

GEO-BASED VISUAL NETWORK ANALYSIS OF EXPORT AND IMPORT PATTERNS IN INTERNATIONAL WHEAT TRADE

Tamás T. SIKOS¹ , Aigul MEIRMANOVA²

DOI : 10.21163/GT_2020.152.09

ABSTRACT:

The aim of this paper is to describe worldwide export and import patterns in wheat trade and investigate their interconnections. The analysis uses Gephi techniques, including Geo layout and Fruchterman-Reingold algorithms. The research reveals the most important objectives of the wheat network by identifying and quantifying exporters and importers, subnetworks and flow pathways by calculating the centrality, network and clustering metrics. The results of the network analysis of the global wheat trade system highlight the scale-free character of the world's wheat supply system.

Key-words: *International wheat trade, GVNA, Gephi analysis, Geo layout and Fruchterman-Reingold algorithms.*

1. INTRODUCTION

What do a chapati, a matza, a bagel, an English muffin or a knäckebröd have in common? The answer is wheat, which is the main source of our staple food and it consists one-fifth of the calories and protein consumed globally. Climate change can affect agriculture in various ways. The crop yields as the wheat is susceptible to climate changes and most countries are less equipped to deal with the impact of heat stress, droughts, storms on their crops. Most exporter countries of wheat which are located in the Earth's southern hemisphere such as Australia, Argentina, India may face a reduction in wheat harvests of between 6% and 23% by 2050 (Wood, 2019). The projected deficiency of the world's wheat consume could be filled up by countries of the Earth's northern hemisphere by broadening their output and expanding crop harvests. As climate change warms the planet, unstable weather patterns and shifting seasons are disrupting the way of crops cultivation. The top wheat exporters such as Russia, the United States, Canada, France, Germany, Ukraine, Romania, Poland and other countries also may face disease, pests and weeds which may cause by frequent and rising temperatures.

The complexity of the international wheat trade necessitates the application of network analysis. The web of the international trade relations in wheat is considered as a network. The nodes of the network are the different countries, while the trade flows of wheat form the edges of the network in order to describe the presence of an import-export trade relations between geographic subnetworks or countries. This approach is widely used in studies of international trade (Chaney, 2014). There is a growing significance of network-based approaches to food system analysis and food safety-based analysis of supply systems in particular (Piniór et al., 2015; Hueston & McLeod, 2012; McDowell, 2017). Network-based approach has been recently applied in empirical studies of international trade (Li et al., 2003; Kastle et al., 2005; Serrano et al., 2007; Bhattacharya et al., 2007). In the present paper it is important to describe the relationships and interdependence of the geographic subnetworks, also the countries included in the given subnetworks. The different dimensions of connectivity can be designated applying complex network indices that measure the number and the density of the trading relationships, the level of clustering (bilateral or multilateral trade), their dispersion or concentration, the centrality for a given node (country).

¹University of Miskolc, Institute of Management Sciences, 3515 Miskolc, Hungary, sikos.tamas@uni-miskolc.hu; National University of Public Service, sikos.tamas@uni-nke.hu

²University of Miskolc, Institute of Management Sciences, 3515 Miskolc, Hungary, aygulmeyr@mail.ru

The trade-network studies has been profoundly deeply manipulated by the so-called "world system" theories, i.e. the notion that one can distinguish between core and peripheral countries. Thus, the structural analysis of the network is an indispensable condition for the completeness and universality of the network approach. The structure of interactions in the network is the object of research in areas such as theory graphs, the theory of random networks, the theory of the "tight world", the theory of scale-free networks.

Climate change including warming, drought and heat stress is highly likely to impact agricultural yields and, thereby, the proliferation of wheat availability. Due to the trade plays a key role in reallocating wheat production, addressing wheat security issues requires understanding the structure of trade networks. Graph theory is a network science modelling approach that studies the graphs made by the nodes and edges identified as relevant for describing a system. It defines metrics to capture the topology of empirical networks and tackle complex patterns of change. The topology in graph theory describes the arrangement of nodes and edges in networks, and gives important information on the pattern of nodes connectivity. Here, we study the world wheat trade network based on the BACI International Trade database and use the countries as nodes and the mass of international wheat flows as edges.

In the framework of this research we intend to test the following objectives through Gephi methods:

- No. 1: The international network of wheat can be characterized by a small number of central countries and a high number of peripheral countries. No. 1_a: The central countries are linked in a densely-connected core, while the majority of participant countries are relatively atomized around the periphery of the network. No. 1_b: The peripheral countries which are unable to cultivate and provide wheat to other countries due to its unsuitable weather and climate conditions, despite of this the peripheral countries export wheat to other countries;
- No. 2: A small number of countries which are located in one geographic subnetwork sell more than a big number of countries which are located in other geographic subnetworks with suitable conditions of weather for wheat cultivation. No. 2_a: Therefore, there are countries that both do not export and do not import wheat being outliers in the network graph. No. 2_b: The top importer of wheat is located in Asian or African subnetwork with a big population.

In the introduction section, we outlined two research objectives. Next, we present an overview of methodological issues related to the analysis of inter-country trade patterns and the nature of analyzing network data. Particular attention was paid to weighted network data and the corresponding centrality measures. This is followed by an overview of key results related to three types of weighted centrality measures, results from visual inspection of the network, and the optimal solution from weighted generalized modeling. The article concludes with a discussion of results and their implications; particularly with regard to export-import patterns of geographic subnetworks.

2. DATA USED AND METHODS

The heuristic value of network visualizations was first discovered in the second period of the 20th century in the early school of social network analysis or SNA (Wasserman & Faust, 1994). Jacob Moreno, founder of such approach, accurately stated that "the expression of an individual position can be better visualized through a sociogram than through a sociometric equation" (Moreno, 1953). Moreno and his disciples set the standards for the visualization of networks (Freeman, 2009). The analysis of social networks gained significant popularity in communication research, economics, geography, computer science, organizational and developmental research. GVNA or geo-based visual network analysis is the data visualization subsystem of the SNA and it is an important component of high-quality data mining systems, especially those focused on processing large volumes of information. Three main characteristics of data visualization tools are: the nature of the data; visualization methods and samples through which data can be represented; possibilities of interaction with visual images and methods for better data analysis. The visualization data tools can work with

the following types of data: one-dimensional data such as one-dimensional arrays, time series; two-dimensional data such as points of two-dimensional graphs, geographical coordinates; algorithms and programs such as information flows, debugging operations, etc.

The data used in the present article retrieved originally from the United Nations Statistical Division, but cleaned by BACI International Trade research and expertise teams using methodology of harmonization (BACI International Trade database, 2019). The dataset represents countries that exporting and importing wheat around the world in 2018-2019. The top exporters of wheat are Russia, the United States, Canada, Australia and France, whereas the top importers are Egypt, Indonesia, Algeria, Italy and Japan.

Gephi software was used as a tool to visualize and spatialize different network graphs due to its flexibility and multi-task architecture that brings new possibilities to work with complex datasets and produce valuable visual results (Jacomy et al., 2014). Gephi can represent graphs through various layouts such as ARF layout, Circular layout, Concentric layout, DAG layout, Dual Circle layout, Force Atlas layout, Fruchterman-Reingold layout, Geo layout, Hiveplot layout, Isometric layout, Layered layout, Maps of countries layout, Multipartite layout, Network Splitter 3D layout, OpenOrd layout, Radial Axis layout, Yifan Hu layout, Yifan Hu proportional layout (Cherven, 2015). Moreover, the abovementioned layout algorithms can be classified among themselves. The Geo layout, Maps of Countries layout, Isometric layout, Network Splitter 3D layout, Fruchterman-Reingold layout and other geo-based network layouts are categorized as Geographic layouts. The network data of the present article has a geographic component, such as longitude and latitude of each country, Geo layout and Fruchterman-Reingold layout algorithms were used to display geo-based network of the worldwide wheat trade flow.

The micro-analytical level is one of key constituent elements of *Geo-based Visual Network Analysis* which is used to evaluate node centrality (degree, closeness, betweenness), network measures (eccentricity, graph density, average path length), clustering measures (modularity). In **Table 1** the graph $G = (V, L)$ represents the group of dyadic connections among the vertices (nodes). The dimension of the graph is $V = (1, \dots, 94)$, where $N = 94$ is the overall amount of countries included in the analysis (De Benedictis & Tajoli, 2011). The link dimension of the graph is $L = (1, \dots, 1041)$, where $M = 1041$ is the overall number of existing links. The links (i.e. edges) are directed, routing from the exporter country, i , to the importer country, j , and $L_{ij} \in \{0, 1\}$. The binary information defines the existence, 1 , and non-existence, 0 , of a trade link. By adding W and P the graph will be transformed into the network: $N = (G(V, L); W; P)$, where W comprises all characteristics of the links, P comprises all characteristics of the nodes (e.g. attributes, geographical coordinates). The trade network is generally distinguished by solid heterogeneity. The 1041 existing edges in 2018-2019 conform to a density of the network of 0.119 (i.e the proportion of the number of existing connections to the total number of maximum connections possible), showing 12% of nodes are connected with one another.

Table 1.

Overview of graph statistics of the graph

No. Countries	94
No. Links	1041
Graph density	0.119
Centrality measures:	
In-degree centrality	36
Out-degree centrality	58
Closeness centralization	1
Betweenness centralization	0.12679
Network measures:	
Eccentricity	6
Average path length	2.397
Clustering measures:	
Modularity	0.222

This is neither an exceptionally high number, nor could it be considered very low.

Centrality statistics provide the framework to compare the roles played by various nodes within a single network and the centrality measures help to define aggregate patterns within the graph. Centrality measures are categorized into three main groups: a) degree centrality, that measures how a node is connected to other nodes; b) closeness centrality, that measures how easily a node can be reached by other nodes; c) betweenness centrality, describing how important a node is in terms of connecting other nodes (Jackson, 2010). Centrality statistics is an appropriate tool in analyzing wheat trade data.

In a given network there are two measures of degree centrality: in-degree centrality, measuring the number of edges pointing to ego, and out-degree centrality, measuring the number of edges going from ego. The out-degree is the total number of countries toward which country i is exporting;

$$C_{Dout} = \frac{\sum_{j \neq i}^N L_{ij}}{(N-1)} \quad (1)$$

whereas in-degree is the total number of countries from which country i is importing:

$$C_{Din} = \frac{\sum_{j \neq i}^N L_{ji}}{(N-1)} \quad (2)$$

The given dataset in the present article has an *in-degree centrality* of 36% and an *out-degree centrality* of 58%. The importing countries are receiving edges 36% of the countries in the network, while is sending edges to 58% of them. One of the most commonly used global centrality measures is the closeness centrality. It is a measure of how close (in terms of topological distance) a node is with respect to all other nodes. The shortest path between country i and country j is called the geodesic distance between i and j :

$$C_c = \frac{(N-1)}{\sum_{j \neq i}^N D_{ij}} \quad (3)$$

The *closeness centrality* is equal to $C_c = 1$, which indicates there are no far-flung outliers lying at a great distance from the hub of the graph. All countries are on average 1 topological step away from all other countries in the network. The betweenness centralization is the variation in the betweenness centrality of vertices divided by the maximum variation in betweenness centrality scores possible in a network of the same size.

The *betweenness centrality* is equal to $C_b = 0.12679$, which indicates showing the strength of the economic links between countries in the network. The *eccentricity* is tightly related to diameter for each vertex within the network. Every vertex in the network have a value equal to 6, indicating how many nodes match the criteria. *Average path length* can provide insight into the general structure and connectedness of a network. Average path length measure is 2.397 which is below the eccentricity indicator showing that the network is efficient, and information can easily flow across the graph.

Modularity was used for graphs as one of the measures of the network. The modularity is used for network partition into clusters and for disclosure of community structure (Dedić & Stanier, 2017). The *Modularity* statistic places individual nodes into an aggregated group or cluster based on shared characteristics. Output for this function is simply an integer value starting at 0.222.

3. RESULTS AND DISCUSSIONS

Fruchterman-Reingold algorithm had been applied in order to generate a more precise map of the given dataset, pushing the most connected nodes away from each other whereas flattening the nodes that are connected to the hubs in clusters around them. The geographical coordinates of research locations with latitudinal, longitudinal and attribute values were imported into the Gephi, the overall quantity of the nodes were 94 and appropriate edges were 1041 in the network graph. The various communities in a network have distinct characteristics such as node degree, centrality, betweenness etc. A numerous real-world problem may be indicated as the structured feature by applying graphs.

Each node corresponds to a country and each edge performs an export-import relationship between countries. The colored nodes represent wheat exporters around the world (**Fig. 1**). The present geo-based network displays a clear pattern of a core of interconnected countries where shown as categorized clusters of countries which supply wheat around the world.

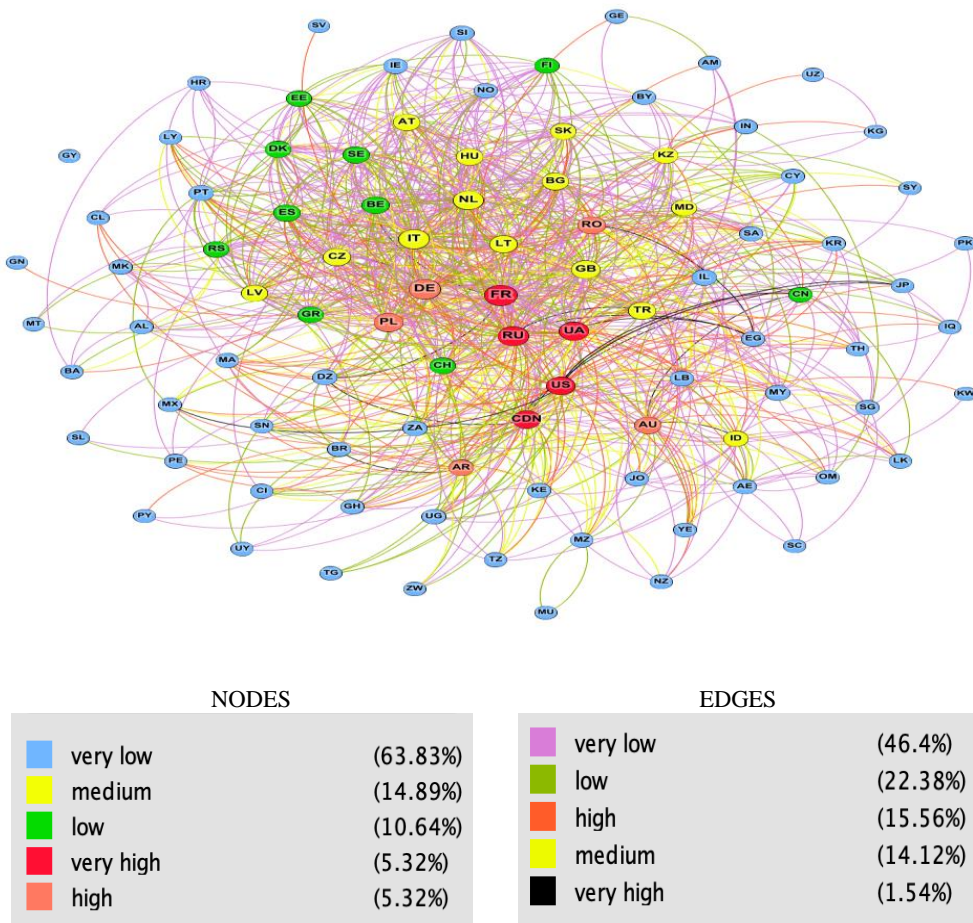


Fig. 1. The worldwide wheat exporters by Fruchterman-Reingold algorithm, 2018-2019.

The wheat exporter-countries are classified by attributes which are identified by to how many countries they export wheat to. The categorization is identified by red colored nodes – to very high amount of countries export to (5.32%), orange colored nodes – to high amount of countries export to (5.32%), yellow colored nodes – to medium amount of countries export to (14.89%), green colored

nodes – to low amount of countries export to (10.64%), blue colored nodes – to very low amount of countries export to (63.83%). Each node performs wheat exporter country which leads to identify ‘hub’ supplier countries and to categorize exporters in order to highlight the largest players in the market (**Fig. 1**). The majority of countries are rather peripheral, but some countries play a central role in wheat trade flows. The identification of local centres and the most important trade flows is critical to determine the most vulnerable parts of the network (Ślusarczyk, 2017).

Clustering methods will do this through specialized algorithms that interpret network patterns into distinct groupings (clusters) of similar nodes (Cherven, 2015). Clustering of the network graph by nodes according to the sizing and coloring options, therefore beforementioned options were created according to the categories to which they belong to. The same approach can be applied using node size to scale values according to their degree level (David & Tukey, 1977). The size of vertices is proportional to the number of countries where exporters send wheat to: a large circle indicate that the corresponding country is among the main partner for a large number of importing countries. A typical approach is to have scaled coloring based on a measurable attribute, with orange and red colored nodes corresponds to a high and very high number of countries they export wheat to. The network graph represents how the wheat trade network is polarized around the United States of America, Canada, Russia, Ukraine, France, Australia, Argentina, Poland, Germany, Romania, acting as the trajectory of the wheat trade system.

A typical approach is to have scaled coloring based on a measurable attribute, with yellow and green colored nodes corresponds to a medium and low number of countries they export wheat to. The centrality of such countries as India, the United Kingdom, Latvia, Lithuania, Kazakhstan, Turkey, Moldova, Bulgaria, Italy, the Netherlands, Hungary, Czech Republic, Slovakia, Austria, China, Switzerland, Belgium, Serbia, Greece, Spain, Sweden, Estonia, Finland and Denmark are evident and these countries have a supportive position in the wheat trade system. Other countries represented in blue colored nodes are peripheral countries in the network graph, indicating that the international network of wheat can be characterized by a small number of central countries and a high number of peripheral countries. This means that the central countries are interconnected in a densely-linked core, while the majority of peripheral countries are relatively atomized around the circumference of the network. As a summary it can be stated that Fruchterman-Reingold layout algorithm determines the hubs and peripheral states of the international wheat trade network, which can be considered by a scale-free network. In Barabási’s and Bonabeau’s (2003) view, a scale-free network can be characterized by a relatively low number of nodes, which have a high number of connections to other nodes. The results have supported Objectives No. 1, 1_a due to the international trade in wheat can be designated by a few hubs, around which there are numerous peripheral countries.

A typical approach is to have scaled coloring based on a measurable attribute with blue colored nodes corresponds to a very low number of countries they export wheat to. Some peripheral countries as el Salvador, the Seychelles, Sri Lanka, Kuwait, Singapore, Oman, Thailand, Jordan, Yemen, Japan, Iraq, Syria, Senegal, Israel, South Korea, Cyprus, Bosnia and Herzegovina, Macedonia, Croatia, Malta, Libya, Guinea, Kyrgyzstan, Uzbekistan, Guyana, Zimbabwe, Uruguay, Sierra Leone, Mauritius, Mozambique, Tanzania, Ghana, Kenya, Togo etc. are unable to provide wheat to other countries due to its weather and climate conditions that is unsuitable for cultivation wheat. Despite of this the abovementioned countries export wheat to other countries, supporting Objective No. 1_b.

Geo layout algorithm was applied in order to spatialize a full geo-based network of wheat trade flow. Nodes repulse each other like charged particles, while edges attract their nodes, such as springs. These forces create a movement that converges to a balanced state (Jacomy et al., 2014). The network represented the separate countries as the nodes in a directed spatial network, based on the geographic coordinates (latitude/longitude) for each country. Edges between node pairs were weighted by the value of wheat exports (US\$) transferred between each origin and destination country in 2018-2019 (**Fig.2**). The categorization is identified by black colored edges – exporter send wheat to importer in very high value between \$330M - \$3B (consists of 1.54%), orange colored edges – exporter send wheat to importer in high value between \$30M - \$330M (consists of 15.56%), yellow colored edges – exporter send wheat to importer in medium value between \$9M - \$30M (consists of 14.12%), green

colored edges – exporter send wheat to importer in low value between \$1M - \$9M (consists of 22.38%), purple colored edges – exporter send wheat to importer in very low value between \$1K - \$1M (consists of 46.4%).

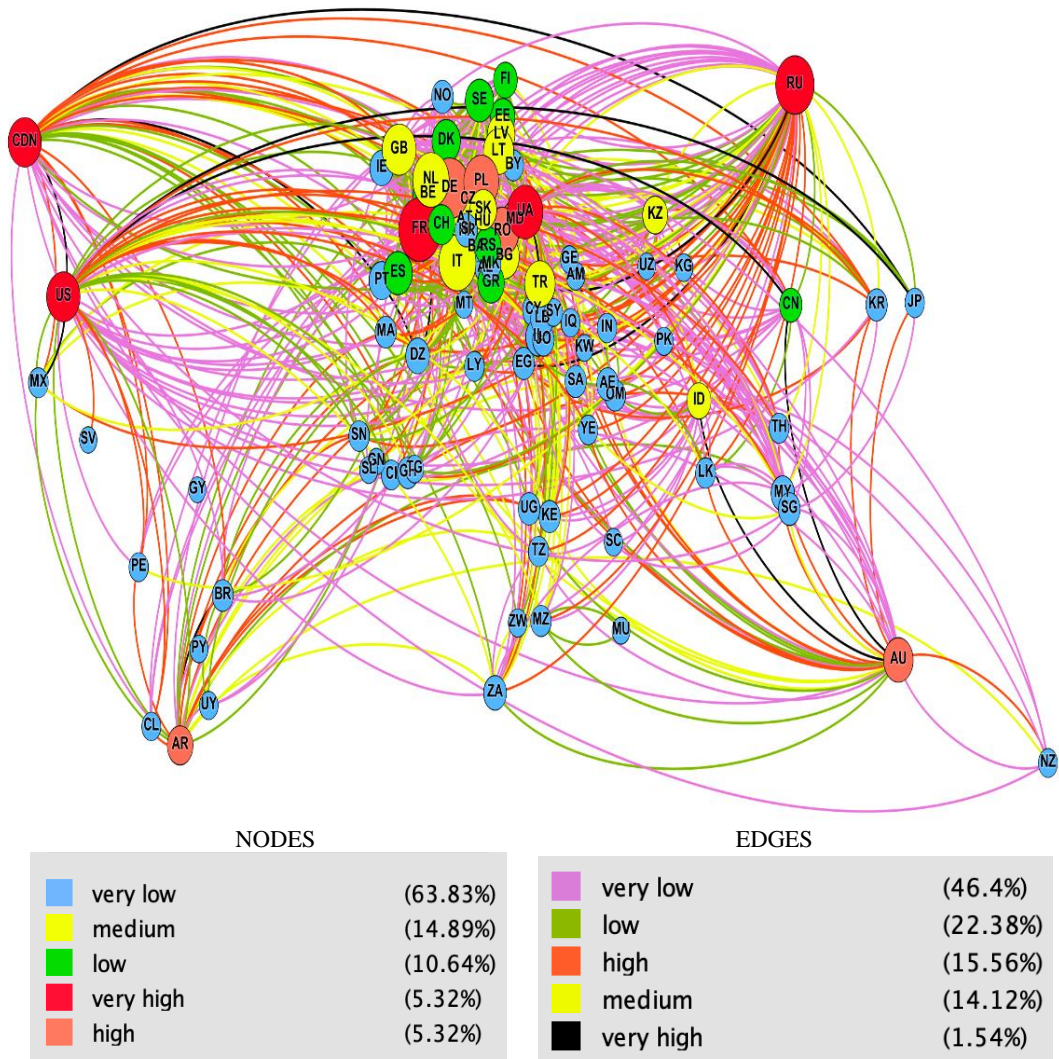


Fig. 2. Graphical representation of the world trade flow of wheat ((1) the number of nodes is proportional to the number of outflows, (2) the number of edges is proportional to the amount of the export value).

Classifications by geographical sub-networks is not only identify the main trends of the worldwide trade flow, but also analyze, compare continents with each other. But any classification is useful only when it is possible to identify clear criteria, classification features that enable us unambiguously evaluate certain phenomena. International trade export in wheat considered as geographically segmented, with the major role played by European sub-network (export value is \$22.7B), while the North American sub-network with its main players such as Canada, the United States are located in a decentralized position (export value is \$11.4B) (**Fig.2**). The third hub is Oceanic sub-network (export value is \$4.88B) supporting Objective No. 2 which claims that one geographic sub-network export more than other sub-network with more countries with suitable conditions of weather for cultivating wheat in comparison with the South American sub-network (export value is

\$2.71B), the Asian sub-network (export value is \$900M), the African sub-network (export value is \$33.2M) from **Fig.2**.

There are inward and outward flows between countries, based on the network shown in **Fig.2** we notice that:

- Some edges between countries do not exist, due to some countries such as Mongolia, Afghanistan, Colombia, Venezuela, Ecuador, Mauritania, Mali, Chad, Sudan, Angola, Mozambique etc. do not export wheat to other countries and Turkmenistan, Chad, Namibia, Botswana, Laos do not import wheat. Therefore, there are countries Turkmenistan, Chad, Namibia, Botswana, Laos both do not export and do not import wheat being outliers in the network graph supporting Objective No. 2_a which claims that there are number of countries do not export and do not import;
- The top importers of wheat around the world in the African sub-network are Egypt (import share 9.8%) and Algeria (import share 4%), in the Asian sub-network are Indonesia (import share 6.1%) and Japan (import share 3.3%), in the South American sub-network is Brazil (import share 2.5%), in the North American sub-network is Mexico (import share 2.5%); in the European sub-network is Italy (import share 3.6%), supporting Objective No. 2_b which claims that the top importer located in Africa or in Asia with big population.

4. CONCLUSION

We examined a number of issues related to the international wheat trade system applying the tools of Network Analysis. The world wheat trade network is represented graphically and analytically through the indices describing the network's properties, such as in-degree centrality, out-degree centrality, closeness centralization, betweenness centralization, eccentricity, average path length and modularity. The analysis shows a role of countries in shaping the wheat trade network. This approach has relevant implications in demonstrating some valuable insights about the group of wheat exporters and importers around the world: what are the subgroups it consists of, who are the most influential players, what are the peripheral players and outliers, etc. The results of a network analysis of the global wheat trade system highlight the scale-free character of the world's wheat supply system. This information is important for the organizations of international trade control systems during temperature changes and when some countries are less equipped to deal with the impact of heat stress on their crops.

REFERENCES

- Baci International Trade Database (2019). – Available from <https://oec.world/en/profile/hs92/1001/>. [Accessed December 2019].
- Barabási, A. & Bonabeau, E. (2003) Scale-free networks. *Scientific American*, 288 (5), 50-59.
- Bhattacharya, K., Mukherjee, G., Saramaki, J., Kaski, K. & Manna, S. (2007) The International Trade Network: Weighted Network Analysis and Modeling. *Journal of Statistical Mechanics: Theory and Experiment*, 2008 (02), 2002.
- Chaney, T. (2014) The network structure of international trade. *American Economic Review*, 104 (11), 3600-3634.
- Cherven, K. (2015) *Mastering Gephi Network Visualization*, Packt Publishing Ltd., Birmingham.
- David, F.N & Tukey, J. W. (1977) Exploratory Data Analysis. *Biometrics*, 33 (4), 768.
- Dedić, N. & Stanier, C. (2017) Towards Differentiating Business Intelligence, Big Data, Data Analytics and Knowledge Discovery. *Lecture Notes in Business Information Processing*, 114–122.
- De Benedictis L. & Tajoli L. (2011) The World Trade Network. *The World Economy*, 34 (8), 1417-1454.
- Freeman, Linton C. (2009) Methods of Social Network Visualization. *Encyclopedia of Complexity and Systems Science*, 1-17.
- Hueston, W. & McLeod, A. (2012) *Overview of the global food system: Changes over time/space and lessons for future food safety. Improving Food Safety through a One Health Approach*, National Academies Press, Washington.

- Jacomy, M., Venturini, T., Heymann, S. & Bastian, M. (2014) Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. *PLoS ONE*, 9 (6), e98679.
- Jackson, M. (2010) *Social and Economic Networks*, Princeton University Press, New Jersey.
- Kastelle, T., Steen, J. & Liesch, P. (2005) Measuring globalisation: an evolutionary economic approach to tracking the evolution of international trade. Paper presented at the *DRUID Summer Conference on Knowledge, Innovation and Competitiveness: Dynamics of Firms, Networks, Regions and Institutions*.
- Li, X., Jin, Y. Y., & Chen, G. (2003) Complexity and synchronization of the World trade Web. *Physica A: Statistical Mechanics and its Applications*, 328, 287-296.
- Moreno, J. (1953) *Who Shall Survive? Foundations of sociometry, group psychotherapy and psychodrama*, Beacon House Inc., New York.
- McDowell, T. C. (2017) The complexity of the international food industry. *International Journal of Society Systems Science*, 9 (1), 1-28.
- Piniór, B., Conraths, F.J., Petersen, B. & Selhorst, T. (2015) Decision support for risks managers in the case of deliberate food contamination: the dairy industry as an example. *Omega*, 53(1), 41-48.
- Serrano, A., Boguna, M., & Vespignani, A. (2007) Patterns of dominant flows in the world trade web. *Journal of Economic Interaction and Coordination*, 2, 111-124.
- Ślusarczyk, B. (2017) Prospects for the shared services centers development in Poland in the context of human resources availability. *Polish Journal of Management Studies*, 15(1), 218-231.
- Wasserman, S. & Faust, K. (1994) *Social Network Analysis: Methods and Applications: Methods and Applications*, University Press, Cambridge.
- Wood, J. (2019) *These 5 foods are under threat from climate change*. [Online] Available from <https://www.weforum.org/agenda/2019/08/these-5-foods-are-under-threat-from-climate-change/>. [Accessed January 2020].