Network effects govern the evolution of maritime trade

Zuzanna Kosowska-Stamirowska[a,1,2]

[a]Géographie-Cités, CNRS & Université Paris 1 Panthéon-Sorbonne, 75006 Paris, France

Edited by Jon Kleinberg, Cornell University, Ithaca, NY, and approved April 15, 2020 (received for review April 24, 2019)

Maritime transport accounts for over 80% of the world trade volume and is the backbone of the global economy. Global supply chains create a complex network of trade flows. The structure of this network impacts not only the socioeconomic development of the concerned regions but also their ecosystems. The movements of ships are a considerable source of CO_2 emissions and contribute to climate change. In the wake of the announced development of Arctic shipping, the need to understand the behavior of the maritime trade network and to predict future trade flows becomes pressing. We use a unique database of daily movements of the world fleet over the period 1977–2008 and apply machine learning techniques on network data to develop models for predicting the opening of new shipping lines and for forecasting trade volume on links. 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 temporal data. We further confirm it with a natural experiment, involving traffic redirection from the port of Kobe after the 1995 earthquake. Our forecasting method enables researchers and industry to easily model effects of potential future scenarios at the level of ports, regions, and the world. Our results also indicate that maritime trade flows follow a form of random walk on the underlying network structure of sea connections, highlighting its pivotal role in the development of maritime trade.

maritime trade | transport networks | network science | evolving networks | machine learning

In global supply chains, centers of production and consumption are far away from each other, creating a complex network of trade flows. Over 80% of all cargo, in terms of volume, is carried by sea, accounting for 70% of the total value of international trade (1), with ships being the least expensive means of transportation in terms of marginal cost per item (2). Maritime transport is regarded as the backbone of global trade and of the global economy (3).

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 (4, 5), especially for marine species. At the same time, the future of the maritime transport industry is inextricably linked to climate change: The movements of ships contribute significantly to global CO_2 emissions (6), and, conversely, future shipping routes are likely to be affected by the consequences of climate change. With the development of Arctic shipping becoming a reality, the need to understand the behavior of the system of maritime trade flows and to forecast their future evolution reasserts itself.

Despite the obvious and crucial importance of maritime logistics to the world economy, only a few works provide a detailed overview of the global distribution of maritime trade flows (4), and even fewer analyze their long-term evolution and the rules which govern it (7, 8).

Current research on maritime transport and ports tends to focus on particular operators or segments of shipping industry, regions, and specific snapshots in time (9–12), whereas studies analyzing historical evolution of the maritime network as a whole are scarce (13). This is presumably due to the difficulty of development or acquisition of a global temporal database on maritime trade flows. The only comprehensive statistical source allowing for the representation of global maritime trade flows accurately and over a long period, developed by the main maritime insurer, Lloyd's, has only begun to be exploited, due to restricted access to these sources and lack of adequate technical tools. In this study, we aim to fill this gap by using data on daily movement of the world fleet between 1977 and 2008 provided by Lloyd's List Intelligence. We treat the available data (16.9 million recorded ship voyages) with tools from complex systems and machine learning on graphs. Our goal is to enable the extraction of economically viable information about trade and trade volumes from data that are not only port-specific but also depend on the broader structure of connections between the ports established by the movements of ships. This approach is motivated by the fact that vessels are functionally comparable to road or rail infrastructure and should be seen as such while designing transport or trade policies (14). In this paper, we introduce tools to model maritime trade flows and simulate the effects of potential shocks or changes to the system at local, regional, and global scale.

Network generative models are used to explain mechanisms of network evolution by assigning probabilities of creation to individual potential links based on characteristics of the involved nodes. From the point of view of complexity science, generative models for real-world networks have been in the spotlight (15–17) ever since the acclaimed preferential attachment model was proposed by Barabási and Albert (18). Our work develops a

Significance

Over 70% of the total value of international trade is carried by sea, accounting for 80% of all cargo volumes. Maritime trade flows impact both the economic development of the concerned regions and their ecosystems. 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 pollution. This work performs a rigorous and comprehensive analysis of maritime trade flows on a global scale over a long period, taking into account aspects of evolution, reactions to shocks, and discovering predictive models. We use machine learning methods to uncover the single dominant dynamics that govern the evolution of the structure and the intensity of maritime trade.

Author contributions: Z.K.-S. designed research, performed research, analyzed data, and wrote the paper.

Competing interest statement: Z.K.-S. is affiliated with NavAlgo.

This article is a PNAS Direct Submission.

[1]Present address: NavAlgo, 91120 Palaiseau, France.

[2]Email: zuzanna@navalgo.com.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1906670117/-/DCSupplemental.

First published May 26, 2020.

generative model of link opening specifically for the maritime trade network. We also apply a similar methodology to explain maritime trade flow volumes over existing links. In both scenarios, this study does so based on temporal, historical network data.

Our research design consists in uncovering the rules governing the evolution of maritime trade network ex post and is almost entirely data driven. The network features, which are the input variables fed into the considered learning methods, include both a preselected set of conventional network measures found with a metaoptimization technique and port characteristics suggested by the literature. We then feed combinations of these features into different learning methods to automatically construct the best models for predicting link creation and future trade flows. This process produces models that hit the sweet spot between prediction accuracy and the amount of data needed as model input. We believe that the scientific approach proposed in this work, which consists in automatic learning of network generative models directly from available temporal data (analysis of newly created edges and weight changes between two time frames), may also be of independent interest in other areas of complexity science.

In this study, we have found that models relying on features of the pure network topology, in some cases augmented by additional information on sea distances, are the most powerful for predicting link creation and estimating trade flows in maritime trade for all vessel types. In all contexts, we obtain the most accurate forecasts when relying on one specific network feature: the number of common neighbors between a pair of ports, understood here as the number of other ports that are simultaneously trading partners for both of the ports in question (see Fig. 1 for illustration). We find that the relative quality of models is stable across the studied period (1977–2008).

We have evaluated the performance of models relying on network features against classical gravity models that use distances and data on demographic and economic development to describe the affinity between ports, an approach popular in geographic and economic literature (2) and frequently used as a baseline for evaluating the significance of other effects [such as the network connectivity index (19)]. Our research methodology encompasses both types of models and automatically tunes their parameters (where applicable) to their best performance. We have found that classical gravity models perform very poorly and that, in the case of maritime trade flows, they are very far from a "fact of life" (20). Indeed, our results suggest that classical gravity is almost as poor a baseline model for maritime trade as a hypothesis of random uniform link creation that requires no input. We also benchmarked against a tailored variant of the gravity model which uses, as weights, port throughputs in the gravity equation (4), with affinity between ports moderated through their sea distance and measures of cultural and historical ties [such as sharing a common language, country, or colonial origin (21)], in a combination tuned to optimal performance. We call this model port gravity. We have found port gravity not to perform well in predicting link creation and to perform moderately well in predicting actual trade flows. It seems that these models miss an important piece of the story: that affinity between ports is influenced by network effects which arise from preexisting network infrastructure. The nonnetwork variables conventionally used in the gravity models are no more than a partial proxy for some of these effects (*SI Appendix*, section L).

Armed with this knowledge, we have additionally tested the effect of common neighbors in a natural experiment—the destruction of the port of Kobe by an earthquake in 1995 and the subsequent redirection of trade flows forced by this unfortunate event. We have found that the simple number of common neighbors shared with Kobe in 1994 successfully identifies ports taking over traffic from Kobe, at least as well as and, in some ranges, better than the models relying on economic data. This connection further supports our claim that network effects, specifically, network-based affinity, govern the evolution of maritime trade.

Looking more generally, our findings support a vision of trade in which units of goods follow a form of random walk on the underlying network structure. The observed effects provide hints as to the precise nature of this random walk process which turns out to be local, as it relies on information accessible only in the port's neighborhood. This effect is observed despite the nature of the maritime industry, which involves strong concentration of capital and very few global key players controlling most of the world's shipping market (22).

Results

Dataset and Approach. We use a unique, temporal dataset of daily movements covering the majority of the world fleet (*SI Appendix*, section A) between 1977 and 2008, developed by the main maritime insurer, Lloyd's, licensed to the European Research Council (ERC) "World Seastems" project. The raw database has been cleaned and treated to extract only meaningful movements of ships (*SI Appendix*, section B).

For the learning process, we construct yearly snapshots of the maritime trade network, where ports stand for nodes, and links are created by ship voyages between two ports in a given year. A link is created as soon as the total vessel dead weight tonnage (DWT) transported along it exceeds a certain threshold (*SI Appendix*, section C, as well as *SI Appendix*, section K where we consider an alternative network definition for liner shipping).

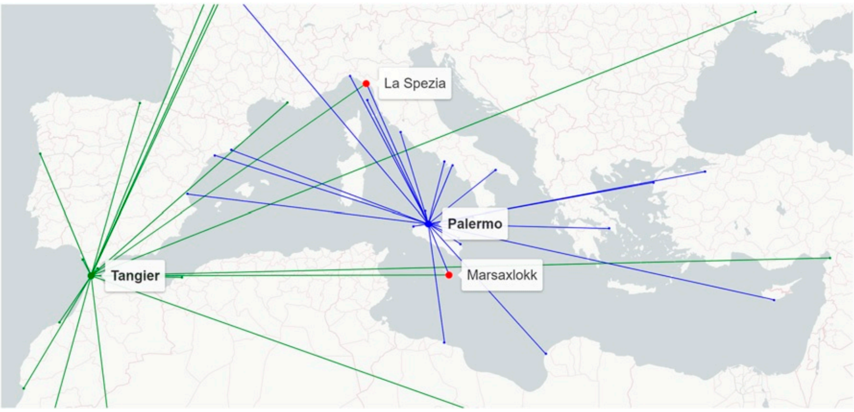


Fig. 1. The concept of common neighbors. The two common neighbors (or "common connections") of the ports of Tangier and Palermo are Marsaxlokk (Malta Freeport) and La Spezia, represented by the full red nodes. This example comes from the studied network of container carriers in 2007.

Modeling maritime trade as a network has recently become a common practice in the industry (23) and is justified by the relative rigidity of sea connections, resulting from both technological reasons (port types, infrastructure, depth, capacity etc.) and business reasons. In our study, the overlap of links appearing in the network in two successive years is in the range of 55 to 65% (reference value for model of uniform independent random connections: 0 to 1%, depending on year and vessel type), with more-frequented links being preserved more frequently.

We construct a separate network for each of the main commercial vessel types: container carriers, dry bulk carriers, general cargo vessels, petroleum tankers, and liquefied natural gas tankers (accounting, in total, for 93% of the world fleet DWT in the database). This division into subnetworks is made because it is known that different vessel types follow different movement patterns (4, 13), and some use specialized ports.

In the modeling process, we first consider a broad set of features (characteristics of ports and the links between them), suggested by the literature on network formation (24–26) and works on spatial interactions and international trade (2, 27, 28). The considered port characteristics include demographic and gross domestic product (GDP) economic variables, such as population potential and country (*SI Appendix*, section D) and port throughput, as well as features that arise from the comparative place of the port in the broader network of maritime connections: so-called network centrality measures. Link characteristics include features that arise from relations between the pairs of ports: sea distance, hop distance, number of trade partners in common (common neighbors measure), or cultural and historical ties. A model F based on port and link features produces, in general, a prediction value $w_{i,j}$ which, depending on the context, describes the probability of opening a link between a pair of ports, i, j, or describes the flow value between these ports. This is written as

$$w_{ij} = F(\mathit{PortFeatures}_i, \mathit{PortFeatures}_j, \mathit{LinkFeatures}_{ij}). \qquad \textbf{[1]}$$

The following research design ensures that we discover the best of the models constructed with the available features. In order to preselect dominant features which are strongly reflected in the observed network structure, we apply a metaoptimization technique: symbolic regression (29) on formulae of generative network models for a static snapshot of the network.

We then feed learning methods with combinations of these dominant features, as well as other data suggested by the literature on gravity in trade, and evaluate the predictive quality of the obtained models (*Materials and Methods* and *SI Appendix*, section D). This learning approach generalizes and subsumes any ad hoc polynomial combinations of features (sometimes known as indices; cf. refs. 30 and 31) which could be designed by hand. In all cases, in this final step, we use features from each year Y to predict new links and flows in year Y+1. The models are cross-validated on random cuts of the network into training and testing sets (see *Materials and Methods*). This approach has allowed us to identify the best models based on their predictive quality, while minimizing researcher bias. In this way, we have been able to look beyond and possibly bypass the most intuitive gravity-based models.

Model for Predicting Link Creation. We first obtain a model for predicting link creation in the maritime trade network. We do so by testing models consisting of different sets of features that include network measures and port characteristics. We compare the relative performance of these models by analyzing the proportion of the number of correctly predicted link openings to the number of all new links that were created in Y+1. This measure was chosen to best reflect the goals of the model in real-world applications, where the identity of the created links is crucial, and, in this sense, it is more demanding than network similarity measures used in the literature of network generative models (29) which consider overall network structure and statistics.

The most successful model that we discover turns out to be the same for all vessel types, and its relative quality is stable across the studied period. It relies on the number of common neighbors between a pair of ports tempered by sea distance between them (Fig. 2). Intuitively, this means that the probability of link creation in the maritime trade network is proportional to the number of neighbors the two ports have in common and decreases the farther apart these ports are from one another.

We find that models which rely on the number of common neighbors, even when they are very simple, deliver much better predictions of link creation than any gravity model. Our network generative model with common neighbors and sea distance, which uses no external economic data, correctly predicts 19 to 24% of edges created from one year to another depending on the vessel type, on average, over the studied period (Fig. 2 and *SI Appendix*, sections G, H, J, and K). Moreover, a parameter-free variant of our model using only a single variable—the number of common neighbors—achieves comparable prediction results at 19 to 23% (lower by at most 1 percentage point [pp]). For completeness, we compare the obtained models to those previously proposed for the purpose of estimating maritime trade flows. We find that port gravity, which uses data on port performance (throughputs) as weights, predicts, on average, 14 to 20% of links depending on the vessel type (worse, by 3 pp to 6 pp, than common neighbors with sea distance). The contribution of features based on cultural and historical ties to the performance of port gravity is negligible in all cases, not exceeding 1 pp.

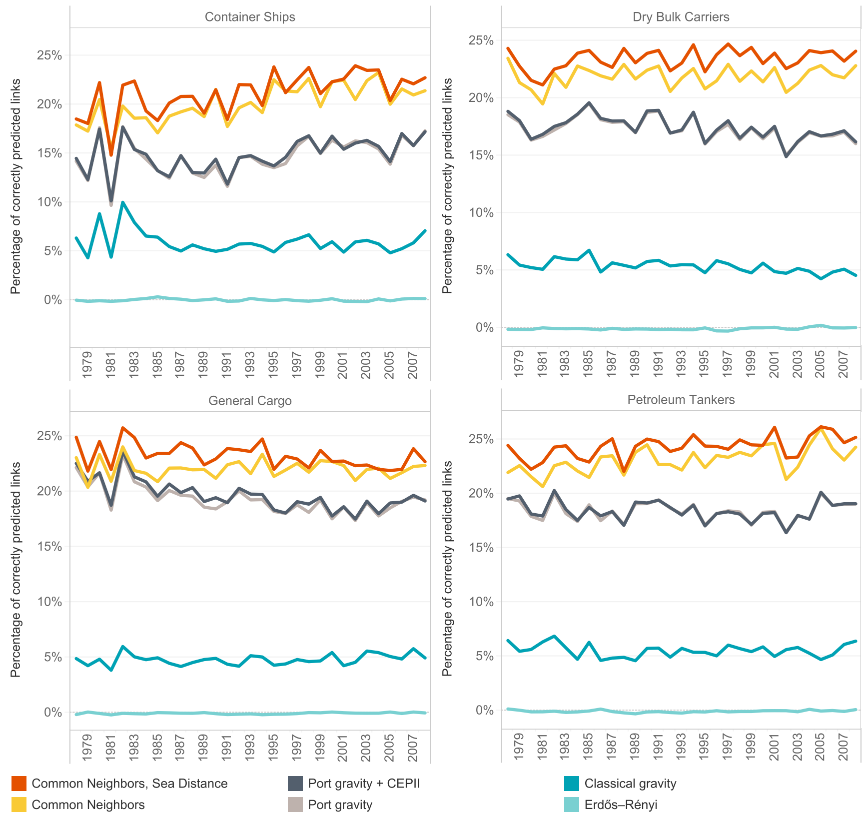
The advantage of the common neighbors-based models is even more apparent when identifying those potential links which are the most likely to be opened in the coming year (e.g., identifying new links with a fixed low ratio of false positives). The performance of the model relying only on the number of common neighbors, compared to port gravity, is illustrated by the sample maps in Fig. 3 which show the set of interregional link openings predicted with at least 50% accuracy by each of the models (less than 50% of false positives). Depending on the precise setting, the simple parameter-free common neighbors model correctly predicts 2 to 5 times more link openings than port gravity, even when the parameters of the latter model are tuned specifically to each instance (also *SI Appendix*, section I).

The obtained results show that geography (distance) plays a role in networks built by some types of vessels, whereas the effect of common neighbors is stable across vessel types, capturing a universal and crucial feature for the creation of direct links in maritime trade. The link prediction quality delivered by models relying only on common neighbors is, on average, very close to that of models which additionally include sea distance. We note that the number of common neighbors is only very weakly correlated with sea distance, and the sign of the correlation depends on the vessel type (*SI Appendix*, section L).

We further note the poor performance of the classical gravity model (2, 32) that relies on population potentials of ports, country GDP values, and sea distance. This model, on average, manages to correctly predict 5 to 7% of newly created edges.

We also observe that combining data on GDP values or population potentials with the number of common neighbors does not lead to an increase in prediction accuracy of the common neighbors model, which indicates that these two variables have no further added value for predicting links in maritime trade (*SI Appendix*, section M).

Interpretation of the Common Neighbors Effect. One possible explanation for the good performance of the common neighbors effect when predicting link creation in the maritime trade network is the tendency to shortcut routes. The number of common



Model	Container Carriers	Bulk Carriers	General Cargo	Petroleum Tankers	LNG Tankers
Common Neighbors and Sea Distance	21%	23%	23%	24%	19%
Common Neighbors	20%	22%	22%	23%	19%
Port gravity	15%	17%	20%	18%	14%
Port gravity (excl. cultural ties)	15%	17%	19%	18%	14%
Classical gravity	6%	6%	5%	6%	7%

Fig. 2. Percentage measure of correctly forecasted link openings by different models between 1978 and 2008. Predictions for year Y+1 are based on features from year Y. Averaged results over the period are provided in the table. In the design of the measure, the number of links to be predicted by each model was set to the number of links actually created between the consecutive years, thus, false positives ratio = false negatives ratio = (100% − percentage measure). The relative quality of models for link forecasting is stable across the studied period. The models relying on the number of common neighbors are consistently the best, for all vessel types.

neighbors between a pair of ports corresponds to the number of two-hop paths in the network between them. Two ports sharing a large number of neighbors can be expected to have a higher volume of indirect trade flow between them which passes via the ports they are both connected to. Creating a direct connection instead of using intermediary ports reduces transport costs by avoiding costly transshipment.

In this context, the common neighbors effect in link creation is also consistent with the representation of transport flow as a Markov process on the network (33). Consider a simple example of a memoryless model in which a unit of trade located at a port i is equally likely to be transported directly to any of the neighbors of i. In a regular network topology (i.e., one with identical degrees), the distribution of ports at which the unit may be located after two steps of such a walk corresponds mathematically to that given by the common neighbors measure for the port of origin, normalized over all possible destination ports. We discuss more-advanced Markovian models in subsequent sections, where the random walk is biased toward certain neighbors.

In other domains, the number of common neighbors has been proven to work successfully as an affinity function in the case of mining hidden links in social networks (34) and as an efficient technique for identifying links in theoretical block models of community-type networks which are large and dense (35). We note that the results of our study show that the common neighbors model for link creation in maritime trade network is consistently superior to a preferential attachment model for link creation, which is covered in our work by using node characteristics, such as degree (*SI Appendix*, section N). Port gravity is a special variant of preferential attachment of links in which port throughputs replace degree.

Model for Forecasting Flow of Cargo on Links. After the warmup with link creation, we now obtain models for forecasting flows in the maritime trade network, expressed as tonnage (DWT) transported over the existing links. We evaluated the performance of models consisting of the same feature sets as previously, in combinations extended to the weighted scenario (*SI Appendix*, section F). We tested both models of the form given by formula Eq. **1**, where w_{ij} is understood as an estimate of the absolute flow value between ports i and j in year Y+1, as well as models in which w_{ij} represents the share of the flow outgoing from port i toward port j. Most generally, in the latter case, the estimated flow value in year Y+1 between ports i and j is computed as

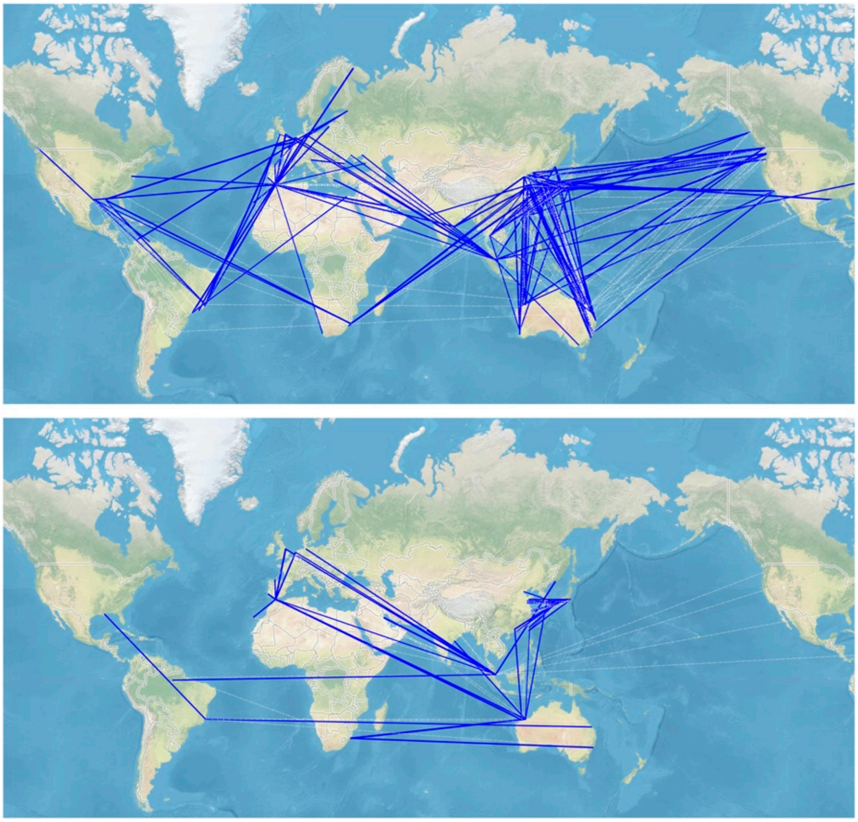


Fig. 3. Forecast of interregional link openings for dry bulk carriers in 2008 based on 2007 data. Predictions made with at least 50% accuracy using (*Top*) common neighbors model (306 predictions, of which 153 are correct; parameter-free model) and (*Bottom*) port gravity model (84 predictions of which 42 are correct; model parameters tuned to problem instance). Correctly predicted links are marked in blue (line width increasing with actual DWT of opened link), and false positives are marked in dashed white. Interregional links are chosen for visual clarity. Extending considerations to all links in the network leads to quantitatively similar results.

$$\Phi\left(\frac{w_{ij}}{\sum_{k\in N(i)} w_{ik}} DWT_i, \frac{w_{ji}}{\sum_{k'\in N(j)} w_{jk'}} DWT_j\right), \qquad [2]$$

where DWT_i represents the total DWT passing through port i in year Y, $N(i)$ is the network neighborhood of port i, that is, the set of direct trade partners of port i in year Y, and Φ represents a function fitted by a learning method. In the rest of the paper, we refer to the models following Eq. **2** as the "outflow models." In this study, unweighted network data and port throughputs (port DWT) from year Y were used to predict flows in year Y+1.

We find that the outflow models relying on the number of common neighbors and sea distance to preferentially assign shares of cargo to the outgoing links deliver the best forecasts of the actual flows of cargo in maritime trade network, irrespective of vessel type. For the vessel types considered, they achieve a coefficient of determination for log-flow in the range 0.33 to 0.45 (averaged over the studied period; also *SI Appendix*, section I) and outperform all other models tested (Fig. 4 and *SI Appendix*, Table S2). For the alternative network definition reflecting liner shipping patterns for container vessels, the coefficient of determination is 0.48 (*SI Appendix*, section K). The relative predictive quality of these outflow models is stable over the studied period.

The common neighbors feature appears to be particularly well suited to the outflow model. In our study, assigning flow shares to outgoing links based on sea distance, only (without the common neighbors feature), reduces the coefficient of determination by 0.10 to 0.14. By contrast, the importance of including sea distance in addition to common neighbors in the outflow model varies across vessel types. For dry bulk carriers and for tankers, the outflow model based only on common neighbors offers good prediction quality (coefficients lower by 0.03 to 0.05 than for the outflow model also including sea distance), whereas, for container and general cargo vessels, it performs significantly worse (coefficients lower by 0.12 than the model also including sea distance) (*SI Appendix*, Table S2).

Comparing the outflow model relying on common neighbors and sea distance with the best absolute flow models considered (i.e., those following Eq. **1**, which include in particular port gravity), we observe, for all vessel types, a significant gap in prediction quality in favor of the outflow model. This difference is at least 0.06 for all vessel types and is most marked for container and general cargo vessels (difference of 0.11).

Within the class of models following Eq. **1**, port gravity performs best. However, for all vessel types apart from tankers, it is closely rivaled by a model relying only on common neighbors with sea distance, which does not require information about port DWT (coefficient differences of 0.01 to 0.03 with respect to port gravity). Neglecting historical and cultural ties in the port gravity model has a very slight effect on the coefficient of determination, reducing it by 0.01.

Classical gravity models which use population, GDP, and sea distance between ports do not perform well, and their ability to predict future trade volumes is close to that of models of uniformly random network formation.

Interpretation of Outflow Models. Assuming a sufficiently simple form of function Φ (also *SI Appendix*, section E), the outflow models can be interpreted in the following way: Every unit of trade outgoing from a port follows a weighted random

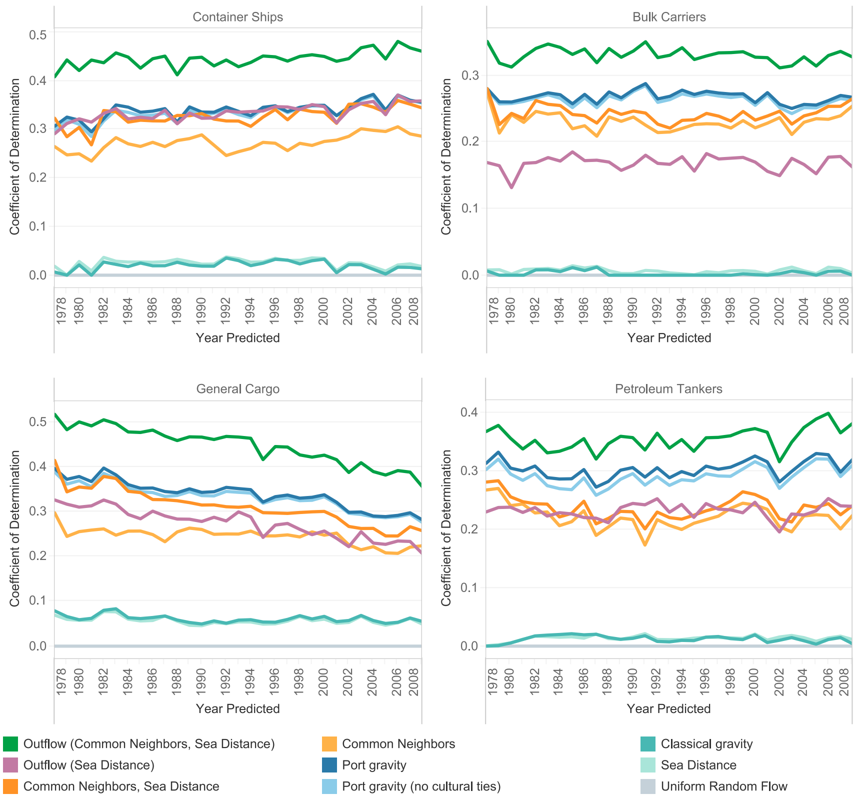


Fig. 4. Predictive quality of models for tonnage transported on links of the maritime trade network, 1978–2008. Values are obtained using the better of the two regressors: random forest and OLS averaged over 24 runs. The best prediction quality is obtained consistently with outflow models including the common neighbors feature.

walk. In this way, the total throughput of a port is distributed among its neighbors, proportionally to the weights described by the model.

The best outflow model relies on weights w_{ij} computed as a function of the number of common neighbors between the ports and their sea distance. Thus, the share of outgoing trade between a port i and its neighbor j increases if the receiving port j ranks high in the number of common neighbors with i and decreases with the distance between the ports.

In the more illustrative case of the outflow model depending only on common neighbors (which already achieves very good performance for some vessel types), the matrix of weights w for the flow process is compactly given as $w = G \odot G^2$, where G is the adjacency matrix representing the unweighted topology of the network, G^2 is the matrix square of G (describing the number of two-hop paths between node pairs), and $\odot$ represents element-wise multiplication of matrices, ensuring that flows only follow existing links. Thus, in this case, the flow weights depend only on the unweighted network topology through an algebraic transformation. This allows us to draw an analogy between the outflow model and a process of shortcutting two steps of an unweighted random walk (i.e., a walk with weight matrix G) into a single step: A walker simulates where she would end up after two steps, then checks whether there is a direct link and, if there is one, goes through this shortcut. When the graph is regular, this equivalence is mathematically precise, since the matrix G corresponds to the random walk transition matrix (up to scaling). We find such a lookahead property of the walker, who can plan its voyage two steps in advance, natural in the context of economic exchanges. For the best outflow model depending on both common neighbors and sea distance, the weight matrix is computed with an additional dependence on the sea distance matrix SD, as $w = G \odot G^2 \odot SD^{-\alpha}$, where $\alpha > 0$ is a deterrence parameter (typically, $\alpha \in [0.25, 0.50]$, depending on the vessel type, with little sensitivity to the precise parameter value).

Extending the link creation models which were considered in the previous section, the outflow model makes use of the common neighbors feature to perform a reinforcement of direct routes (shortcuts) between pairs of ports which have many trading partners in common. We conclude that the number of common neighbors between ports appears to be the key feature controlling the actual volumes of trade flow in the maritime trade network.

Traffic Redistribution after the Closing of the Port of Kobe. We confirm our findings for the mechanisms which govern the evolution of the network by looking at a localized event of profound significance. On 13 January 1995, the port of Kobe, one of the busiest container ports in the world, was destroyed by a very severe earthquake (36). The port had to shut down completely, and the traffic which was originally supposed to flow to Kobe had to be redirected to other ports. This event created a window of opportunity for other ports to capture Kobe's traffic, with consequences lasting long after the port of Kobe reopened. Kobe has never regained its previous importance in the maritime system, with traffic never reaching the pre-1995 level.

This unfortunate event creates a rare natural experiment in which a perfectly exogenous shock results in the closure of a single important port, forcing an adaptation of flows in the network. The event caused a redirection of traffic on the order of 200 million DWT, which is almost of an order of magnitude less than the year-to-year traffic changes in the global network in the 1990s (mean change +1,100 million DWT year-to-year, $\sigma = 800$

million DWT). Consequently, we used more-fine-grained network data and information about the topology of shipping connections to reinforce the signal-to-noise ratio. The signal related to the Kobe shock was isolated by restricting our considerations to the flows of an unbiased representative subnetwork H, with a cutoff point depending on the share of trade with Kobe before 1995 (see *Materials and Methods*). Between 1994 and 1995, the exceptional change of Kobe's trade volume in exchanges within subnetwork H was -113 million DWT ($\sigma = 7$ million DWT), whereas the exceptional traffic increase of the remaining ports of the subnetwork H in 1995, excluding Kobe, is given as +111 million DWT ($\sigma = 46$ million DWT), corresponding well to the traffic lost by Kobe. The reported values of exceptional changes are controlled for year-to-year traffic change patterns in the period 1990–2005, excluding the years 1994 and 1995; see *SI Appendix*, section P for details on detrending. The strength of the Kobe shock for the considered subnetwork H is thus at least 2.4σ. Moreover, the 1995 increase of traffic in the subnetwork H is clearly the largest such event over an even longer time span (from the oil price shocks until the end of the available data [*SI Appendix*, Table S5]). This means that the redistribution of trade after the Kobe shock was absorbed by the subnetwork H itself. This also allows us to draw statistically significant conclusions about the redistribution of flows in the subnetwork H, even if signals are weaker for some individual ports.

We find that, in the short term, traffic redirection from Kobe to other ports in 1995 is linked to the number of common neighbors with Kobe of the port in question directly before the cataclysm. The 15 ports which had the highest number of common neighbors with Kobe in 1994 (*SI Appendix*, Fig. S13) accounted for 50% of the DWT traffic on the links of the subnetwork H. In 1995, these ports captured a disproportionately large amount of traffic: 98 million DWT (after controlling for year-to-year changes; $\sigma = 16$ million DWT). Thus, the ratio between the detrended traffic increase of these 15 ports and the detrended traffic loss of Kobe is likely to be disproportionately high at 85% ($\sigma = 19\%$).

The common neighbors measure does not rely on any economic data besides unweighted network topology. We also considered measures which included data on throughputs and sea distances: port gravity, port throughput, and DWT exchange with Kobe pre-1995. In our study, these methods prove much more volatile and do not achieve comparable statistical significance to the number of common neighbors with Kobe, nor do they appear to provide any advantage, on average (*SI Appendix*, Fig. S14).

We propose the following interpretation for the observed effect of common neighbors. First, the number of common neighbors can be seen as a measure of similarity of nodes in trade networks (37, 38). From a business perspective, a port having a higher number of neighbors in common with Kobe was able to serve the same markets as Kobe did before the earthquake and was therefore a good replacement. Second, by a purely topological argument, traffic flowing through Kobe from one port i to another port j is most easily rerouted using existing link infrastructure, via a third port l which has both i and j as its common neighbors with Kobe. Thus, the larger the number of common neighbors between Kobe and a port l, the larger this redirection may be expected to be. Overall, this is consistent with the view that trade flows follow a form of random walk on the underlying network structure, with the next hop of the walk being chosen from the set of the available neighbors.

Discussion

Concluding Remarks. Maritime transport is "the backbone of global trade and the global economy" (3). Here we present the rules according to which the system built by ports and ship passages evolves over time. We have found that the opening of new shipping lines and the volume of flows of cargo on existing links depend on the underlying network structure. The topological feature which offers the best prediction results, in both cases, is the number of common neighbors shared by the two ports. This contribution improves our understanding of a crucial part of our economy and may be used to simulate future scenarios.

We find that the evolution of the maritime trade system can be explained by looking only at the unweighted network topology of shipping connections (sometimes augmented with sea distances). On the one hand, the models of evolution of the unweighted topology which we put forward do not depend on trade volumes (DWT or ship counts). On the other hand, in the flow forecasting scenario, we can predict how ports distribute volumes of their trade along the network links adjacent to them, by considering only the unweighted topology and sea distances on existing links. Using port total DWT, we then obtain good forecasts of link DWT in absolute terms. The quality of such forecasts is significantly better than that obtained with gravity-like models, even those using port DWT.

The discovered importance of the number of common neighbors for forecasting the opening of links in the maritime trade network and the good performance of the outflow models provide an indication that trade flows follow a form of random walk on the underlying network topology. They also give us further insights into the precise form of this random walk. It has been remarked that tankers follow Brownian motion, in contrast to container carriers which follow regular movement patterns (4). However, our research suggests that actual containers (trade flows) also follow a form of random walk on the network, thus giving rise to the observed network effects.

Our research implies that the status quo structure of the maritime transportation network is a much stronger factor for predicting future trade flows than population potentials or GDP of countries. This observation suggests that it is the transportation links and their structure that foster the development of maritime trade, rather than the other way around. We believe this to be a piece of the puzzle in the perpetual debate on which comes first: economic development or transport infrastructure (39).

Trade models have been studied extensively in the economics literature (40). To the best of our knowledge, the only other works that develop a flow model for maritime trade rely on gravity (19) or other economic variables (8). Our work is a comprehensive study that proposes a network generative model as well as a model for forecasting flows in a maritime trade network, both of which are developed and tested on a large, temporal database of ship movements. We note that our results are remarkably stable throughout the period analyzed (1978–2008).

We are not aware of any similar trade or transportation network for which comparable temporal effects have already been observed. We note that the setting of our study is significantly different from studies of nontemporal data, where the notion of link prediction is typically equated with the question of detecting hidden links in a static network (30, 41), allowing the reconstruction of the current network topology based on incomplete data. In the hidden link setting, the common neighbors effect has been singled out for social networks (34). Of the variety of approaches that have been developed for link prediction in complex networks (31, 38), relatively few have been validated in forecasting settings on real-world temporal datasets, typically for datasets related to social networks and communication patterns in online media (cf. e.g., ref. 42). Whereas, in studies of social networks, the common neighbors measure is often treated as a baseline for other methods, up to now, for transportation networks, it has usually been treated as one of many similar network indices, and perhaps not the most intuitive.

From the perspective of complex network methodology, our approach to temporal networks provides a simple and robust data-driven framework for selecting optimal combinations of relevant sets of features, minimizing research bias. It is sufficiently powerful to subsume so-called local network approaches given by so-called indices (31), as well as other polynomial and nonpolynomial combinations of network features and external ones. The obtained models can typically be analyzed by human researches, analogously to formulae obtained through symbolic regression (29). Our methodology also allows for the inclusion of historical data as node and link features, enabling certain stochastic learning approaches (42).

While noting the good prediction quality of the proposed models using the common neighbors feature, the models do not explain the maritime trade flow in its entirety. The discovered importance of the number of common neighbors implies that network evolution is a consequence of a cumulation of network effects and exogenous shocks. We have provided for at least a partial identification of the network effects concerned. However, our models do not reveal to what extent the unexplained effects are directly attributable to exogenous shocks. Further network evolution models can be constructed using machine learning methods tuned to temporal network data (43) and higher-order flow models with memory (33), most likely at the cost of sacrificing simplicity, usability, and interpretability of the models. One such attempt to extend our results may involve considering nonlocal network features of nodes by considering network representations which capture the spectral structure or random walk structure of the network, with benchmark references possibly provided by the seminal deepwalk (44) and node2vec packages (45) and their hierarchical variants (46). Some preliminary results obtained using such graph embeddings are discussed in *SI Appendix*, section H. An alternative approach may rely on node embeddings in low-dimensional hyperbolic spaces (47, 48).

We also note that the passages of ships do not correspond directly to the flows of goods and that we analyze ship movements, not the bills of lading. Finally, we note that maritime trade does not capture the entirety of the world trade. Further studies on the world trade network with tools from complexity science, like ref. 49, may be performed to investigate the nature of trade flows complementary to the maritime realm.

Advantages of the Proposed Models. The strength of the proposed forecasting models lies in their simplicity, the fact that the common neighbors measure is parameter-free, and the fact that the models require small input size to deliver accurate predictions. The models for both forecasting link opening and flow values require the knowledge of the direct network neighborhood of ports, without the need for knowledge of the global network topology or the development of large economic datasets. In the case of forecasting link openings, it is just the one-hop network neighborhood of the two ports and sea distance which are needed to assign creation probability to an individual link. Indeed, the set of plausible candidate link openings for a port is both restricted and easy to find by local search. In the case of forecasting flow on a link, one needs to know only the throughput volumes of the two ports concerned, sea distances on the existing links of each of the two ports, and the network topology to the extent of the two-hop network distance from each of the two ports. Our models can thus be seen as local, which makes them substantially different from, for example, the preferential attachment model, which, in its basic form, requires global knowledge of the network.

Finally, the fact that common neighbor-type network effects govern evolution of the maritime trade provides strong evidence for the irreversibility of shocks, as changes to the structure of shipping lines are likely to trigger subsequent changes in the common neighbors measure, propagating throughout the network. The observed network effects also offer an intuitive and relatively simple way of estimating the lasting consequences of shocks.

Areas of Application. We trust our work will benefit researchers in fields such as economics, ecology, and even studies of cultural exchanges in archeology (50), providing them with tools to model the evolution of maritime trade and to simulate the effects of potential shocks or changes to the system at local, regional, and global scales.

Most pressingly, according to Russian prognosis for Arctic shipping, traffic on the Northern Sea Route is expected to rise to 80 million tons of shipments per year by 2025 (51). This rapid development generates economic, environmental (52), political, and social (53) challenges which are at the center of attention of multiple governments. The development of Arctic shipping is likely to impact not only the countries that are directly involved in the Arctic routes and ports due to their geographical position but also those from whom the traffic will be redirected and those who can become new transshipment hubs. This will be the first time since the blockage of the Suez Canal in the 1950s that sea distances will be effectively modified. Understanding the model underlying maritime trade is essential to predict the response of the system. In the common neighbors model, new connections or flow depend on the previous form of the network, and the effects of shocks propagate over its entirety, as one reconfiguration modifies creation probabilities and tonnages for multiple other links. Shortening of sea distances in the network as a result of Arctic shipping can be seen as an exogenous shock and will certainly be a trigger for the opening of new sea connections. Based on this research, we expect further reconfiguration of the network to depend on network effects. In the case of Arctic shipping, the identity of ports that will be the first to serve cross-Arctic voyages will be of crucial importance to further network evolution, along with their network of connections. This stands in contrast to predictions which could potentially be obtained with gravity models, which are not sensitive to the order in which links are created.

Our results also make possible the forecasting of maritime trade flows on shipping lines for the purpose of business intelligence and decision support tools. Simulations of shock scenarios can be crucial for port operators, shipping companies, and policy makers. One can model a closing or appearance of a port or passage, or sudden increases and falls in trade volumes on individual links. These simulations can help to conduct cost—benefit analyses of development projects for port infrastructures, canals, and opening of new shipping lines by shipping companies. They can also facilitate vulnerability analyses for policy makers. Governments can potentially benefit from the models developed in this work in order to position (benchmark) themselves with respect to other players in maritime trade. These models can also help them to take informed decisions, reached by testing different strategies and their impact, taking into account the entire maritime system and not just the local perspective.

Another area of application of our models is related to epidemiology and bioinvasions, such as the spread of species which are often transported in the ballast water of ships across the globe or by insects, birds, and other animals that come onboard a ship. Cruise ships are also known to disseminate infectious diseases such as influenza (54). Our models may help researchers recreate or model future vectors of spread of epidemics or bioinvasions, without recourse to a complete global database.

With respect to all potential applications, we highlight that the classical models of gravity, in which one, first of all, takes into account the economic and demographic development of pairs of places and the distance between them, are not well adapted to maritime trade, even as a baseline model. It appears that the network structure of maritime connections bears much more information about the future of maritime trade and that it is therefore of crucial importance for economic development.

Materials and Methods

Learning Model Selection. Manually selected combinations of features were fed into appropriately chosen learning methods.

The considered task of forecasting link openings is one of binary classification (deciding whether a given potential link is created or not). In the setting where we require the classifier to return a link creation probability, we found that the simple logit classifier provided almost optimal predictions for the considered feature combinations. This optimality was verified, in particular, by comparison to a bucketing classifier for low-dimensional models (e.g., for the case of a parameter-free model such as common neighbors, this classifier corresponds to ranking potential links by the considered feature). The results presented in the paper correspond to values obtained with logit. The results for port gravity presented in Fig. 3 were optimized by exhaustive search over the parameter space of log-linear models.

For the flow forecasting task, we aim at predicting log values of DWT transported over the links based on features from the previous year. We applied supervised learning with training and test sets of equal size. We tested a number of regressors, including multilayer perceptron neural networks, random forest, and linear regression models (including ordinary least squares [OLS]). We then compared their performance to the outcome of metaoptimization process over regressors and metaparameters obtained using the AutoML framework [auto-sklearn toolkit (55)]. We found that a random forest classifier with suitably tuned metaparameters (the same for all tests) consistently gave cross-validation testing results on par with the best learning models chosen by AutoML (and was also the model most frequently returned by this framework). We also used OLS as a staple backup, which has the advantage of involving a small number of internal parameters with clear interpretation. The respective loss functions are described in *SI Appendix*, section I.

Training and Testing Sets. The training and testing sets were constructed in the following manner. In each run, the set of nodes was split into two subsets, where each port entered either the training or testing set with equal probability. The cut was consistent over all annual snapshots of the network considered in a given run. Only links between pairs of nodes belonging to the same set were considered, thus approximately 1/4 of the links entered the training set and 1/4 entered the testing set, and the rest were discarded from the run. In this way, both the training and the testing sets provided representative samples for the whole network, while minimizing possible autocorrelation between the two sets. Node features were computed based on parameters of the entire network. Experiments were performed over 24 runs. The metaparameters of the regressors (specifically, random forest) were chosen so that the predictive quality of the models on the training and testing sets was comparable. For the classifiers and regressors with fewer internal parameters (OLS, logit), we also used the same approach with training and testing sets to avoid the effect of tuning model parameters to single data points (i.e., large ports such as Singapore which control a large part of the world's traffic).

In the case of forecasting link openings, we have restricted our attention to the links absent from the network in year Y for potential addition in Y+1. In the case of flow forecasting, flows in year Y+1 were predicted on those edges which were present in year Y so as to avoid introducing bias.

We do not provide an explicit statistical significance analysis of the predictive quality of our models. We note that all available historical data (31 year-to-year experiments) point to the consistent superiority of models relying on common neighbors with respect to all gravity models. These experiments are ordered by time. While this is a weaker property than independent trials, most statistical significance formulas developed for independent trials carry over rigorously to this setting, including those for binomial proportion confidence intervals. For example, if we were to abstractly assume that there exists a probability value $\alpha \in [0\%, 100\%]$ such that the common neighbors model performs better than port gravity at forecasting link openings with probability α in any year-to-year experiment, we obtain, from our study, a Clopper–Pearson interval of $\alpha \in [88.7\%, 100\%]$, for a confidence level of 95% ($P = 0.05$).

The performed averaging over multiple (24) runs minimizes noise due to randomness in choice of training sets and learning algorithm convergence, thus smoothing the obtained quantitative results while not affecting qualitative results.

The Kobe Shock: Construction of Subnetwork. The subnetwork H used to isolate signal after the Kobe earthquake was constructed in the following way. First, we identified ports for which the trade with Kobe accounted for at least 10% of their total throughput in 1994, and not less than 10,000 DWT in moves recorded in the database in that year. There were 16 such ports. Then we added their direct network neighbors to the node set of the subnetwork along with the links relaying them with the identified ports, to form the link set of the subnetwork. Thus, the subnetwork H is almost bipartite and, technically, is a subgraph of the entire network but not an induced subgraph. We retained the aggregated DWT value transported over the links of the subnetwork H in each year. Considering a union over all of the years 1994–1996, there were 614 active ports in this subnetwork. The analysis of detailed effects of the Kobe shock is made under moderate assumptions on normality and weak correlation of year-to-year DWT changes for port subsets within subnetwork H, which are justified by patterns observed in the dataset (*SI Appendix*, section P).

Data Availability Statement. By permission of Lloyd's List Intelligence, the anonymized network snapshots used in this study will be made available to any researcher, for scientific purposes, upon request made to the author (https://github.com/zuzannastamirowska/maritime). Access to the complete data source is also possible on a commercial basis. The data are also available in paper form, for example, at the British Library.

ACKNOWLEDGMENTS. I gratefully acknowledge the support of Lloyd's List Intelligence, related to sharing data for this study. The research leading to these results has received funding from the ERC under the European Union's Seventh Framework Program (FP/2007-2013)/ERC Grant Agreement 313847 "World Seastems" and was conducted, in part, at the Complex Systems Institute of Paris. I am obliged, for valuable discussions and suggestions related to the manuscript, to Arnaud Banos, Ray Rivers, and Laurent Viennot. I gratefully acknowledge the help of David Fourquet, who provided samples of source code for symbolic regression in real-world networks which was refactored and adapted for the preselection step of model features. Finally, I thank Adrian Kosowski for further source code samples and for a statistical review of the paper.

1. J. Hoffman, S. Sirimanne, Eds., Executive Summary (United Nations Conference on Trade and Development, 2017).
2. J. P. Rodrigue, C. Comtois, B. Slack, *The Geography of Transport Systems* (Routledge, 2016).
3. B. Ki-Moon, Maritime transport is 'backbone of global trade and the global economy', says Secretary-General in message for International Day. https://www.un.org/press/en/2016/sgsm18129.doc.htm. Accessed 1 March 2019.
4. P. Kaluza, A. Kölzsch, M. T. Gastner, B. Blasius, The complex network of global cargo ship movements. *J. R. Soc. Interface* **7**, 1093–1103 (2010).
5. G. Concepcion *et al.*, Resolving natural ranges and marine invasions in a globally distributed octocoral (genus Carijoa). *Mar. Ecol. Prog. Ser.* **401**, 113–127 (2010).
6. F. Creutzig *et al.*, Transport: A roadblock to climate change mitigation? *Science* **350**, 911–912 (2015).
7. Z. Kosowska-Stamirowska, C. Ducruet, N. Rai, Evolving structure of the maritime trade network: Evidence from the Lloyd's shipping index (1890–2000). *J. Shipp. Trade* **1**, 10 (2016).
8. L. Tavasszy, M. Minderhoud, J. F. Perrin, T. Notteboom, A strategic network choice model for global container flows: Specification, estimation and application. *J. Transport Geogr.* **19**, 1163–1172 (2011).
9. F. G. Laxe, M. J. F. Seoane, C. P. Montes, Maritime degree, centrality and vulnerability: Port hierarchies and emerging areas in containerized transport (2008–2010). *J. Transport Geogr.* **24**, 33–44 (2012).
10. D. Doshi, B. Malhotra, S. Bressan, J. S. L. Lam, Mining maritime schedules for analysing global shipping networks. *Int. J. Bus. Intell. Data Min.* **7**, 186–202 (2012).
11. Y. Hu, D. Zhu, Empirical analysis of the worldwide maritime transportation network. *Physica A* **388**, 2061–2071 (2009).
12. C. Ducruet, T. Nottebom, The worldwide maritime network of container shipping: Spatial structure and regional dynamics. *Global Network* **12**, 395–423 (2012).
13. C. Ducruet, Multilayer dynamics of complex spatial networks: The case of global maritime flows (1977–2008). *J. Transport Geogr.* **60**, 47–58 (2017).
14. A. Baird, Redefining maritime transport infrastructure. *Proc. Inst. Civ. Eng.* **163**, 29–33 (2010).
15. T. Verma, F. Russmann, N. A. M. Araújo, J. Nagler, H. J. Herrmann, Emergence of core-peripheries in networks. *Nat. Commun.* **7**, 10441 (2016).
16. E. Strano, V. Nicosia, V. Latora, S. Porta, M. Barthélemy, Elementary processes governing the evolution of road networks. *Sci. Rep.* **2**, 296 (2012).
17. V. Arora, M. Ventresca, Action-based modeling of complex networks. *Sci. Rep.* **7**, 6673 (2017).

18. A. L. Barabási, R. Albert, Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).
19. M. Fugazza, J. Hoffmann, Liner shipping connectivity as determinant of trade. *J. Shipp. Trade* **2**, 1 (2017).
20. A. V. Deardorff, Determinants of bilateral trade: Does gravity work in a neoclassical world? https://www.nber.org/papers/w5377. Accessed 10 December 2019.
21. T. Mayer, S. Zignago, Notes on CEPII's distances measures: The geodist database. http://www.cepii.fr/CEPII/en/bdd_modele/presentation.asp?id=6. Accessed 5 May 2020.
22. United Nationals Conference on Trade and Development, "Structure, ownership and registration of the world fleet" in *United Nations Conference on Trade and Development,* J. Hoffman, S. Sirimanne, Eds. (United Nations, 2017), pp. 28–31.
23. "Maritime transport connectivity" in *United Nations Conference on Trade and Development*, J. Hoffman, S. Sirimanne, Eds. (United Nations, 2017), pp. 99–111.
24. M. Newman, *Networks: An Introduction* (Oxford University Press, New York, NY, 2010).
25. R. F. Betzel, D. S. Bassett, Generative models for network neuroscience: Prospects and promise. *J. R. Soc. Interface* **14**, 20170623 (2017).
26. M. Bell *et al.*, Network growth models: A behavioural basis for attachment proportional to fitness. *Sci. Rep.* **7**, 42431 (2017).
27. K. Head, T. Mayer, "Gravity equations: Workhorse, toolkit, and cookbook" in *Handbook of International Economics* (Elsevier, 2014), vol. 4, pp. 131–195.
28. J. Q. Stewart, A measure of the influence of a population at a distance. *Sociometry* **5**, 63–71 (1942).
29. T. Menezes, C. Roth, Symbolic regression of generative network models. *Sci. Rep.* **4**, 6284 (2014).
30. T. Zhou, L. Lü, Y. C. Zhang, Predicting missing links via local information. *Euro. Phys. J. B* **71**, 623–630 (2009).
31. V. Martínez, F. Berzal, J. C. Cubero, A survey of link prediction in complex networks. *ACM Comput. Surv.* **49**, 69 (2016).
32. W. Isard, Location theory and trade theory: Short-run analysis. *Q. J. Econ.* **68**, 305–320 (1954).
33. M. Rosvall, A. V. Esquivel, A. Lancichinetti, J. D. West, R. Lambiotte, Memory in network flows and its effects on spreading dynamics and community detection. *Nat. Commun.* **5**, 4630 (2014).
34. D. Liben-Nowell, J. Kleinberg, The link-prediction problem for social networks. *J. Am. Soc. Inf. Sci. Technol.* **58**, 1019–1031 (2007).
35. P. Sarkar, D. Chakrabarti, P. J. Bickel, "The consistency of common neighbors for link prediction in stochastic blockmodels" in *Advances in Neural Information Processing Systems 28*, C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, R Garnett, Eds. (Curran Associates, Inc., 2015), pp. 3016–3024.
36. G. Horwich, Economic lessons of the Kobe earthquake. *Econ. Dev. Cultural Change* **48**, 521–542 (2000).
37. F. Saracco *et al.*, Inferring monopartite projections of bipartite networks: An entropy-based approach. *New J. Phys.* **19**, 053022 (2017).
38. L. Lü, T. Zhou, Link prediction in complex networks: A survey. *Physica A* **390**, 1150–1170 (2011).
39. D. A. Aschauer, Is public expenditure productive? *J. Monetary Econ.* **23**, 177–200 (1989).
40. O. Gianmarco, D. Puga, Agglomeration in the global economy: A survey of the new economic geography. *World Econ.* **21**, 707–731 (1998).
41. J. Yang, X. D. Zhang, Predicting missing links in complex networks based on common neighbors and distance. *Sci. Rep.* **6**, 38208 (2016).
42. C. Chiu, J. Zhan, Deep learning for link prediction in dynamic networks using weak estimators. *IEEE Access* **6**, 35937–35945 (2018).
43. A. Ortega, P. Frossard, J. Kovacevic, J. M. F. Moura, P. Vandergheynst, Graph signal processing: Overview, challenges, and applications. *Proc. IEEE* **106**, 808–828 (2018).
44. B. Perozzi, R. Al-Rfou, S. Skiena, "Deepwalk: Online learning of social representations" in *The 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '14*, S. A. Macskassy, C. Perlich, J. Leskovec, W. Wang, R. Ghani, Eds., (Association for Computing Machinery, New York, NY, 2014), pp. 701–710.
45. A. Grover, J. Leskovec, "Node2vec: Scalable feature learning for networks" in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16*, B. Krishnapuram *et al.*, Eds. (Association for Computing Machinery, New York, NY, 2016), pp. 855–864.
46. H. Shao, L. Wang, Y. Ji, Link prediction algorithms for social networks based on machine learning and HARP. *IEEE Access* **7**, 122722–122729 (2019).
47. M. Nickel, D. Kiela, "Poincaré embeddings for learning hierarchical representations" in *Advances in Neural Information Processing Systems 30*, I. Guyon *et al.*, Eds. (Curran Associates, Inc., 2017), pp. 6338–6347.
48. G. Alanis-Lobato, P. Mier, M. A. Andrade-Navarro, Efficient embedding of complex networks to hyperbolic space via their Laplacian. *Sci. Rep.* **6**, 30108 (2016).
49. C. A. Hidalgo, R. Hausmann, The building blocks of economic complexity. *Proc. Natl. Acad. Sci. U.S. A.* **106**, 10570–10575 (2009).
50. T. Evans, R. Rivers, C. Knappett, "Physical and relational networks in the aegean bronze age" in *European Conference of Complex Systems*, J. Jost, F. Reed-Tsochas, P. Schuster, Eds. (Complex Systems Society, 2006), p. 81.
51. A. Kireeva, C. Digges, Russian executives weigh in on arctic development plans. *The Maritime Executive* (2018). https://www.maritime-executive.com/editorials/russian-executives-weigh-in-on-arctic-development-plans. Accessed 1 March 2019.
52. A. W. Miller, G. M. Ruiz, Arctic shipping and marine invaders. *Nat. Clim. Change* **4**, 413–416 (2014).
53. G. K. Hovelsrud, B. Poppel, B. van Oort, J. D. Reist, Arctic societies, cultures, and peoples in a changing cryosphere. *AMBIO* **40**, 100–110 (2011).
54. S. Soto, Human migration and infectious diseases. *Clin. Microbiol. Infect.* **15**, 26–28 (2009).
55. M. Feurer *et al.*, "Efficient and robust automated machine learning" in *Advances in Neural Information Processing Systems 28*, C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, R. Garnett, Eds. (Curran Associates, Inc., 2015), pp. 2962–2970.