

Thèse de doctorat de l'Université Paris 1 Panthéon-Sorbonne  
Ecole Doctorale de Géographie de Paris  
UMR CNRS 8504 Géographie-cités

# Évolution et robustesse du réseau maritime mondial : une approche par les systèmes complexes

Evolution and Robustness of the Maritime Trade Network:  
A Complex Systems Approach

*Présentée par*

Zuzanna Kosowska-Stamirowska

Sous la direction d'Arnaud Banos

Date de soutenance : 17 décembre 2019

*Composition du jury :*

<b>Lena Sanders</b>	Directrice de Recherche, CNRS (Présidente du jury)
<b>Jan Hoffmann</b>	Docteur, Chief of Trade Logistics Branch, UNCTAD (Examineur)
<b>Ray Rivers</b>	Professeur, Imperial College London (Rapporteur)
<b>Jean-Paul Rodrigue</b>	Professeur, Hofstra University (Rapporteur)
<b>Arnaud Banos</b>	Directeur de Recherche, CNRS (Directeur)

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement No. [313847] "World Seastems".

# Resumé

Plus de 70% de la valeur totale du commerce international est acheminée par voie maritime, ce qui représente 80% de la totalité du fret en volume. En 2016, le secrétaire général de l'ONU a attiré l'attention sur le rôle du transport maritime, qu'il décrit comme «l'épine dorsale du commerce mondial et de l'économie mondiale». Les flux du commerce maritime ont un impact non seulement sur le développement économique des régions concernées, mais également sur leurs écosystèmes. Les navires en mouvement étant notamment un important vecteur de propagation pour les bioinvasions. En parallèle, l'avenir du secteur du transport maritime est inextricablement lié au changement climatique : les mouvements de navires contribuent de manière significative aux émissions mondiales de CO<sub>2</sub>, de NO<sub>x</sub> et de SO<sub>x</sub>, avec des émissions de CO<sub>2</sub> comparables à celles de l'Allemagne et des émissions de NO<sub>x</sub> et de SO<sub>x</sub> comparables à celles des États-Unis. Le développement de la navigation dans l'Arctique devenant une réalité, la nécessité de comprendre le comportement de ce système et de prévoir les futurs flux du commerce maritime s'impose.

Malgré l'importance évidente et cruciale de la logistique maritime pour l'économie mondiale, très peu de travaux fournissent une vue détaillée de la répartition mondiale des flux commerciaux maritimes, et encore moins n'analysent leur évolution sur le long terme et les règles qui les régissent. Dans cette thèse, nous utilisons une base de données unique sur les mouvements quotidiens de la flotte mondiale entre 1977 et 2008, fournie par l'assureur maritime Lloyd's, afin de constituer un réseau complexe des flux du commerce maritime où les ports représentent les nœuds et les liens sont créés par les traversées de navires.

Dans cette thèse, nous effectuons une analyse du réseau du commerce maritime qui est entièrement *data-driven*. Nous utilisons des outils issus de la science de la complexité et de *Machine Learning* appliqués aux données de

réseau pour étudier les propriétés de ce réseau et développer des modèles de prévision des ouvertures de nouvelles lignes maritimes et des volumes des flux commerciaux futurs sur des liens.

L'application du *Machine Learning* pour analyser les flux commerciaux sur le réseau nous paraît être une nouvelle approche par rapport à l'état de l'art. Cette approche nécessitait une sélection et une modification soigneuses des outils d'apprentissage automatique existants pour les adapter aux données de type réseau et sur des flux physiques. Les résultats de la thèse suggèrent que le commerce maritime suit une marche aléatoire sur la structure sous-jacente du réseau.

Après le nettoyage et le prétraitement nécessaires des données brutes, nous analysons les propriétés structurelles des couches composées de 5 types de navires : porte-conteneurs, vraquiers secs, navires de charge, pétroliers et méthanières - représentant plus de 90% du DWT total dans la base de données. Les propriétés de base du réseau du commerce maritime et leurs évolutions au cours de la période étudiée sont analysées.

Le réseau maritime exhibe une croissance super-linéaire du nombre de liens par rapport au nombre de noeuds, alors même que le diamètre du réseau reste stable, ce qui laisse soupçonner une densification du réseau. Nous découvrons également que la distribution des degrés des différentes couches et du réseau agrégé est proche d'une distribution géométrique, mais le *fit* n'est pas idéal. En général, nous concluons que le réseau vise à améliorer sa propriété d'expansion tout en ayant des propriétés de navigabilité quasi optimales, étant donné la distribution de ses degrés.

Nous examinons ensuite une question primordiale du point de vue des autorités portuaires : «Quelles sont les mesures de centralité les plus pertinentes pour le succès d'un port, si l'on prend comme critère de succès l'augmentation du tonnage traité dans le port?». Nous constatons que la mesure dont la corrélation est la plus forte avec une hausse du DWT des ports est la *current flow closeness*. Cette mesure de centralité reflète la position centrale d'un nœud dans le réseau pour les échanges avec d'autres nœuds suivant des modèles de *random walks*.

Nous passons ensuite à la découverte des règles qui régissent l'évolution du réseau du commerce maritime en découvrant les modèles de croissance

du réseau directement à partir des données du réseau. Nous trouvons que l'évolution de ce système est régie par une règle simple qui provient de la science des réseaux et qui repose sur le nombre de voisins en communs (*Common Neighbors*) entre des paires de ports. Ce constat est cohérent sur les trois décennies de données. En outre, nous trouvons qu'un modèle de commerce basé sur les flux – dans lequel les ports sont censés échanger avec leurs voisins de réseau proportionnellement au nombre de leurs voisins communs – fournit les prévisions de volume de flux commerciaux les plus précises d'une année à l'autre.

Nous confirmons ces résultats en étudiant la réaction du réseau aux chocs. Grâce à une expérience naturelle impliquant une redirection du trafic du port de Kobe après le tremblement de terre de 1995, nous constatons que le trafic a été redirigé de préférence vers les ports qui avaient le plus grand nombre de voisins communs avec Kobe avant la catastrophe naturelle.

Ensuite, en simulant des attaques ciblées sur le réseau du commerce maritime, nous analysons les meilleurs critères qui permettraient de maximiser les dommages causés au réseau, ainsi que la robustesse générale du réseau face aux différents types d'attaques.

Tous ces résultats suggèrent que les flux commerciaux maritimes suivent une forme de marche aléatoire sur le réseau des connexions maritimes, ce qui fournit la preuve d'une vision nouvelle de la nature des flux commerciaux.

# Abstract

Over 70% of the total value of international trade is carried by sea, accounting for 80% of all cargo in terms of volume. In 2016, the UN Secretary General drew attention to the role of maritime transport, describing it as “the backbone of global trade and of the global economy”. Maritime trade flows impact not only the economic development of the concerned regions, but also their ecosystems. Moving ships are an important vector of spread for bioinvasions. Shipping routes are constantly evolving and likely to be affected by the consequences of Climate Change, while at the same time ships are a considerable source of air pollution, with CO<sub>2</sub> emissions at a level comparable to Germany, and NO<sub>x</sub> and SO<sub>x</sub> emissions comparable to the United States. With the development of Arctic shipping becoming a reality, the need to understand the behavior of this system and to forecast future maritime trade flows reasserts itself.

Despite their scope and crucial importance, studies of maritime trade flows on a global scale, based on data and formal methods are scarce, and even fewer studies address the question of their evolution. In this thesis we use a unique database on daily movements of the world fleet between 1977 and 2008 provided by the maritime insurer Lloyd’s in order to build a complex network of maritime trade flows where ports stand for nodes and links are created by ship voyages.

In this thesis we perform a data-driven analysis of the maritime trade network. We use tools from Complexity Science and Machine Learning applied on network data to study the network’s properties and develop models for predicting the opening of new shipping lines and for forecasting future trade volume on links. Applying Machine Learning to analyse networked trade flows appears to be a new approach with respect to the state-of-the-art, and required careful selection and customization of existing Machine Learning tools to make them fit networked data on physical

flows. The results of the thesis suggest a hypothesis of trade following a random walk on the underlying network structure.

After the necessary cleaning and preprocessing of the raw data, we analyse structural properties of the layers consisting of 5 different vessel types: Container Ships, Dry Bulk Carriers, General Cargo Vessels, Petroleum Tankers, and LNG Tankers—accounting for over 90% of total DWT in the database. The basic network properties and their changes over the study period are considered. The network is seen to display super-linear growth of the number of edges with the number of nodes, while the network diameter remains stable, hinting towards network densification. We also find that the degree distribution of different layers and the aggregated network is close to a geometric distribution but the fit is not ideal. In general, we conclude that network aims at enhancing its expansion property while exhibiting its quasi optimal navigability, subject to the degree distribution. We then consider a primary question from the perspective of port authorities: “Which centrality measures are the most pertinent for the success of a port expressed as an increase in throughput?”. We find that the measure with the strongest correlation with an increase in port throughputs is current flow closeness. This centrality measure reflects the central position of a node in the network for exchanges with other nodes following random walk models.

We then turn to uncovering rules which govern the evolution of maritime trade networks by learning the network growth models directly from the network data. We find that the evolution of this system is governed by a simple rule from network science, relying on the number of Common Neighbors between pairs of ports. This finding is consistent over all three decades of data. Additionally, we find that a flow-based model of trade, in which ports are expected to trade with their network neighbors in proportion to the number of their Common Neighbors, provides the most accurate trade flow volume forecasts in a year-to-year horizon.

We further confirm these results by studying the reaction of the network to shocks. Thanks to a natural experiment, involving traffic redirection from the port of Kobe after the 1995 earthquake, we find that the traffic was redirected preferentially to ports which had the highest number of Com-

mon Neighbors with Kobe before the cataclysm. Then, by simulating targeted attacks on the maritime trade network, we analyze the best criteria which may serve to maximize the harm done to the network and analyse the overall robustness of the network to different types of attacks. All these results hint that maritime trade flows follow a form of random walk on the network of sea connections, which provides evidence for a novel view on the nature of trade flows.