an important raw material for industrial production. Under the background of multiple risks and crises in the global economic development, it has become a profound issue to ensure the safe supply of oil. Based on the data of oil trade from 2001 to 2021, this paper discusses the influential factors of oil trade from the perspective of dualistic dependence. The results show that there are reciprocal effects, hierarchical effects, divergence effects, convergence effects and time-dependent effects in the oil trade network. These structures facilitate the evolution of oil trade network. The country border and common language have positive effect on the petroleum trade network. That is, if two countries have a common border or a common language, the possibility of oil trade links between the two countries will increase.

In order to guard against the oil supply security problem which may be brought by the complex economic environment and political relations, our country needs to take measures from many aspects. First, establish good diplomatic relations with the oil exporting countries. There is an obvious trade hierarchy in oil trade, and the trade relationship tends to export from the oil exporting power to other countries. To establish good relations with countries rich in oil resources is conducive to the sound development of oil trade relations between the two countries. Second, strengthen commodity trade cooperation with neighboring countries. Country border is an important factor influencing the formation of oil trade network. The existence of hierarchical effect in oil trade will lead to the expansion of the oil trade gap, which will lead to trade imbalance in the long run. Therefore, strengthening commodity trade with neighboring countries can maintain trade balance and contribute to China’s safe oil supply.

Acknowledgements

This work was supported by the Youth Natural Science Foundation of Wuhan Donghu University(Grant number: 2023dhzk018).

References


