The road most travelled: The impact of urban road infrastructure on supply chain network vulnerability⋆.

Nadia M. Viljoen · Johan W. Joubert

Abstract Making a supply chain more resilient and making it more efficient are often diametrically opposed objectives. Managers have to make informed trade-offs when designing their supply chain networks. There are many methods available to quantify and optimise efficiency. Unfortunately the same cannot be said for vulnerability and resilience. We propose a method to quantify the impact that a supply chain's dependence on the underlying transport infrastructure has on its vulnerability. The dependence relationship is modelled using multi-layered complex network theory. We develop two metrics relating to the unique collection of shortest path sets namely redundancy and overlap. To test the relationships between these metrics and supply chain vulnerability we simulate progressive random link disruption of the urban road network and assess the impact this has on Fully Connected, Single Hub and Double Hub network archetypes. The results show that redundancy and overlap of the collection of shortest paths are significantly related to supply chain resilience, however under a purely random disturbance regime they hold no predictive power. This paper builds a foundation for a new field of inquiry into supply chain vulnerability by presenting a flexible mathematical formulation of the multilayered network and defining and testing two novel metrics that could be incorporated into supply chain network design decisions.

Keywords supply chain vulnerability · urban road network · multilayer complex network · shortest paths · link betweenness

1 Introduction

Ask any operations manager on the warehouse floor, in the distribution centre's cross-dock or at the receiving bay of the retail store and they will confirm that disruptions in supply chain activities are far less exceptional and far more costly than commonly believed. It is not only the natural disasters and terrorist attacks that cost supply chains billions in turnover, the less extraordinary realities of traffic congestion, power outages, internet network failure and even industrial action can be just as harmful. The impact of these disruptions could range from missed

Nadia M. Viljoen
Centre for Transport Development, University of Pretoria, South Africa
Tel.: +27 12 420 2205
Fax.: + 27 12 362 5103
E-mail: viljoenn@gmail.com

Johan W. Joubert
Centre for Transport Development, University of Pretoria, South Africa
Tel.: +27 12 420 2843
Fax.: + 27 12 362 5103
E-mail: johan.joubert@up.ac.za

delivery time windows to stock-outs or unplanned overtime. Although these impacts are of a smaller scale, over time they could lead to death by a thousand cuts for the supply chain.

Supply chain risk is "*anything* that (disrupts or impedes) the information, material or product flows from original suppliers to the final delivery of the final product to the ultimate end-user" (emphasis ours)(Peck, 2006). To define supply chain vulnerability, we start with the notion that it is the susceptibility of a supply chain to such disruptions (Wagner and Bode, 2006). Four pertinent trends are driving increased supply chain vulnerability: an increase in natural and man-made disasters, a rise in supply chain complexity, heightened financial pressure on supply chain operations and fiercer global competition (Wagner and Neshat, 2010). Therefore both the internal characteristics of a supply chain (e.g. network design, supplier relations, quality control procedures) and the external circumstances (e.g. political climate, condition of infrastructure and the weather) make a supply chain vulnerable. In this paper we define supply chain vulnerability as "the degree to which supply chain configuration and external circumstances make a supply chain susceptible to the damaging effects of a disruption."

More than 90% of CEO's surveyed by the World Economic Forum confirmed that Supply Chain Risk Management (SCRM) has been elevated to a top-level priority (Chacon et al., 2012). These days executives are not only tasked to mitigate the negative consequences of risk, but to proactively build resilience into their supply chains. Building resilience requires that one prioritises the basic function of customer service (efficacy) over the conservation of resources (efficiency) in the ongoing quest for financial sustainability (Bhatia et al., 2013). This demands a precarious balancing act. "How vulnerable is our current network design in City A? By what margin could we lessen this vulnerability by changing our facility locations? What impact would such a change have on our bottom line in the short, medium and long term?" Answering such questions requires a deep and quantified understanding of the supply chain's vulnerability — herein lies the SCRM research gap.

While there is an abundance of qualitative definitions, frameworks and taxonomies on the topic of supply chain risk, very few quantitative methodologies and metrics have been developed (Chacon et al., 2012; Heckmann et al., 2015). Furthermore, the majority of literature fixates on the *what* and the *when* of a disruption and how to mitigate the resulting consequences (Heckmann et al., 2015) instead of delving deeper to understand what drives the vulnerability in the first place (Rao and Goldsby, 2009). This paper proposes a quantitative approach to assessing the dependence of a supply chain on the underlying transport infrastructure and how that dependence makes it more, or less vulnerable to disruptions.

Quantifying the dependence of a supply chain on the underlying transport infrastructure it utilises is not trivial. Infrastructure is physical, measurable and definitive. Supply chains are a conglomeration of interconnected business relationships that cannot be easily defined or measured and are subject to change. Initially, researchers modelled supply chains as "sequential dyadic relationships", but this simplistic view has recently been replaced by a complex network approach that better captures the reality of a supply chain's relations (Bellamy and Basole, 2013; Choi et al., 2001; Choi and Wu, 2009; Hearnshaw and Wilson, 2013; Pathak et al., 2007; Thadakamalla et al., 2004). Although the cited authors use different terminology, the discerning quality of a complex network approach is that the whole

is greater than the sum of its parts. In other words, one cannot fully distill the properties of a supply chain by merely dividing up and studying its functions in isolation. Rather, one should model how these functions interact as they do in reality to appreciate emergent trends and behaviours. However, a supply chain also interacts with its environment, for example the transport infrastructure it uses when freight moves between facilities. Although the application of complex network theory in transportation systems has been quite prolific (see Section 2), modelling the interaction between a supply chain and the transport infrastructure is rather novel. We borrow concepts from the erupting field of multilayered complex networks to capture this interaction.

The scope of our study focusses on supply chain networks within urban areas where supply chain facilities are usually concentrated (Holl and Mariotti, 2017). These areas predominantly exhibit the grid-like structure reminiscent of urban centres. These networks often face volatile demand with short notice periods, tight transport lead times, fierce competition and fickle customers. Urban centres also frequently experience traffic congestion and disruptions due to roadworks or accidents, infrastructure or equipment failure. The reality and impact of the supply chain's daily dependence on urban transport infrastructure warrants a targeted investigation.

Within a metropolitan area there is a range of supply chain facilities including factories, warehouses, distribution centres, and retail outlets. These facilities seldom belong to the same organisations, but rather to different supply chain partners. *How* freight moves between these facilities is the art and science of supply chain network design. The inter-firm relations and intra-firm business rules prescribe, for example, whether retail stores can reallocate stock amongst themselves or whether stock first has to be channeled through some central distribution centre. The supply chain network design that emerges from these relations and business rules we capture as the *logical* layer of our multilayered network. In practice there are common archetypes of such network designs of which we study three in this paper.

The logical layer of the the multilayered network is placed on the *physical* layer, which represents the urban road network. This physical layer is modelled as a regular, directed grid network — a typical approximation of the general topology of cities around the world (Ortigosa and Menendez, 2014). Through a set of progressive disruptions we then investigate how the different network design archetypes are impacted as the physical layer degrades.

This work is a novel application of complex network concepts to the study of supply chain vulnerability in an urban setting. The modelling approach and metrics presented here build the foundation for a new field of inquiry.

This paper is structured as follows: Section 2 presents the definition of multilayered complex networks and especially discusses applications of single and multilayered networks to transportation and supply chain problems. Further it describes the status quo in measuring vulnerability in single and multilayered networks. Section 3 describes the formulation of our multilayered network, the definition of metrics related to redundancy, overlap and efficiency and the progressive random link removal strategy used to degrade the grid network. Section 4 describes the specific datasets used. The results are presented in Section 5 with an exploration of the predictive power of specific metrics in Section 6. The paper is concluded with a description of follow-up research questions in Section 7.

2 Related work

2.1 Multilayered network definition

Natural and engineered systems most often have multiple layers of interaction among heterogenous actors/nodes. Take for example the relationships between colleagues all working together in a department. Some may be no more than work acquaintances, others may work together on projects and therefore have a closer relationship, and then there are those who have become friends outside of work and compete together in the volleyball league. This scenario describes one type of layered interaction where different layers present different types of relationships. Consider now the same group of colleagues but focus only on the project team relationships. As time goes by project teams change, giving rise to a unique set of relationships at each point in time. The second type of layered interaction can thus be temporal. To better capture the interdependent intricacies of such layered interactions, network scientists have rapidly developed multilayered methods stemming from complex network theory. The definitive shift of focus from single layer networks to the multilayer concept is evident in the slew of papers published by leading authors between 2014 and 2016. The network science community regard the study of multilayered networks as a new frontier in many areas of science and a rapidly expanding movement that will stimulate interdisciplinary research (Boccaletti et al., 2014; Tsiostas and Polyzos, 2017).

Multilayered networks behave very differently to single layer networks. Researchers have shown, both analytically and empirically, that we can't simply deduce multilayered characteristics by studying the single layer components in isolation. This is particularly true for the topics of resilience (how long a network can remain functional under attack/error) and spreading processes (how quickly information/disease/innovation spreads to network nodes) (Boccaletti et al., 2014). Therefore, in our study, apart from the fact that a single layer perspective does not adequately capture the dependency between a supply chain and the transport network it uses, insights drawn regarding the vulnerability of the supply chain may simply be untrue when considering only a single layer.

With multilayered research exploding simultaneously amongst different groups of network scientists, a diverse cloud of definitions and mathematical formulations has mushroomed around the topic. Although there has been effort to cement a common vernacular, scientists have not yet reached a point of consensus and one is faced with many different alternatives to describe and formulate multilayered networks. Scientists have also sought to define unifying frameworks that relate the many different types of multilayered networks such as multiplex, interdependent, multilevel, hypergraphs, temporal etc. The interested reader is specifically referred to Kivelä et al. (2014) for a comprehensive description and categorisation of recent studies and Boccaletti et al. (2014) for detailed structural descriptions of multilayered networks, their metrics and dynamics. Other more concise reviews include that of Lee et al. (2015) and Danziger et al. (2014) who focus on resilience and percolation, and Salehi et al. (2015) who focus on spreading processes. In this paper we adopt the formulation of Boccaletti et al. (2014) as we find it most intuitive and relevant to the application at hand.

The *multilayer network* is a pair $\mathcal{M} = (\mathcal{G}, \mathcal{C})$ where $\mathcal{G} = \{G_m; m \in \{1, 2, \ldots, M\}\}$ is a family of M individual graphs $G_m = (X_m, E_m)$ which each represent a layer of $\mathcal{M}$.

In this generic formulation, α and β refer to layers of $\mathcal{G}$ such that $\alpha, \beta \in \{1, 2, \ldots, M\}$ and $\alpha \neq \beta$. The set of nodes in a layer G_α are denoted by $X_\alpha = \left\{ x_1^\alpha, \ldots, x_{N_\alpha}^\alpha \right\}$ and $E_\alpha \subseteq X_\alpha \times X_\alpha$. The set of interconnections between nodes in G_α and G_β with $\alpha \neq \beta$ are defined by

$$\mathcal{C} = \{E_{\alpha,\beta} \subseteq X_\alpha \times X_\beta; \alpha, \beta \in \{1, \ldots, M\}, \alpha \neq \beta\} \tag{1}$$

Therefore the elements of $E_{\alpha,\beta}, \alpha \neq \beta$ are *interlayer* connections while elements of E_α are the *intralayer* connections.

In Section 3 we start from the foundation of this simple generic formulation to define our multilayered network and extend the concepts to define the metrics tracked throughout our experiments.

2.2 Complex network applications in supply chains and transportation

Complex network applications have bred useful and relevant fields of inquiry in an impressive array of different fields: neural networks, social networks, ecological systems, information technology, citation networks, power grids, water supply, internet networks, financial systems and, of course, transport systems and supply chains. However, there is still a bridge to be built to link the work done in multilayered transport systems to single layered supply chain models. This section describes the most relevant studies and status quo of applications in these two fields.

2.2.1 Supply chain applications

Hearnshaw and Wilson (2013) have set a standard in applying complex network theory to supply chains. Taking great care to capture the different supply chain facets in their modelling approach, they propose that an efficient supply chain adheres to a hub-and-spoke network design. Such a network has a heterogenous degree distribution and is labelled "scale-free". Their proposition states that efficient supply chains have short average path lengths between nodes and reasonably well-connected clusters as indicated by high clustering coefficients. This work exclusively considers an efficiency-based view and does not consider the vulnerability of the supply chain. In fact, scale-free networks are known to be highly vulnerable to targeted attacks (O'Kelly, 2015).

The first work to tie supply chain vulnerability concepts to network metrics is by Thadakamalla et al. (2004) and actually precedes the thorough treatment of Hearnshaw and Wilson (2013). The authors set out to identify what supply chain network design would be most resilient by remaining functional after repeated errors or attacks. According to them a resilient supply chain

> "maintains connectivity between the majority of its nodes; does not suffer a significant increase in the average shortest path length between nodes; has well-defined clusters that offer many alternative shortest paths and can automatically re-wire itself after disruption to establish functionality" (Thadakamalla et al., 2004).

They empirically investigate three complex network alternatives namely random graphs, small-world models and scale-free networks to ascertain which resilience properties are inherent in which designs. Their results show that no single network design satisfies all components but that a hybrid design is probably most resilient. Although closer to our goal of studying supply chain vulnerability, this study still presents the supply chain network as a single layer of relational connections.

We have not found any complex network applications that study the dependence between supply chain network design and spatial networks such as transportation, power grid or communication networks. The insights from such studies would add great value to strategic decision making in supply chain network design.

2.2.2 Transportation applications

Single layer studies on transportation systems are extensive covering air transport (Lordan et al., 2016, 2015; Lordan, 2014; Lordan et al., 2014b; Sun et al., 2017; Zanin et al., 2009; Zanin and Lillo, 2013), maritime networks (Bartholdi and Jarumaneeroj, 2014; Ducruet et al., 2010a,b; Ducruet and Notteboom, 2012; Ducruet, 2016; González Laxe et al., 2012; Hu and Zhu, 2009; Fraser et al., 2014; Mohamed-Chérif and Ducruet, 2016; Pais Montes et al., 2012; Tsiostas and Polyzos, 2015; Viljoen and Joubert, 2016), train, bus, tram and subway systems (Criado et al., 2007; Kurant and Thiran, 2006a; Mouronte and Benito, 2012; Ouyang et al., 2014; Sen et al., 2003), and the structural properties and vulnerability of road networks (Crucitti et al., 2006; Duan and Lu, 2014; Gudmundsson and Mohajeri, 2013; Háznagy et al., 2015; Jiang, 2007; Jiang and Claramunt, 2004; Porta et al., 2012; Reggiani et al., 2011; Rupi et al., 2015; Strano et al., 2009; Zadeh and Rajabi, 2013).

The vulnerability of passenger transport systems, particularly within the urban environment, has also been a topic of ardent study amongst network scientists ((Mattsson and Jenelius, 2015) and references therein) and (Bóta et al., 2017; Taylor and Susilawati, 2012; Rodríguez-Núñez and García-Palomares, 2014; Rupi et al., 2015). However, we found that very few studies explicitly consider the impact of urban transport vulnerability on the movement of freight using a complex network approach. The underrepresentation of freight systems in transport planning in general is a recognised fact (Joubert and Axhausen, 2013) and the situation is no different in network vulnerability studies. Darayi et al. (2017) weighted a multi-modal freight transport network with commodity flows calculated by a typical four-step model. The purpose of their study was to evaluate the multi-industry impacts of transport network disruptions. While insightful, their study did not focus exclusively on the urban reality. It also used gravity modelling as a proxy for freight movement instead of actual freight activity. Joubert and Axhausen (2013) were the first to explicitly address freight movement using complex networks. The authors constructed a network from real-life commercial vehicle activity chains, illustrating how centrality measures can be used to identify key logistics facilities in Gauteng, South Africa. However, their work also only considers the single layer of relationships between logistics facilities.

Multilayered network studies in transportation have been more recent and mostly present a network $\mathcal{M}$ where the individual layers (G_α) are different transport modes. The global air transport and maritime networks offer rich examples

of multilayered networks where the layers present different airline or airline alliance networks (Cardillo et al., 2015; Tsiotas and Polyzos, 2015), different maritime cargo types (Ducruet, 2013; Kaluza et al., 2010) and even multimodal layers (Parshani et al., 2011). Of particular interest to this study are the metropolitan transport applications motivated by the importance of multimodal transport as the mainstay of urban mobility. A context familiar to many is the multimodal interaction between the street network and subway systems of New York and London. Strano et al. (2015) use a multilayered approach to study the impact that subway speeds and design have on congestion and balanced accessibility. Broadening the scope to all modes within the British public transport system, Gallotti and Barthelemy (2014, 2015) quantify the impact of intermodal waiting times on the network's efficiency and compare the anatomy of multimodal trips across cities. Adding socioeconomic considerations, Lotero et al. (2016) construct six multiplex representations of urban mobility in Bogotá and Medellín for different socioeconomic stratas, proving that such factors have an extraordinary impact on the structure of the networks. Another study provides equations to predict the onset of congestion in multimodal networks and shows that the efficiency offered by multimodal systems can actually unbalance transportation loads and result in unexpected congestion (Solé-Ribalta et al., 2016).

In all these models, each layer G_α represents a physical layer along which goods or people can be transported. This is in contrast to the multilayered network we will develop. In our network the supply chain layer represents a *logical* layer defined by the supply chain network design. These links represent inter-firm relations or intra-firm business rules. Freight cannot travel along these ethereal links but instead travels along the urban road network which represents the *physical* layer of the network. Notably two other multilayer transportation studies also place a *logical* layer on a *physical* layer. Kurant and Thiran (2006b) investigate the centrality metrics of public transport systems in three European geographies by defining $\mathcal{M}$ where $G_\alpha = (X_\alpha, E_\alpha)$ as the transport infrastructure and $G_\beta = (X_\beta, E_\beta)$ as transport intensity, extracted from timetables with $X_\alpha \equiv X_\beta$. Zhuo et al. (2011) study congestion vulnerability on a selection of experimental networks defining $\mathcal{M}$ with $G_\alpha = (X_\alpha, E_\alpha)$ as the transport infrastructure and $G_\beta = (X_\beta, E_\beta)$ as traffic profiles generated by their algorithm.

Our study thus draws insights from a wealth of multilayered applications in transportation systems, keeping in mind that our context and formulation of $\mathcal{M}$ is somewhat unique. Because it is unique, we must evaluate the applicability of the typical approaches to multilayered vulnerability.

2.3 Vulnerability of multilayered networks

Before diving into a discussion of vulnerability studies, let's define three prevalent network characteristics used to measure vulnerability. **Robustness** always refers to the overall connectedness of a network post disturbance. In other words, can messages (or people or freight) still flow from one node to any other (or most other) nodes? **Efficiency** determines how quickly a message can travel from one node to another and is generally expressed in terms of the average shortest path of the network (Crucitti et al., 2004; Thadakamalla et al., 2004). **Flexibility** refers to the network's inherent ability to find alternative paths if the shortest paths are

destroyed. Measuring flexibility is slightly more tricky, but most typically authors have used the clustering coefficient, arguing that well-connected neighbourhoods offer more rerouting options (Thadakamalla et al., 2004; Viljoen and Joubert, 2016).

In multilayered networks, robustness takes centre stage with percolation theory and cascading failure being the preferred methods used to assess it. The following two sections describe these methods and evaluate their applicability to this study.

2.3.1 Percolation theory and the critical point

The connectedness of a network is defined by its *connected components* which are subsets of nodes such that a path exists between all node pairs in each component. The Largest Connected Component (LCC) is the connected component that contains the largest fraction of the entire network's nodes. The LCC is also referred to as the Giant Connected Component. Essentially when the fraction of the network in the LCC is close to one, messages from any node have a high probability of reaching any other node of the network. When this fraction reduces, nodes become disconnected and can no longer be reached by other network nodes. If the fraction drops below some threshold value, the network would become non-functional due to increasing disconnectedness. The fraction of nodes in the LCC of a network (or the size of the LCC) has thus become a convenient way of measuring the robustness of a network (Danziger et al., 2014). The threshold value below which the LCC would become disconnected is, however, instance-specific and depends greatly on the field of application (social systems, transportation, neural networks etc.).

From the discipline of percolation theory we learn that in many cases the fraction of nodes that needs to be removed before the LCC becomes disconnected can be analytically determined. This fraction is denoted by $1 - p_c$ where p_c is referred to as the *critical point*. In other words, when the fraction of nodes removed from a network is $1 - p$ and $p > p_c$, then the size of the LCC remains greater than zero. As $p \to p_c$ the size of the LCC also approaches zero until finally the LCC becomes completely disconnected when $p \leq p_c$ (Danziger et al., 2014; Stauffer and Aharony, 1991).

It has been shown, in single layer networks, that $p_c \sim 0$ for scale-free networks, meaning that in cases of random failures, nearly *all* of the nodes have to be removed before the LCC becomes disconnected (Cohen et al., 2001). Thus scale-free networks are highly robust to random failures. Random (Erdös-Rényi) networks, on the other hand, are highly vulnerable with p_c inversely related to average of the network degree (Cohen et al., 2000).

A significant caveat is that percolation (and the determination of p_c) refers to random failures and not targeted attacks. Cohen et al. (2001) showed that under targeted attack scale-free networks disintegrate long before the calculated p_c is reached. In fact, Pastoras-Satorras and Vespignani (2001) prove that a p_c that takes into account targeted attack cannot be analytically determined.

Despite the fact that p_c is not always a relevant indicator, measuring the size of the LCC remains the most popular robustness metric (Lordan et al., 2014a, 2016; Thadakamalla et al., 2004; Viljoen and Joubert, 2016). However, the impact on the network skeleton — a concept emerging recently from the study of core-

periphery structure (Shekhtman et al., 2014) — could also be a valid interpretation of robustness.

2.3.2 Cascading failures on multilayered networks

Cascading failures is the foremost phenomenon studied in multilayered network vulnerability. If a network $\mathcal{M}$ is interdependent then the intralayer connections $E_{\alpha,\beta}$ are actually called *dependency* links. By removing any node $x_i^\alpha \in E_\alpha$ all nodes in other layers that were connected to x_i^α via *dependency* links are also removed. This removal process "cascades" back and forth between layers until all dependent nodes are removed. The calculation of the *Mutually Connected Giant Component (MCGC)* and the critical percolation properties of cascading failures on a variety of interdependent networks have been studied extensively ((Boccaletti et al., 2014; Kivelä et al., 2014), and references therein).

Studies that use the concepts of cascading failures and percolation to investigate vulnerability make three *a priori* assumptions:

1. $\mathcal{M}$ is interdependent resulting in cascading phenomena;
2. Nodes are removed during failures or attacks; and
3. A network is considered robust when the size of its MCGC is larger than some fraction.

These *a priori* assumptions are not present in our representation of the problem. Firstly, there is no *inter*dependence, only one-way dependence of the supply chain network on the road network. Removing one or more supply chain facilities has no impact on the integrity of the road network and therefore the cascading phenomenon does not occur. Secondly, to model typical urban road disruptions like traffic jams, road works or accidents, one would remove one or more road links, not necessarily intersections (nodes). Removing a node would remove links that are not affected by the disruption. The mechanism of failure or attack is thus focussed on link removal, not node removal, drawing into question the established equations for evaluating percolation. Lastly, in this study we consider a supply chain functional when *all* nodes remain connected to *all* other nodes. The context of our application and the highly conservative definition of supply chain functionality requires us to broaden our view of multilayered vulnerability and find a different way to quantify it.

3 Vulnerability of the multilayered urban supply chain network

In this section we present the mathematical formulation of our multilayered network. We then zoom in on the characteristics of the shortest path sets to define our metrics for efficiency, overlap and redundancy. We conclude the section by describing the progressive random link removal strategy that will be used to assess the vulnerability of $\mathcal{M}$.

3.1 Defining the multilayered network

We adapted the generic formulation slightly for better readability. The indices that name the different network layers (α and β in the generic formulation) are

superscripts in the customised formulation instead of subscripts. This prevents confusion with node indices which are indicated as subscripts.

Let $\mathcal{M} = (\mathcal{G}, \mathcal{C})$ be the multilayered network where $\mathcal{G} = (G^{1K}, G^2)$. G^{1K} represents the *logical* layer with K denoting the type of network design archetype and G^2 represents the road network.

Similar to the representation in Ortigosa and Menendez (2014), the directed, unweighted grid network has m rows and n columns, and the nodes are numbered sequentially from 1 to $m \times n$ as shown in Figure 1a. In this paper we use an example network with 100 nodes organised in a 10×10 configuration (Figure 1b). Nodes are connected with two directed, opposite links, $\leftrightarrows$, instead of one undirected link, $\leftrightarrow$. The assumption is that when a road segment in one direction fails, the associated lane in the opposite direction is not necessarily affected as well. This occurs, for example, when the two opposing lanes of a road are separated by a median strip. Therefore we have $G^2 = (X^2, E^2)$ where:

$$N^2 = m \times n = 100 \tag{2}$$

$$X^2 = \{x_1^2, x_2^2, \ldots, x_{N^2}^2\} \tag{3}$$

$$E^2 = \{e_{st}^2\} \qquad\qquad \forall s, t \in \{1, 2, \ldots, N^2\} | s \neq t \tag{4}$$

where

$$e_{st}^2 = \begin{cases} 1 & \text{if } x_s^2 \text{ was connected to } x_t^2 \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

(a) Generic.

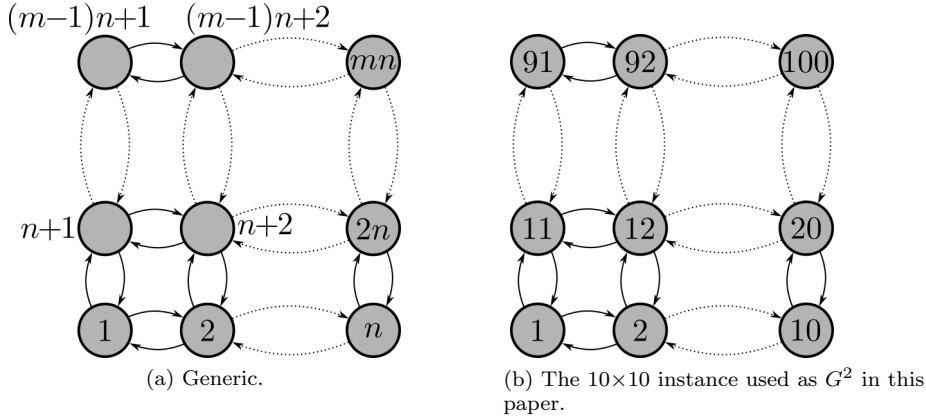(b) The 10×10 instance used as G^2 in this paper.

Fig. 1: Unweighted, directed grid layouts to approximate the urban road infrastructure.

Our three network design archetypes are called the Fully Connected (G^{1F}), Single Hub (G^{1S}) and Double Hub (G^{1D}) archetypes. The nodes represent logistics facilities and the links represent a relationship between two facilities such that these facilities can ship freight directly to each other. Links between networks

are bi-directional, as in G^2. We therefore have $G^{1K} = (X^{1K}, E^{1K}), K \in \{F, S, D\}$ where:

$$N^{1K} = 12 \qquad\qquad \forall K \in \{F, S, D\} \quad (6)$$

$$X^{1K} = \{x_1^{1K}, x_2^{1K}, \ldots, x_{N^{1K}}^{1K}\} \qquad\qquad \forall K \in \{F, S, D\} \quad (7)$$

$$E^{1K} = \{e_{ij}^{1K}\} \,\forall i, j \in \{1, 2, \ldots, N^{1K}\} | i \neq j, \qquad\qquad \forall K \in \{F, S, D\} \quad (8)$$

where

$$e_{ij}^{1K} = \begin{cases} 1 & \text{if } x_i^{1K} \text{ was connected to } x_j^{1K} \quad \forall K \in \{F, S, D\} \\ 0 & \text{otherwise.} \end{cases} \qquad (9)$$

In G^{1F} all the facilities can ship directly to all other facilities. G^{1S} and G^{1D} are hub networks with G^{1S} essentially being a star-network channeling all freight flows through a hub node while G^{1D} is the combination of two star-networks with a single connecting link between the hubs. Figure 2 shows conceptual examples of the three archetypes. For the sake of generality we do not presume the function of the facilities and therefore assume that connection between every node-pair in the respective networks $K \in \{F, S, D\}$ is important.

(a) Fully Connected archetype. (b) Single Hub archetype. (c) Double Hub archetype.
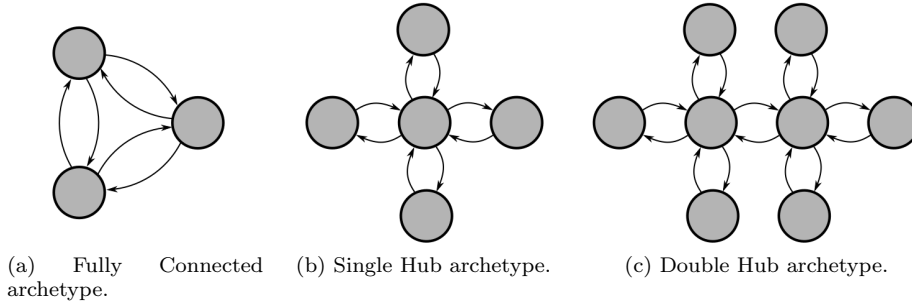
Fig. 2: Conceptual representation of the three supply chain network design archetypes.

Defining the *interlayer* connections, $E^{1K,2}$, requires that we associate the supply chain nodes, X^{1K}, with the grid nodes, X^2. Simplifying assumptions are that the logistics facilities correspond with their nearest intersections on the grid and that at most one logistics facility is associated with each intersection.

3.2 Shortest path sets

With G^2 a regular grid network and G^{1K} a predefined network archetype there is little mystery regarding the metrics of the individual layers. However, each distinct instance of $E^{1K,2}$ generates a unique set of shortest paths which characterises both the efficiency and vulnerability of $\mathcal{M}$.

Figure 3 shows the example of a Double Hub network placed on the grid with nodes 5 and 8 identified as origin and destination, respectively. If only the logical relationships of G^{1D} were to be considered, the shortest path between 5 and 8 would consist of three segments: 5–1; 1–2; 2–8 (Figure 3b). However, the physical constraints of G^2 must also be regarded. Figure 3c presents the three alternative shortest paths between 5 and 1, each of length three. Similarly there are 20 alternative shortest paths between 1 and 2, each of length six (Figure 3d) and two alternative shortest paths between 2 and 8, each of length two (Figure 3e). The length of the shortest paths between node 5 and 8 is thus 11 and there are 120 alternatives.
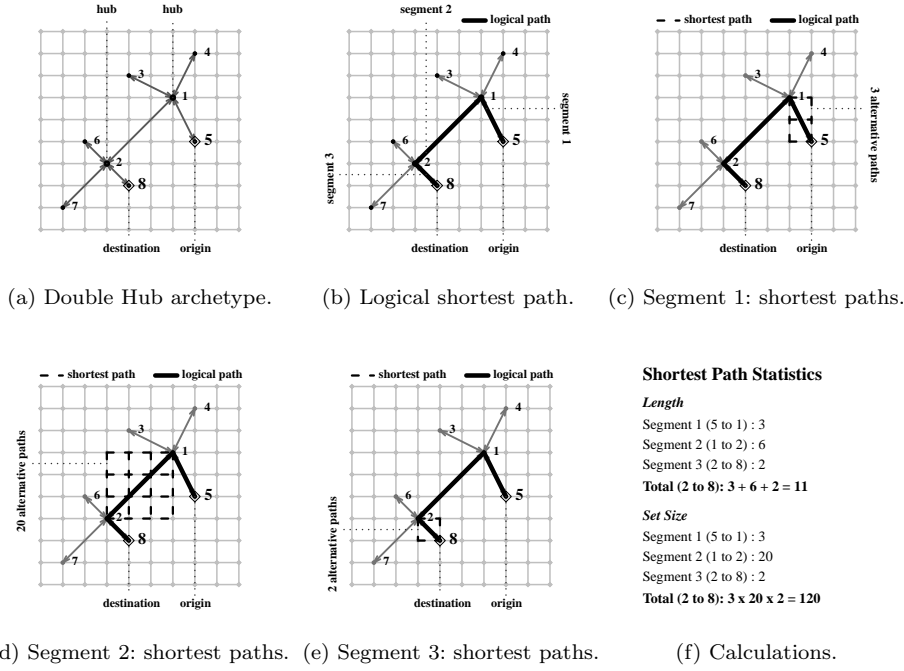

(a) Double Hub archetype.   (b) Logical shortest path.   (c) Segment 1: shortest paths.

(d) Segment 2: shortest paths. (e) Segment 3: shortest paths.   (f) Calculations.

Fig. 3: Calculating the length and number of shortest paths on $\mathcal{M}$ by adhering to both G^{1K} and G^2 constraints.

All metrics pertaining to shortest paths refer to a specific realisation of $\mathcal{M}$, therefore in the definitions that follow we drop the subscripts relating to the layers and network archetype for simplicity's sake. Generally we define:

$$\mathcal{S}_{ij} = \{SD_{ij}, SI_{ij}\} \tag{10}$$

where SD_{ij} is the subset of all shortest path sets between node-pairs (x_i^{1K}, x_j^{1K}) that are directly connected in G^{1K}:

$$SD_{ij} = \{s^1, s^2, \ldots, s^{P_{ij}}\} \qquad \forall e_{ij}^{1K} \in E^{1K} \tag{11}$$

and SI_{ij} is the subset of all shortest path sets between node-pairs that are indirectly connected in G^{1K}:

$$SI_{ij} = \{s^1, s^2, \ldots, s^{P_{ij}}\} \qquad\qquad \forall e_{ij}^{1K} \notin E^{1K} \qquad (12)$$

There were two statistics of interest namely the length of a shortest path and the number of alternative shortest paths between two nodes. Therefore:

$$L_{ij} \equiv \text{length of a shortest path between node } x_i^{1K} \text{ and } x_j^{1K} \qquad (13)$$
$$\forall i, j \in \{1, 2, \ldots, N^{1K}\}, i \neq j$$

$$P_{ij} \equiv \text{number of alternative shortest paths between node } x_i^{1K} \text{ and } x_j^{1K} \qquad (14)$$
$$\forall i, j \in \{1, 2, \ldots, N^{1K}\}, i \neq j$$

We then define the *collection* of shortest paths as

$$\mathcal{C}(\mathcal{S}_{ij}) = \bigcup_{i,j} \mathcal{S}_{ij} \qquad\qquad \forall i, j \in \{1, 2, \ldots, N^{1K}\}, i \neq j \qquad (15)$$

$\mathcal{C}(\mathcal{S}_{ij})$ is the unique fingerprint of each realisation of $\mathcal{M}$ and therein lies the key to comparing vulnerability across different network archetypes.

3.3 Efficiency, resilience and robustness

Earlier we defined **vulnerability** as *'the degree to which supply chain configuration and external circumstances make a supply chain susceptible to the damaging effects of a disruption'*.

In this specific study the *supply chain configurations* we are testing are the three network archetypes, the *external circumstance* is the urban road network and the *disruption* is the iterative removal of links from the grid.

One *damaging effect* could be that after a disruption, the length of the shortest paths (L_{ij}) between two nodes becomes longer, that is the network becomes less efficient. We measure the **efficiency** of $\mathcal{M}$ as the average shortest path length:

$$\bar{L} = \frac{\Sigma_{i,j,i \neq j} L_{ij}}{N^{1K}(N^{1K} - 1)} \text{ where } i, j \in \{1, 2, \ldots, N^{1K}\} \qquad (16)$$

where N^{1K} was the number of nodes in G^{1K}.

We thus define **resilience** as:

> The ability to maintain a level of efficiency in $\mathcal{M}$ despite the removal of links from G^2.

The other *damaging effect* could be that one or more facilities become disconnected from the rest. As already mentioned, this study holds a very conservative view of supply chain functionality. *All* nodes must be connected to *all* other nodes, directly or indirectly, for the network to be considered functional. Thus $\mathcal{S}_{ij} \neq \emptyset$ for all $x_i^{1K}, x_j^{1K} \in X^{1K}; x_i^{1K} \neq x_j^{1K}$, where $\emptyset$ denotes an empty set. We thus define **robustness** as:

> The degree to which destruction of G^2 can be endured before connection is lost between one or more node-pairs in G^{1K}.

The hypothesis is that there is a relationship between characteristics of $\mathcal{C}(\mathcal{S}_{ij})$ and the resilience and robustness of $\mathcal{M}$. By means of deductive reasoning, two characteristics of $\mathcal{C}(\mathcal{S}_{ij})$ are identified as potential indicators of the relationship.

3.3.1 Redundancy of $\mathcal{C}(\mathcal{S}_{ij})$

The more alternative shortest paths there are between a node-pair (x_i^{1K}, x_j^{1K}), the less likely it would be that a random removal of links could destroy *all* the shortest paths in $\mathcal{S}_{ij}$. (Recall that the number of alternative shortest paths is measured by P_{ij}.) As long as there remains one of the alternative shortest paths in $\mathcal{S}_{ij}$ the L_{ij} will be the same as it was before the removal of the links. Therefore it is deduced that there must be a relationship between the P_{ij} and the **resilience** of $\mathcal{M}$.

The distribution of $\sum P_{ij}$ is greatly skewed with a long tail (see Section 5.1.2). This is because P_{ij} increases exponentially with the diagonal distance between two nodes on G^2 and for indirectly connected connected node-pairs P_{ij} is the product of the set sizes of the underlying directly connected node-pairs. The median of P_{ij} is therefore a better estimator of the distribution's centrality than the mean. The number of sets in $\mathcal{C}(\mathcal{S}_{ij})$ will always be even and therefore the formula for the median is:

$$\tilde{P} = (P_{\text{mid}} + P_{\text{mid}+1})/2 \tag{17}$$

where $\boldsymbol{P}$ is the ordered set of P_{ij}, so that $P_{\text{mid}} \in \boldsymbol{P}$, and

$$\text{mid} = (N^{1K}(N^{1K} - 1) + 1)/2 - 0.5 \tag{18}$$

with N^{1K} the number of nodes in G^{1K}.

3.3.2 Overlap within $\mathcal{C}(\mathcal{S}_{ij})$

Consider a shortest path set with $P_{ij} = 20$, where all 20 shortest paths share one specific link. If that link were removed it would matter very little that there had been 20 initial paths as all would have been broken simultaneously. Imagine another scenario where two different shortest path sets share common links, the destruction of those links would affect multiple paths in both sets. While greater redundancy in $\mathcal{C}(\mathcal{S}_{ij})$ could imply better resilience, it doesn't take into consideration the overlap within $\mathcal{C}(\mathcal{S}_{ij})$.

In the few studies that also consider link removal strategies, link betweenness is used to identify overlap (Viljoen and Joubert, 2016). Link betweenness considers all the shortest paths in a network and then calculates the fraction of those paths that contain a specific link to obtain that link's betweenness score. In effect one is measuring the network's dependence on that one link in maintaining its current collection of shortest paths. A similar approach is used to calculate overlap.

Every shortest path set SI_{ij} between indirectly connected node-pairs in G^{1K} is a combinatorial product of the sets SD_{ij} of those directly connected node-pairs that comprise the logical path between the indirectly connected pair. In the example of Figure 3:

$SI_{58} = SD_{51} \times SD_{12} \times SD_{28} = 3 \times 20 \times 2 = 120$

Therefore, when calculating the overlap within $\mathcal{C}(\mathcal{S}_{ij})$ we only consider the subsets SD_{ij} as these are the building blocks of the subsets SI_{ij} and every SI_{ij} will, by definition, overlap 100% with the subsets SD_{ij} that it is comprised of.

To calculate the overlap we first define the network G^γ which is the subset of all the nodes and links in G^2 that are included in the shortest paths of all the directly connected node-pairs in G^{1K}. Therefore $G^\gamma = (X^\gamma, E^\gamma)$, where the node set is defined by:

$$X^\gamma \subseteq X^2 \mid x_u^\gamma \in SD_{ij} \tag{19}$$

and the links by:

$$E^\gamma \subseteq E^2 \mid e_{uv}^\gamma \in SD_{ij} \tag{20}$$

with $u, v \in \{1, 2, \ldots, N^\gamma\}$.

The weight of each link w_{uv}^γ is thus the sum of the shortest paths that use that link. But the absolute value of w_{uv}^γ alone is not representative of the overlap. A higher value of w_{uv}^γ could merely indicate that there are many more shortest paths in total and not that $\mathcal{C}(\mathcal{S}_{ij})$ critically depends on that link.

A better indication would be the heterogeneity of the distribution of w_{uv}^γ. High heterogeneity would imply that a few links have comparatively high weights and are pivotal to the integrity of the shortest path sets. Therefore, the kurtosis[1] of the distribution of w_{uv}^γ is used as an indication of the critical dependence of $\mathcal{C}(\mathcal{S}_{ij})$ on a few links. We use $\kappa(w_{uv}^\gamma)$ to denote the kurtosis of the distribution of w_{uv}^γ.

To test the hypothesis that there is a relationship between the characteristics of $\mathcal{C}(\mathcal{S}_{ij})$ and the vulnerability of $\mathcal{M}$ we monitor $\kappa(w_{uv}^\gamma)$, $\tilde{P}$ and $\bar{L}$ as we progressively destroy G^2.

3.4 Progressive random link removal strategy

The urban road network is frequently disturbed by congestion, roadworks, accidents, and infrastructure failure. Removing links from G^2 instead of nodes, as most other vulnerability studies do, is a better representation of these disruptions. Furthermore, the way in which an urban road network is disturbed can be described as a combination of *random failures* and *targeted attacks*. Streets that are more central could be more likely to experience congestion or higher traffic loads could increase the likelihood of accidents (targeted attacks). However, accidents, equipment failure and roadworks could really happen anywhere on the grid at any time (random failure). To first create a baseline for this field of inquiry, this study regards the impact of the random removal of links from G^2. Ongoing studies are considering the impact of targeted link removal and comparing that to the results in this paper.

The random link removal strategy is progressive, meaning that multiple iterations are performed during which a set number of links are randomly selected and removed. The metrics described in Section 3.3 are measured after each iteration and the process repeated until $\mathcal{M}$ becomes disconnected.

---

[1] In this article kurtosis refers to *Pearson's measure of kurtosis* which is the unadjusted fourth standardised moment of a distribution. This is not to be confused with *excess kurtosis*.

4 Datasets and simulation

4.1 Generating samples of $\mathcal{M}$

(The algorithms used to generate the datasets described in this section are detailed in Viljoen and Joubert (2017) and the datasets themselves are publicly available on Mendeley (Joubert and Viljoen, 2017))

For each of the three network archetypes $G^{1K}, K \in \{F, S, D\}$, we generate 500 instances of $\mathcal{M}$ with $N^{1K} = 12$ for all three archetypes and G^2 a 10×10 square grid as shown in Figure 1b. The difference between the instances of $\mathcal{M}$ is the random association of X^{1K} to nodes in X^2. Therefore the interlayer connections, $E^{1K,2}$ are unique for each instance.

4.2 Progressive link removal simulation

Each network $\mathcal{M}$ undergoes a progressive random link simulation that performs a series of successive link removals until G^{1K} becomes disconnected. Algorithm 1 describes how links are randomly removed from G^2.

If a node in X^2 has only incoming links and no outgoing links it is called a *sink* as no paths can go out from the node, whereas if the node has only outgoing links it is called a *source*. If a node is completely disconnected with no incoming or outgoing links it is called an *isolate*. If a node in X^2 is associated to a node in X^{1K} and becomes a sink, source or isolate, that instance of $\mathcal{M}$ is broken as G^{1K} becomes disconnected. We extend the longevity of the instances by first evaluating if a random selection of links would render any nodes in X^{1K} sinks, sources or isolates. If it does we repeat the random selection up to a maximum of five times. If five random selections all have a similar result, the last selection of links is removed from G^2 and that instance of $\mathcal{M}$ is broken.

Algorithm 1: Link removal from G^2

    Input   : G^2, X^{1K}, $A^{[1K,2]}$
1 //Grid network, node set of SC network and adjacency matrix of the interlayer connections
    Output: $G^{2'}$
2 //Grid network after link removal
3 $repeat \leftarrow TRUE$
4 $count \leftarrow 0$
5 **while** $repeat$ **and** $count \leq 5$ **do**
6   Randomly select 18 links from G^2; //5% of the original 360 links in G^2
7     $G^{2''} \leftarrow G^2$ with 18 links removed;
8   Use $A^{[1K,2]}$ to determine if any nodes in X^{1K} are sources, sinks or isolates in $G^{2''}$;
9     $count = count + 1$;
10   **if** *no sources, sinks or isolates in* X^{1K} **then**
11         $G^{2'} \leftarrow G^{2''}$;
12         $repeat \leftarrow FALSE$;
13   Return $G^{2'}$

---

5 Results

The progressive link removal simulation is executed on all 500 instances for each network archetype. The simulation is continued until all of the 500 instances are broken. Therefore, the results in this section plot the distributions of the various metrics as measured from a sample of 500 instances for each network archetype.

First we describe the initial characteristics of $\mathcal{C}(\mathcal{S}_{ij})$ by assessing the distributions of $\bar{L}$ and $\tilde{P}$ for the three network archetypes. Thereafter we determine the relationship between $\mathcal{C}(\mathcal{S}_{ij})$ and the resilience of $\mathcal{M}$ throughout the simulation. Lastly, we test whether the characteristics of $\mathcal{C}(\mathcal{S}_{ij})$ could be predictors of robustness and resilience.

$\mathcal{C}(\mathcal{S}_{ij})$ is calculated throughout the simulation by a combination of author-defined Java classes and classes from the JUNG library. These processes were performed on a single core of a Dell PowerEdge 910 running Linux (Ubuntu 16.04LTS), with an Intel Xeon E7540 processor.

The calculation of $\kappa(w_{uv}^{\gamma})$, $\tilde{P}$ and $\bar{L}$ and all other statistical analyses were performed in R version 3.2.4 (R Core Team, 2013) on a Macbook Pro with 8 GB 1867 MHz DDR3 memory and a 2,7 GHz Intel Core i5 processor running OS X El Capitan version 10.11.6.

5.1 Initial characteristics of $\mathcal{C}(\mathcal{S}_{ij})$

The initial characteristics of $\mathcal{C}(\mathcal{S}_{ij})$ depend greatly on the network archetype. Here we compare specifically the distributions of the initial shortest path lengths and the set sizes.

5.1.1 Initial distributions of the shortest path lengths

Figure 4a plots the distributions of $\bar{L}$, showing that the Fully Connected network has significantly shorter paths than the other two archetypes. This directly results from the fact that all node-pairs are directly connected, not requiring rerouting through a hub node. A Kolmogorov-Smirnov test (KS-test) comparing the distributions of $\bar{L}$ for the Single Hub and Double Hub archetypes rejects the H_0 that the distributions are similar with $p = 0.0047$. The lower mean and wider spread in both tails of the Double Hub is explained by the structure of G^{1D}. Intra-hub paths (paths between nodes that share a common hub node) in the Double Hub are generally shorter than the paths of the Single Hub because in the former the placement of the two hubs on the grid effectively splits the grid and therefore the intra-hub nodes are closer to their respective hubs (Viljoen and Joubert, 2017). On the other hand, the inter-hub movements are longer as they have to be rerouted through two hubs. Intra-hub paths account for 45% of the network, inter-hub for 38% and the remainder of the paths are links between one hub and a node assigned to the other hub and vice versa. This last group of paths are not distinctly longer or shorter between the two archetypes. As the majority of the links in the Double Hub will have shorter lengths, $\bar{L}$ is lower.

The diameter of a network is the length of its longest shortest path denoted by $\max(L)$. Figure 4b plots the distributions of $\max(L)$, where once again the Fully Connected archetype shows far shorter paths than the others. This time, the

KS-test fails to reject the H_0 that the distributions of $\max(L)$ for Double Hub and Single Hub are similar with $p = 0.29$.

The Fully Connected archetype is most efficient by a large margin, whereas the clear winner between the two hub archetypes would be instance-specific, with a slight prejudice in favour of the Double Hub.
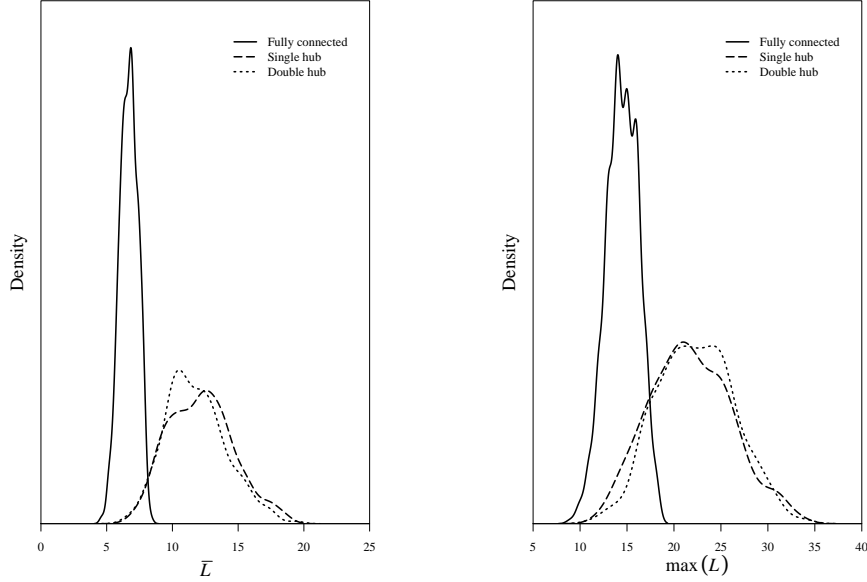
Archetype	Min	Mean	Max
Fully connected	4.58	6.82	8.41
Single hub	6.50	12.17	18.83
Double hub	6.38	11.77	19.48

(a) Average shortest path length, $\bar{L}$.

Archetype	Min	Mean	Max
Fully connected	9	14.48	18
Single hub	12	21.79	34
Double hub	11	22.33	34

(b) Diameter, $\max(L)$.

Fig. 4: Distributions of the diameter and average shortest path lengths for each of the three archetypes.

5.1.2 Initial distributions of the shortest path set sizes

The distributions of the sum of the shortest path set size $\tilde{P}$ are very different to that of $\bar{L}$ and $\max(L)$. Figure 5a plots the distribution of the sum of $\tilde{P}$ for all directly connected node-pairs (SD_{ij}). The distributions have very long tails as confirmed by the high kurtosis values. A kurtosis value of 3 indicates that the distribution is mesokurtic, being no more likely to produce outliers than a normal distribution. A value greater than three indicates a leptokurtic distribution that has a greater degree of "tailedness" than the normal distribution and kurtosis

less than 3 indicates a platykurtic distribution with a smaller likelihood than the normal distribution to produce outliers.

For the Fully Connected archetype all nodes are directly connected thus the distributions in the two graphs are the same. For the hub archetypes the distributions in Figure 5b are disproportionately wider due to the fact that 83.3% of the node-pairs are indirectly linked in these archetypes and each of these sets SI_{ij} is the product of the set sizes of its component sets SD_{ij}. The Double Hub has the smallest set sizes for directly connected node-pairs, owing to the fact that its directly connected nodes are closer together on the grid. Interestingly, the kurtosis the Double Hub in Figure 5a also shows that it is far more likely to produce outliers than the other two archetypes.
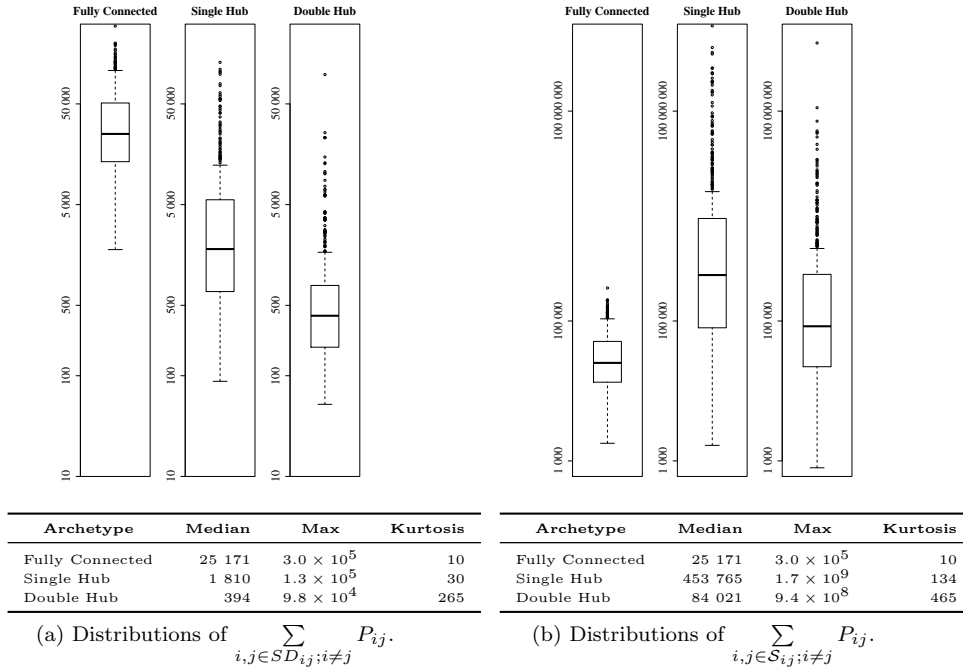

Archetype	Median	Max	Kurtosis
Fully Connected	25 171	3.0×10^5	10
Single Hub	1 810	1.3×10^5	30
Double Hub	394	9.8×10^4	265

(a) Distributions of $\sum\limits_{i,j \in SD_{ij}; i \neq j} P_{ij}$.

Archetype	Median	Max	Kurtosis
Fully Connected	25 171	3.0×10^5	10
Single Hub	453 765	1.7×10^9	134
Double Hub	84 021	9.4×10^8	465

(b) Distributions of $\sum\limits_{i,j \in \mathcal{S}_{ij}; i \neq j} P_{ij}$.

Fig. 5: Analysis of the distributions of the sum of the shortest path set sizes for direct paths (SD_{ij}) and the full network $(\mathcal{S}_{ij})$

When regarding only shortest path lengths one can conclude that there isn't great variance between the different instances of a single archetype, nor is there really a pronounced difference between the two hub archetypes. This is definitely not true when considering the shortest path sets. One randomly generated instance of $\mathcal{M}$ can have vastly different set sizes than the next randomly generated instance.

5.2 Relationship between $\mathcal{C}(\mathcal{S}_{ij})$ and the resilience of $\mathcal{M}$

This section investigates the relationship between $\mathcal{C}(\mathcal{S}_{ij})$ and the resilience of $\mathcal{M}$ in the initial instances of $\mathcal{M}$ and as the simulation progresses. Specifically we seek

correlations between the redundancy of $\mathcal{C}(\mathcal{S}_{ij})$ and the efficiency $\bar{L}$ as well as the overlap of $\mathcal{C}(\mathcal{S}_{ij})$ and the efficiency $\bar{L}$.

5.2.1 Redundancy of $\mathcal{C}(\mathcal{S}_{ij})$

In the initial networks, the logarithm of $\tilde{P}$ is significantly, linearly correlated with $\bar{L}$ as shown in Figure 6 and Table 1. This result is intuitive as both L_{ij} and P_{ij} are functions of the diagonal distance between two nodes.
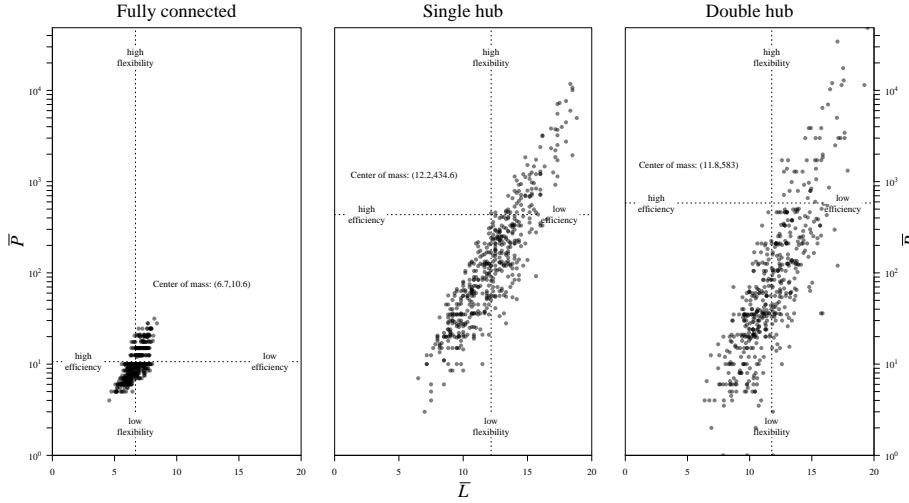

Fig. 6: Shortest set path size versus length in the initial networks.

Table 1: Linear regression of $\tilde{P}$ to $\bar{L}$.

	$\log \hat{y}_F = -0.420 + 0.405 x_F$		$\log \hat{y}_S = -1.800 + 0.540 x_S$		$\log \hat{y}_D = -3.170 + 0.629 x_D$	
std. error	0.115	0.017	0.148	0.012	0.234	0.019
t-value	-3.64	23.59	-12.13	45.27	-13.56	32.30
$Pr(|t|)$	$2.96e^{-4}$	$< 2e^{-16}$	$< 2e^{-16}$	$< 2e^{-16}$	$< 2e^{-16}$	$< 2e^{-16}$

Next we compare the average values of $\tilde{P}$ and $\bar{L}$ taken over all the instances that remain unbroken after a specific link removal iteration. Figure 7a shows that for all three archetypes, the average of $\bar{L}$ increases monotonically as more grid links are removed while the average of $\tilde{P}$ decreases monotonically. Interestingly, the average of $\tilde{P}$ decreases sharply for the hub archetypes, converging with that of the Fully Connected archetype after 30% of grid links are removed. The average of $\bar{L}$, however, does not converge. The changeover between the hub archetypes when 25% of the links are removed is also notable.

The Spearman test is used to test the correlation between $\tilde{P}$ and $\bar{L}$ after each iteration. After each iteration, the sample consisted of all the instances that had

not yet broken. If the sample contained more than five instances, $\tilde{P}$ and $\bar{L}$ was calculated for each instance and the correlation tested. Figure 7b plots the value of Spearman's ρ for each archetype after each iteration and distinguishes between significant (p-value < 0.05) and insignificant (p-value ≥ 0.05) results.

Throughout the simulation there is a significant correlation between these two parameters. Initially the correlation is strongly positive for all three networks but decreases in strength as the networks degrade. The decreasing trend is intuitive. Initially, when the grid is complete, a longer shorter path length automatically implies more alternatives. As the grid becomes more sparse, this is no longer the case.
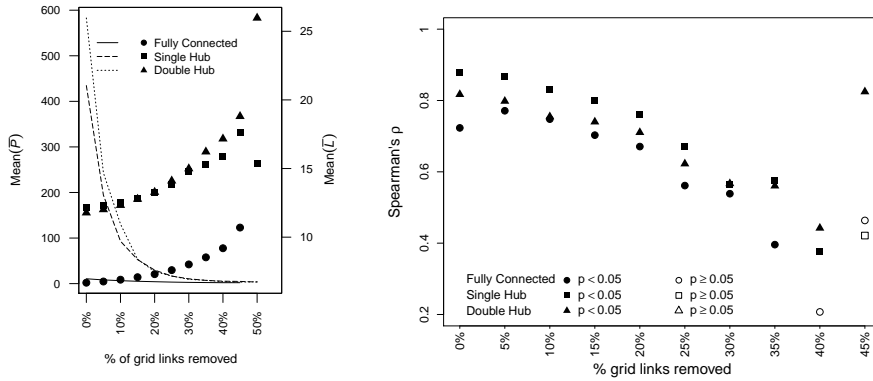
(a) Change in Mean($\tilde{P}$) and Mean($\bar{L}$).

(b) Correlation between $\tilde{P}$ and $\bar{L}$.

Fig. 7: Relationship between $\tilde{P}$ and $\bar{L}$ throughout the link removal iterations.

5.2.2 Overlap within $\mathcal{C}(\mathcal{S}_{ij})$

To compare the overlap in $\mathcal{C}(\mathcal{S}_{ij})$ to resilience in $\mathcal{M}$ we plot the initial values of $\kappa(w_{uv}^{\gamma})$ against $\bar{L}$ (Figure 8). The Fully Connected archetype has the least overlap in $\mathcal{C}(\mathcal{S}_{ij})$ with its distribution of $\kappa(w_{uv}^{\gamma})$ tightly clustered around a kurtosis value of 3.6. The presence of hub nodes actually induces overlap in the shortest path sets, thus the links surrounding the hub nodes are bound to feature on many shortest paths. The Single Hub archetype has only one hub that all other nodes share and thus has more induced overlap than the Double Hub archetype, as evidenced in the higher values for $\kappa(w_{uv}^{\gamma})$.

The Spearman test was repeated to test correlation between $\kappa(w_{uv}^{\gamma})$ and $\bar{L}$. The correlation results were not as definitive (Figure 9b). Fewer correlations were significant and those that were significant were weak. Furthermore the fact that there were both weakly positive and weakly negative correlation results merits discussion. The Fully Connected archetype is primarily positively correlated. Initially these networks have low overlap which increases as fewer alternatives on the grid force the shortest paths to use common links. The hub networks, on the other hand, start off with a higher degree of overlap induced by the hubs. As shortest
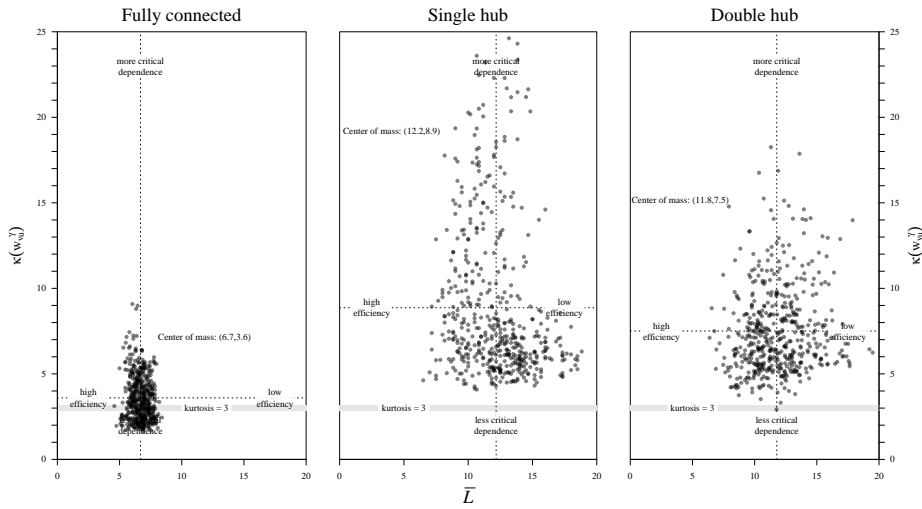
Fig. 8: $\kappa(w_{uv}^{\gamma})$ versus $\bar{L}$ in the initial networks.

paths become longer it actually offers opportunities for diversification in terms of the links. Therefore these correlations are initially negative. However, as the simulation continues, the sparsity of the grid again forces overlap in the shortest paths and thus the correlations become positive.

Figure 9a illustrates how the overlap first increases for all archetypes and then decreasing (sharply in the case of the hub archetypes).
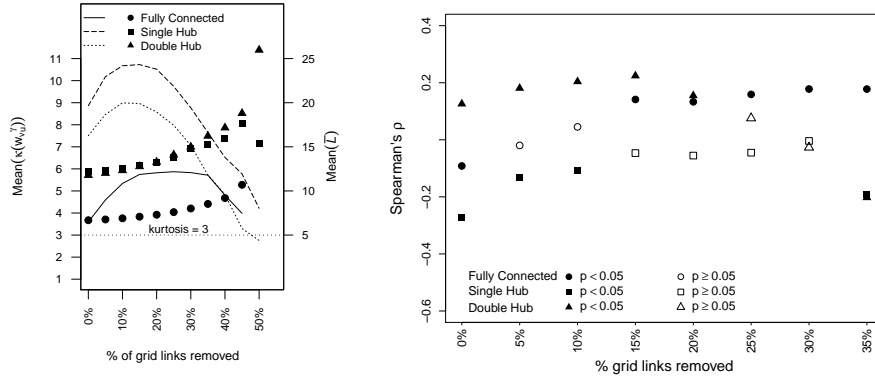
(a) Change in average $\kappa(w_{uv}^{\gamma})$ and $\bar{L}$.

(b) Correlation between $\kappa(w_{uv}^{\gamma})$ and $\bar{L}$.

Fig. 9: Relationship between $\kappa(w_{uv}^{\gamma})$ and $\bar{L}$ throughout link removal iterations.

The characteristics of the shortest path sets for different instances of a specific archetype and between the three archetypes is definitely non-trivial. Furthermore, there are significant relationships between the redundancy and overlap in $\mathcal{C}(\mathcal{S}_{ij})$

and its resilience to progressive random link disturbance. The next question is whether these relationships could be predictors of future resilience and robustness of networks.

6 $\mathcal{C}(\mathcal{S}_{ij})$ as predictor of robustness and resilience

6.1 Aggregate statistics

Figure 10a shows the cumulative percentage of instances that had broken after each iteration of the simulation. The distribution is very similar across the three archetypes with the Fully Connected archetype breaking faster in the mid-range of the simulation. All instances of the Fully Connected archetype were broken after 50% of the links were removed with the final surviving instances of the hub archetypes breaking after 55% link removal. Counter-intuitively, robustness is indifferent to the type of archetype, despite the broadly varying shortest path set characteristics.
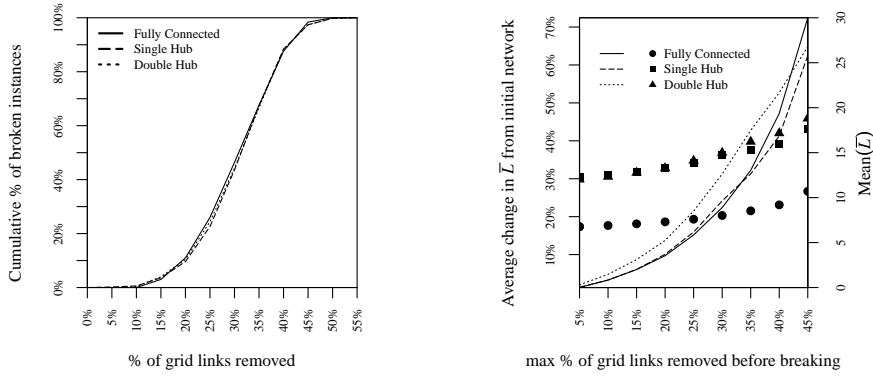
(a) Cumulative percentage of instances that broke with each link removal iteration.

(b) Average percentage change in $\bar{L}$ from the initial networks.

Fig. 10: Comparing efficiency and robustness of instances across link removal iteration.

In terms of resilience, which is the ability to maintain a certain path length despite disruption, we consider the change in the average of $\bar{L}$ from its initial value in Figure 10b. The Double Hub archetype loses its efficiency far more rapidly than the other archetypes but in the end it is the Fully Connected archetype that has degraded the most while the Single Hub archetype degraded the least.

While it seems unlikely that the broadly varying shortest path statistics of the initial archetypes held any predictive capability regarding overall resilience or robustness, it is worthwhile digging deeper to find out if any significant patterns precede a step change in efficiency or break in an instance.

6.2 Pattern-finding in individual instance trajectories

After each progressive link removal iteration, more instances have broken. We seek to identify patterns relating to $\tilde{P}$, $\kappa(w_{uv}^{\gamma})$ or $\bar{L}$ that distinguish those instances that are known to have broken in the next iteration from those that are known to have survived. We do this by defining a distribution of values, X, for those instances that would break and another distribution of values, Y, for the surviving instances. Using a KS-test we determine whether there is any significant difference between these two distributions.

Using the KS-test we test nine sets of distributions:

- $\tilde{P}$: Initial values, terminal values and % change from initial value;
- $\kappa(w_{uv}^{\gamma})$: Initial values, terminal values and % change from initial value; and
- $\bar{L}$: Initial values, terminal values and % change from initial value.

The comparison was made after each iteration between 10% to 50% grid link removal. Overwhelmingly, in 92% of the cases, the KS-test showed that H_0, which states that X and Y are from the same distribution, could not be rejected, meaning that there is no distinguishable difference between the values for those instances that are about to break and those that will survive. Of the 20 cases that yielded p-value < 0.05 and could thus reject H_0, 11 related to $\bar{L}$ for the Fully Connected archetype.

The test was then expanded to include in X those instances whose efficiency would decrease in the next iteration and not just those that would break. The sample sizes of Y only allowed for tests up to 20%, 35% and 40% for the Fully Connected, Single Hub and Double Hub archetypes, respectively. Only 3% of these tests succeeded in rejecting H_0 that X and Y were from different distributions.

Therefore it can be concluded that neither the initial values, the values measured before the final link removal iteration or the rate of change from the initial value for any of the metrics can be used as predictors of which instances would be resilient and/or robust to the next random link disruption.

These results pertain to the context of progressive *random* link disruption which, by definition, cannot easily be predicted. Even though the observed metrics may in many cases have quantified vulnerability in networks, such as a critical dependency or the absence of alternatives, these vulnerabilities may not have been exploited by the next round of link disruptions.

7 Conclusion

The risks facing supply chains and their inherent vulnerability to these risks are increasing, so much so that industry has elevated Supply Chain Risk Management as a top priority. Nonetheless, the quantitative tools required to better capture the trade-offs between supply chain efficiency and supply chain resilience are scant. This paper specifically addressed this gap. We presented a multilayered complex network formulation of three paradigmatic supply chain archetypes on a directed grid network emulating the urban road network. This formulation allowed us to identify three metrics related to the unique collection of shortest path sets $\mathcal{C}(\mathcal{S}_{ij})$ of the multilayered network $\mathcal{M}$ namely the average length of the shortest paths

$(\bar{L})$, the median of the size of the shortest path sets $(\tilde{P})$ and the kurtosis of the link weight of the link betweenness network $(\kappa(w_{uv}^{\gamma}))$.

The initial distributions of these metrics showed that the type of supply chain archetype and the specific randomly generated instance of $\mathcal{M}$ produce very broad ranges of metric values. It also showed a significant relationship between $\tilde{P}$ and $\bar{L}$ and $\kappa(w_{uv}^{\gamma})$ and $\bar{L}$, proving the hypothesis that there is a relationship between the characteristics of $\mathcal{C}(\mathcal{S}_{ij})$ and the resilience of $\mathcal{M}$. However, the robustness of $\mathcal{M}$ (i.e. how quickly the supply chain network became disconnected) is indifferent to these metrics. In addition, the metrics held no predictive power regarding which instances would break or lose efficiency in successive link removal iterations.

In the case of redundancy the lack of predictive power could be due to two reasons:

1. By using the centrality of the distribution of P_{ij} it is possible to overlook the fact that some shortest path sets have very little or no redundancy (i.e. underestimating the left tail of the distribution).
2. Redundancy by itself overlooks the impact of overlap. Thus a highly redundant network could still be critically dependent on one link.

In the case of overlap there are also two potential reasons for the lack in predictive power:

1. Kurtosis may not be the most appropriate measure to capture the characteristics of the right tail of the distribution of w_{uv}^{γ}.
2. The random link removal does not prioritise the removal of these high priority links. Therefore even though there are critically important links, they have an equal probability of removal than any other link.

Ongoing research is investigating additional metrics and refinements to the metrics of this study. In addition simulations are being expanded to include targeted attack strategies.

Although this paper has developed a baseline formulation and initial results for investigating the impact of the dependency of a supply chain on the urban transport network, the following research questions must be addressed before the results can be used in industry boardrooms:

– How close are real-life supply chain networks and urban road networks to the network archetypes presented in this paper and can these results be generalised to these hybrids?
– How does weighting the different layers of $\mathcal{M}$ influence the results and does this provide more valid results?
– What kind of link removal scheme best describes the routine disturbances of urban road networks and how could these be better incorporated?

References

Bartholdi and Jarumaneeroj, P.and Ramudhin, A. (2014). A new connectivity index for container ports. Technical report, The Supply Chain & Logistics Institute, Georgia Institute of Technology.

Bellamy, M. and Basole, C. (2013). Network analysis of supply chain systems: A systematic review and future research. *Systems Engineering*, 16(2):235–249.

Bhatia, G., Lane, C., and Wain, A. (2013). *Building Resilience in Supply Chains*. Risk Response Network. World Economic Forum.

Boccaletti, S., Bianconi, G., Criado, R., Del Genio, C. I., Gómez-Gardeñes, J., Romance, M., Señdina Nadal, I., Wang, Z., and Zanin, M. (2014). The structure and dynamics of multilayer networks. *Physics Reports*, 544(1):1–122.

Bóta, A., Gardner, L., and Khani, A. (2017). Identifying critical components of a public transit systems for outbreak control. *Networks and Spatial Economics*. Available online from https://doi-org.uplib.idm.oclc.org/10.1007/s11067-017-9361-2.

Cardillo, A., Zanin, M., Gómez-Gardeñes, J., Romance, M., García del Amo, A., and Boccaletti, S. (2015). Modeling the multi-layer nature of the European air transport network: Resilience and passengers re-scheduling under random failures. *The European Physical Journal Special Topics*, 215(1):23–33.

Chacon, N., Doherty, S., Hayashi, C., Green, R., and Lever, I. (2012). *New Models for Addressing Supply Chain and Transport Risk*. Risk Response Network. World Economic Forum.

Choi, T., Dooley, K., and Rungtusanatham, M. (2001). Supply networks and complex adaptive systems: control versus emergence. *Journal of Operations Management*, 19:351–366.

Choi, T. Y. and Wu, Z. (2009). Taking the leap from dyads to triads: Buyer-supplier relationships in supply networks. *Journal of Purchasing & Supply Management*, 15:263–266.

Cohen, R., Erez, K., Avraham, D., and Havlin, S. (2000). Resilience of the internet to random breakdowns. *Physical Review Letters*, 85:4626–4628.

Cohen, R., Erez, K., Avraham, D., and Havlin, S. (2001). Breakdown of the internet under intentional attack. *Physical Review Letters*, 86(16):3682–5.

Criado, R., Hernández-Bermejo, B., and Romance, M. (2007). Efficiency, vulnerability and cost: An overview with applications to subway networks worldwide. *International Journal of Bifurcation and Chaos*, 17:2289–2301.

Crucitti, P., Latora, V., Marchiori, M., and Rapisarda, A. (2004). Error and attack tolerance of complex networks. *Physica A*, 340:388 – 394.

Crucitti, P., Latora, V., and Porta, S. (2006). Centrality measures in spatial networks of urban streets. *Physical Review E*, 73:036125.

Danziger, M. M., Bashan, A., Berezin, Y., Shekhtman, L. M., and Havlin, S. (2014). An introduction to interdependent networks. *International Conference on Nonlinear Dynamics of Electronic Systems*, CCIS 438:189–202.

Darayi, M., Barker, K., and Santos, J. (2017). Component importance measures for multi-industry vulnerability of a freight transportation network. *Network and Spatial Economics*. Available online from https://doi-org.uplib.idm.oclc.org/10.1007/s11067-017-9359-9.

Duan, Y. and Lu, F. (2014). Robustness of city road networks at different granularities. *Physica A*, 411:21–34.

Ducruet, C. (2013). Network diversity and maritime flows. *Journal of Transport Geography*, 30:77–88.

Ducruet, C. (2016). The polarization of global container flows by interoceanic canals: geographic coverage and network vulnerability. *Maritime Policy & Management*, 43(2):242–260.

Ducruet, C., Lee, S., and Ng, A. (2010a). Centrality and vulnerability in liner shipping networks: revisiting the Northeast Asian port hierarchy. *Maritime Policy*

and Management, 37:17–36.

Ducruet, C. and Notteboom, T. (2012). The worldwide maritime network of container shipping: spatial structure and regional dynamics. *Global Networks*, 12(3):395–423.

Ducruet, C., Rozenblat, C., and Zaidi, F. (2010b). Ports in multi-level maritime networks: evidence from the Atlantic (1996 – 2006). *Journal of Transport Geography*, 18:508–518.

Fraser, D., Notteboom, T., and Ducruet, C. (2014). Peripherality in the global container shipping network: The case of the Southern African container port system. *GeoJournal*, pages 1–13.

Gallotti, R. and Barthelemy, M. (2014). Anatomy and efficiency of urban multimodal mobility. *Scientific Reports*, 4:6911.

Gallotti, R. and Barthelemy, M. (2015). The multilayer temporal network of public transport in Great Britain. *Scientific Data*, 2:140056.

González Laxe, F., Freire Seoane, M., and Pais Montes, C. (2012). Maritime degree, centrality and vulnerability: port hierarchies and emerging areas in containerized transport (2008 – 2010). *Journal of Transport Geography*, 24:33–44.

Gudmundsson, A. and Mohajeri, N. (2013). Entropy and order in urban street networks. *Scientific Reports*, 3:3324.

Háznagy, A., Fi, I., London, A., and Németh, T. (2015). Complex network analysis of public transportation networks: a comprehensive study. *Models and Technologies for Intelligent Transportation Systems (MT-ITS)*.

Hearnshaw, E. and Wilson, M. (2013). A complex network approach to supply chain network theory. *International Journal of Operations & Production Management*, 33(4):442–469.

Heckmann, I., Comes, T., and Nickel, S. (2015). A critical review on supply chain risk — definition, measure and modeling. *Omega*, 52:119 – 132.

Holl, A. and Mariotti, I. (2017). The geography of logistics firm location: The role of accessibility. *Networks and Spatial Economics*. Available online from https://doi-org.uplib.idm.oclc.org/10.1007/s11067-017-9347-0.

Hu, Y. and Zhu, D. (2009). Empirical analysis of the worldwide maritime transportation network. *Physica A*, 388:2061–2071.

Jiang, B. (2007). A topological pattern of urban street networks: universality and peculiarity. *Physica A*, 384:647–655.

Jiang, B. and Claramunt, C. (2004). Topological analysis of urban street networks. *Environmental Planning B*, 31:151–162.

Joubert, J. and Axhausen, K. W. (2013). A complex network approach to understand commercial vehicle movement. *Transportation*, 40(3):729–750.

Joubert, J. W. and Viljoen, N. M. (2017). Multilayer complex networks, v3. *Mendeley Data*. Available online from http://dx.doi.org/10.17632/268byhmvv5.3.

Kaluza, P., Kölzsch, A., Gastner, M. T., and Blasius, B. (2010). The complex network of global cargo ship movements. *Journal of the Royal Society Interface*, 7:1093 – 1103.

Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., and Porter, M. A. (2014). Multilayer networks. *Journal of Complex Networks*, 2:203–271.

Kurant, M. and Thiran, P. (2006a). Extraction and analysis of traffic and topologies of transportation networks. *Physical Review E*, 74:036114.

Kurant, M. and Thiran, P. (2006b). Layered complex networks. *Physical review letters*, 96:138701.

Lee, K.-M., Min, B., and Goh, K.-I. (2015). Towards real-world complexity: and introduction to multiplex networks. *European Physical Journal B*, 88(48).

Lordan, O. (2014). Study of the full-service and low-cost carriers network configuration. *Journal of Industrial Engineering and Management*, 7(5):1112–1123.

Lordan, O., Sallan, J, E. N., and Gonzalez-Prieto, D. (2016). Robustness of airline route networks. *Physica A*, 445:18–26.

Lordan, O., Sallan, J., Simo, P., and Gonzalez-Prieto, D. (2014a). Robustness of the air transport network. *Transportation Research Part E*, 68:155–163.

Lordan, O., Sallan, J., Simo, P., and Gonzalez-Prieto, D. (2015). Robustness of airline alliance route networks. *Communications in Nonlinear Science and Numerical Simulation*, 22(1–3):587–595.

Lordan, O., Sallan, J. M., and Simo, P. (2014b). Study of the topology and robustness of airline route networks from the complex network approach: a survey and research agenda. *Journal of Transport Geography*, 37:112–120.

Lotero, L., Cadillo, A., Hurtado, R., and Gómez-Gardeñes, J. (2016). *Several multiplexes in the same city: The role of socioeconomic differences in urban mobility*, chapter 9, pages 149–168. Springer International Publishing.

Mattsson, L.-G. and Jenelius, E. (2015). Vulnerability and resilience of transport systems — a discussion of recent research. *Transportation*, 81:16 – 34.

Mohamed-Chérif, F. and Ducruet, C. (2016). Regional integration and maritime connectivity across the Maghreb seaport system. *Journal of Transport Geography*, 51:280–293.

Mouronte, M. L. and Benito, R. M. (2012). Structural properties of urban bus and subway networks of Madrid. *Networks & Heterogeneous Media*, 7:415–428.

O'Kelly, M. (2015). Network hub structure and resilience. *Networks and Spatial Economics*, 15(4):235–251.

Ortigosa, J. and Menendez, M. (2014). Traffic performance on quasi-grid urban structures. *Cities*, 36:18 – 27.

Ouyang, M., Zhao, L., Hong, L., and Pan, Z. (2014). Comparisons of complex networks based models and real train flow model to analyze Chinese railway vulnerability. *Reliability Engineering and System Safety*, 123:38–46.

Pais Montes, C., Freire Seoane, M., and González Laxe, F. (2012). General cargo and containership emergent routes: A complex networks description. *Transport Policy*, 24:126–140.

Parshani, R., Rozenblat, C., Ietri, D., Ducruet, C., and Havlin, S. (2011). Intersimilarity between coupled networks. *Europhysics Letters*, 92(6):68002.

Pastoras-Satorras, R. and Vespignani, A. (2001). Epidemic spreading in scale-free networks. *Physical Review Letters*, 86(14):3200.

Pathak, S., Day, J., Nair, A., Sawaya, W., and Kristal, M. (2007). Complex and adaptivity in supply networks: Building supply network theory using a complex adaptive systems perspective. *Decision Sciences*, 38(4):547 – 580.

Peck, H. (2006). Reconciling supply chain vulnerability, risk and supply chain management. *International Journal of Logistics: Research and Applications*, 9(2):127–142.

Porta, S., Latora, V., Wang, F., Rueda, S., Strano, E., Scellato, S., Cardillo, A., Belli, E., Cárdenas, F., Cormenzana, B., and Latora, L. (2012). Street centrality and the location of economic activities in Barcelona. *Urban Studies*, 49(7):1471–

1488.

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rao, S. and Goldsby, T. (2009). Supply chain risk: a review and typology. *The Interantional Journal of Logistics Management*, 20(1):97–123.

Reggiani, A., Bucci, P., and Russo, G. (2011). Accessibility and network structures in german commuting. *Networks and Spatial Economics*, 11(4):621–641.

Rodríguez-Núñez, E. and García-Palomares, J. C. (2014). Measuring the vulnerability of public transport networks. *Journal of Transport Geography*, 35:50–63.

Rupi, F., Bernardi, S., Rossi, G., and Danesi, A. (2015). The evaluation of road network vulnerability in mountainous areas: A case study. *Network and Spatial Economics*, 15(2):397–411.

Salehi, M., Sharma, R., Marzolla, M., Magnani, M., Siyari, P., and Montesi, D. (2015). Spreading processes in multilayer networks. *IEEE Transactions on Network Science and Engineering*, 2(2):65–83.

Sen, P., Dasgupta, S., Chatterjee, A., Sreeram, P., Mukherjee, G., and Manna, S. (2003). Small-world properties of the Indian railway network. *Physical Review E*, 67:036106.

Shekhtman, L. M., Bagrow, J. P., and Brockmann, D. (2014). Robustness of skeletons and salient features in networks. *Journal of Complex Networks*, 2(2):110–120.

Solé-Ribalta, A., Gómez, S., and Arenas, A. (2016). Congestion induced by the structure of multiplex networks. *Physical review letters*, 116(10):108701.

Stauffer, D. and Aharony, A. (1991). *Introduction to Percolation Theory*. Taylor & Francis, London, second edition edition.

Strano, E., Cardillo, A., Iacoviello, V., Latora, V., Messora, R., Porta, S., and Scellato, S. (2009). Street centrality vs. commerce and service locations in cities: A kernel density correlation case study in Bologna, Italy. *Environmental Planning B*, 36:450–465.

Strano, E., Shai, S., Dobson, S., and Barthelemy, M. (2015). Multiplex networks in metropolitan areas: generic features and local effects. *Journal of the Royal Society Interface*, 12(111):20150651.

Sun, X., Wandelt, S., and Cao, X. (2017). On node criticality in air transportation networks. *Networks and Spatial Economics*. Available online from https://doi-org.uplib.idm.oclc.org/10.1007/s11067-017-9342-5.

Taylor, M. A. and Susilawati (2012). Remoteness and accessibility in the vulnerability analysis of regional road networks. *Transportation Research Part A*, 46:761–771.

Thadakamalla, H., Raghavan, U., Kumara, S., and Albert, R. (2004). Survivability of multiagent-based supply networks: A topological perspective. *IEEE Intelligent Systems*, pages 24–31.

Tsiostas, D. and Polyzos, S. (2015). Analyzing the maritime transportation system in Greece: a complex network approach. *Networks and Spatial Economics*, 15(4):981–1010.

Tsiostas, D. and Polyzos, S. (2017). The complexity in the study of spatial networks: An epistemological approach. *Network and Spatial Economics*. Available online from https://doi-org.uplib.idm.oclc.org/10.1007/s11067-017-9354-1.

Tsiotas, D. and Polyzos, S. (2015). Decomposing multilayer transportation networks using complex network analysis: a case study for the Greek aviation net-

work. *Journal of Complex Networks*, page cnv003.

Viljoen, N. M. and Joubert, J. (2017). Multilayered complex network datasets for three supply chain layouts on an urban road grid. Working paper 062, Centre for Transport Development, University of Pretoria.

Viljoen, N. M. and Joubert, J. W. (2016). The vulnerability of the global container shipping network to targeted link disruption. *Physica A*, 462:396–409.

Wagner, S. and Bode, C. (2006). An empirical investigation into supply chain vulnerability. *Journal of Purchasing and Supply Management*, 12:301–312.

Wagner, S. and Neshat, N. (2010). Assessing the vulnerability of supply chains using graph theory. *International Journal of Production Economics*, 126:121–129.

Zadeh, A. S. M. and Rajabi, M. A. (2013). Analyzing the effect of the street network configuration on the efficiency of an urban transport system. *Cities*, 31:285–297.

Zanin, M., Lacasa, L., Cea, M., and Crístobal, S. (2009). A dynamical model for the air transportation network. *ECMS*, pages 256–261.

Zanin, M. and Lillo, F. (2013). Modelling the air transport with complex networks: a short review. *The European Physical Journal Special Topics*, 215(5):5–21.

Zhuo, Y., Peng, Y., Liu, C., Liu, Y., and Long, K. (2011). Traffic dynamics on layered complex networks. *Physica A*, 390:2401–2407.