

The Impact of Road Networks on Facility Location Modeling: An Interpretation from a Network Topological Perspective

Bin Zhou

Department of Geography and GIS, Southern Illinois University Edwardsville, Edwardsville, Illinois, USA
Email: bzhou@siue.edu

How to cite this paper: Zhou, B. (2025) The Impact of Road Networks on Facility Location Modeling: An Interpretation from a Network Topological Perspective. *Journal of Geographic Information System*, 17, 167-198. <https://doi.org/10.4236/jgis.2025.174009>

Received: April 17, 2025

Accepted: July 20, 2025

Published: July 23, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution-NonCommercial International License (CC BY-NC 4.0). <http://creativecommons.org/licenses/by-nc/4.0/>



Open Access

Abstract

This research investigates the impact of the road network topological structure on facility location modeling. We create four types of road networks, *i.e.*, the radial, the grid, the ring, and the ring-radial networks, and characterize them using centralization indices from social network analysis and connectivity indices from planar network analysis. We perform the p -median location-allocation model on each road network. To control for confounding factors, we incorporate the facility capacity, the road type, the demand locations, and the demand quantity in the model. On each road network and with every combination of confounding factor levels, we perform the p -median location-allocation model 100 times and use the large result samples to conduct spatial and statistical analysis. The radial and ring networks tend to have long facility-demand connecting paths along the radial or ring roads. The grid network has mostly short and localized connecting paths. The ring-radial network has a mix of long and short connecting paths. The optimal models come from the grid network, followed by the ring-radial network. The worst result is from the ring network, followed by the radial network. The analysis of variance confirms this result under all combinations of the four control factors. We attribute the spatial pattern of connecting paths and the model result to the topological characteristics of the networks. The grid network has the highest connectivity and distributed centrality. Both the ring and radial networks have low connectivity. It is interesting that blending these two poor-performing networks into the ring-radial network improves connectivity and generates the second-best result. Understanding the relationship between the road network topological structure and facility location modeling implies that transportation development should be focused on improving connectivity rather than just building more roads.

Keywords

Facility Location, Location-Allocation, Road Network, Graph Topology, Connectivity, Centrality

1. Introduction

The facility location problem locates productive/service facilities and allocates customers to these facilities [1] [2]. An essential component in all facility location modeling is the travel/shipping distance. While some works use the Euclidean distances [3], many use road network distances which are displayed in a minimum travel distance matrix over origin and destination locations. Some location-allocation models include routing as part of the modeling [4] [5]. Routing refers to the search for optimal paths with desired attributes such as the minimum distance. Routing is often performed on graphs that consist of nodes and edges. Nodes represent road junctions and end points, or communities located on roads; edges represent road segments between nodes. In fact, a distance matrix is the product of routing among origin and destination pairs.

Given the prominent role of road networks in facility location optimization, it is surprising that there has not been much research on the impact of road networks' structure on facility location modeling, with a couple of exceptions as discussed in the literature review. Although distances and routing constitute part of location-allocation modeling, a road network is more than just distance. A road network spreads over space with edges being connected by nodes in unique ways, which are eventually manifested in topological characteristics such as network centrality, connectivity, etc. These topological attributes, in turn, affect routing and the resultant distance matrices, and ultimately the outcome of location models. Although distance matrices embody the impact of these topological attributes, understanding and recognizing the relationship explicitly may help improve location modeling. For communities that intend to attract business investment, building roads and improving road quality are often a strategy. However, improving road conditions is more than building roads. Our study will show that improving road network structures and their topological characteristics is just as important.

In brief, it is in the interests of academic research and public policy to understand the impact of road networks on facility location modeling. In this paper, we set out to take a step in this direction. We create four types of road networks, *i.e.*, the radial, grid, ring, and ring-radial networks, and characterize them using centralization indices from social network analysis and connectivity indices from planar network analysis. To isolate possible confounding factors, we run models with control factors restricting the facility service capacity, road types, demand point locations, and demand quantities at various demand sites (see details in the Methods section). We use a heuristic algorithm in routing, which brings in certain randomness in optimization. As a result, on every road network, we perform models

with each combination of confounding factors 100 times to obtain large samples, which are then used in spatial and statistical analyses.

The remainder of this paper is organized as follows. The next section discusses related literature, followed by the methods used. Section 4 presents the result of the geographical analysis of two model types to illustrate the spatial patterns associated with topological structure using different networks. Section 5 expands the discussion to all model variants with the statistical method of the analysis of variance. The last section completes the study.

2. Literature Review

2.1. Facility Location Problem

Facility location problems originated in the early 1600s as the Fermat Problem, which sought to identify a point connecting three others with a minimum total distance [6]. Weber sets the problem in the context of a business firm that minimizes the total shipping cost when producing a product for a market and using localized materials [7]. Over time, algorithms were designed to solve the Weber problem on a continuous plane [8], and on networks [9] [10].

The Weber problem eventually becomes the p -median location-allocation problem. This involves setting up a p number of facilities to serve a certain number of demand locations with the minimum travel cost [11]-[13]. An example is to set up fire department stations across a city so that fire events can be dealt in a timely manner. In situations where service capacities at facilities are limited (such as the service capacities of homeless shelters), it is a capacitated p -median problem where the service provided at each facility cannot go beyond its capacity.

2.2. Recent Research

While the facility location problem was established as an important topic in spatial modeling in the 1960s, efforts to widen the study have been sustained over time. In recent years, models have been developed with uncertainty parameters. Roozkhosh and Farimani [14] develop a hub location model where congestion may cause delays, leading to uncertainty. In their model, the probability distribution is built on a Monte Carlo simulation. Similarly, Li *et al.* [15] propose a multi-objective emergency logistics model with uncertain information. Instead of using probability theory, they base uncertainty parameters on uncertainty theory. In yet another study, Saif and Delage [16] design a model with distributionally robust optimization where uncertainty is governed by a probability distribution that is itself subject to uncertainty. Similarly, Li *et al.* [17] develop a model under demand uncertainty. Celik and Ok [18] develop a model for locating the optimal locations of electric vehicle charging stations. Instead of using conventional mathematical programming, they use the genetic algorithm, which allows learning during the modeling.

2.3. Incorporating GIS and ABM

Recent decades have seen the adoption of multi-agent simulations and the inte-

gration of GIS in location-allocation modeling. The examples of agent-based modeling (ABM) include a recent study by Han *et al.* [19], who solved a model of charging station locations for electric vehicles. Similarly, Sun *et al.* [20] use a two-stage model to determine the optimal sites to set up printers on a university campus, minimizing the students' travel distance to the printing service. The first stage uses optimization to obtain the printer sites, while the second stage validates and visualizes the result with ABM. In a somewhat similar framework, a study by Bartkowski *et al.* [21] optimizes biophysically optimal land-use and then compares it with results from an ABM to determine better solutions. Zhao and Zhang [22] use an ABM and the generic algorithm to optimize a multi-stage production process.

Efforts in integrating Geographic Information Systems (GIS) with facility location studies take advantage of GIS functionalities in processing and handling spatial data [23]. For example, Chen *et al.* [24] use buffering and overlaying techniques to find optimal tourism locations, creating alternative metaheuristic algorithms for optimization. Karim and Awawdeh [25] utilize GIS network analysis to map out the accessibility of 12 types of services in a city and then use a location-allocation model to determine deficiencies in services that can be targeted for improvement. Vafaeinejad *et al.* [26] solve a generalized location-allocation problem, the Vector Assignment Ordered Median Problem, which sees a specific location-allocation problem as a special case of the general problem.

2.4. Location-Allocation Modeling and Network Structure

The issue of the impact of topological characteristics on location-allocation modeling does not seem to get much attention in the literature. Some studies touch upon the relationship between network distances or shapes and spatial modeling. Perreur and Thisse [27] discuss how movements on road networks are affected by the geometry of the roads. Schilling *et al.* [28] investigate how distance matrices built with Euclidean distance and path distance or randomly created may affect the computational efforts in solving a p -median problem. There have also been discussions regarding the advantages of using network distance vs. Euclidean distance in facility location modeling [3] [29].

Peeters and Thomas [30] generate the first major study we can find that specifically investigates the impacts of road network shape on facility location modeling. They design road networks of various shapes on a lattice space with evenly distributed points. They connect points in various ways and create what they call regular and radial networks. The regular networks have local connections in cardinal (horizontal and vertical) and diagonal directions, while the radial networks have long roadways extending diagonally across open spaces. They perform p -median models on these networks and conclude that the denser, regular road networks generate a lower average distance in modeling. On the other hand, radial roads, especially those with no peripheral links, lead to higher distances in model solutions. Peeters and Thomas [30] also use the maximum distance between de-

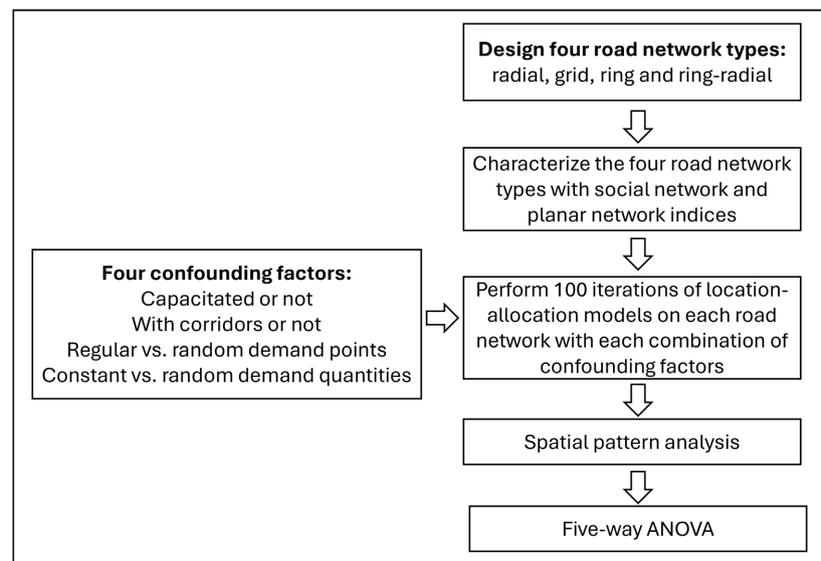
mand points and their closest facility as an alternative measure of modeling optimality. They argue that some network shapes, such as dense grids with diagonal radial lines, naturally generate better modeling results than other shapes, such as sparse grids, coarse radial lines, etc. A more recent study by Zhao *et al.* [31] explores the impact of road system complexity on facility location modeling. The authors run p -median models on road networks of increasing complexity by using the simplest European highways, then adding different grades of national roads, and further adding different grades of regional roads into the model. They conclude that there is a limit to the benefit from increasing road network complexity in the p -median problem modeling. It is interesting that while Peeters and Thomas [30] find denser road systems with more roads generate better results, Zhao *et al.* [31] suggest such benefits may not increase indefinitely.

Works by Peeters and Thomas [30] and Zhao *et al.* [31] are seminal and inspirational. The shortcoming is that they did not characterize the road system using the widely known network topological measures. This disconnect prevents them from generalizing their findings at a deeper theoretical level.

In summary, location-allocation modeling has evolved over time to models on continuous space and network space, various model types (demand coverage, minimum costs, aversion of risks, etc.), and with new modeling tools including GIS and ABM. However, at the heart of location-allocation modeling is the routing on road systems, which is shaped by the topological structure of road networks. It is in this respect that no extensive efforts have been seen from the research community. This paper seeks to contribute to our understanding of the impact of network structures on facility location modeling.

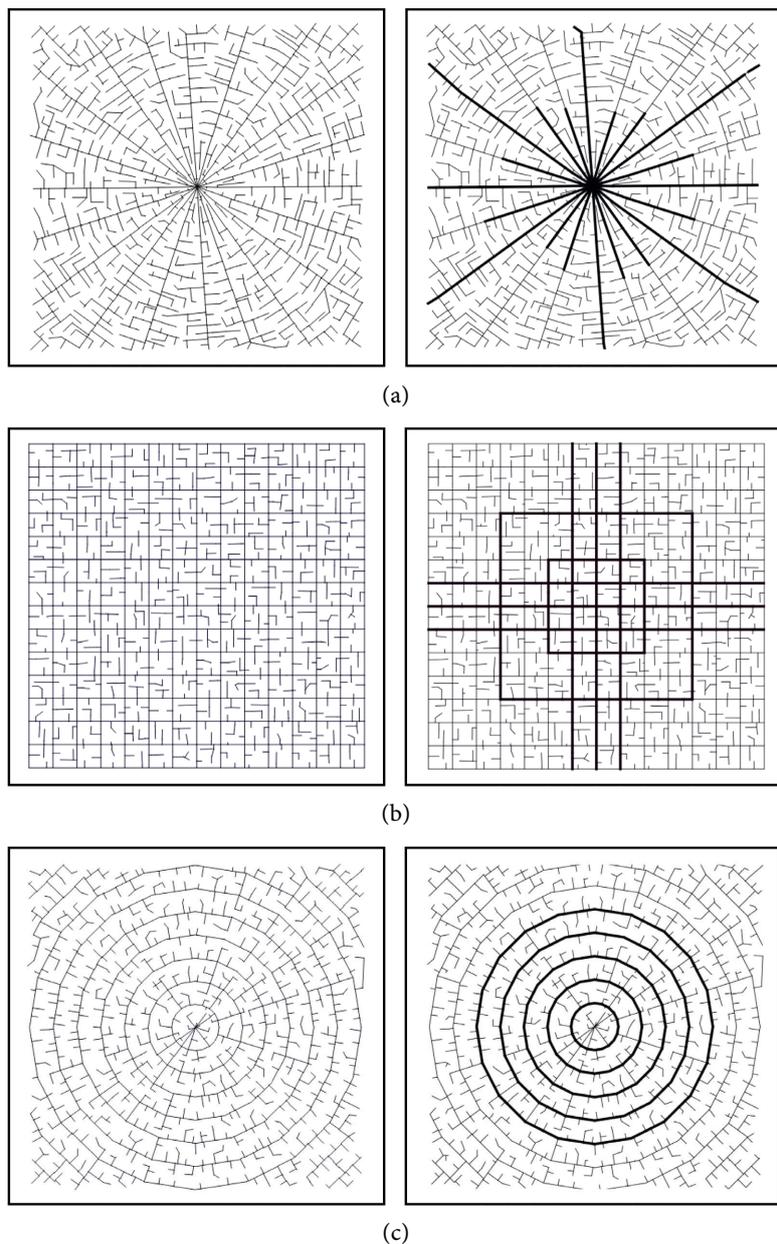
3. The Methods

The following flowchart illustrates the methods used in the study. We will detail and explain what is stated in the flowchart sequentially below.



3.1. Four Road Networks

For the purpose of this study, we design four types of road networks, the radial, grid, ring, and ring-radial networks, as shown by maps in **Figure 1**. At this juncture, it is sufficient to focus on the maps on the left in **Figure 1**. These road networks are created in a square space of equal size and are designed to accentuate distinctly different road patterns of the radial, grid, ring, and ring-radial structures. While the radial, grid and ring networks are inspired by the star (or hub-spoke, polar, center-oriented, radial, diameter, arcantial), lattice (or grid, hierarchical-mesh), and ring topologies in general discussion of network types [7] [32] [33], hybrid networks are also often found in reality. Thus, we create the ring-radial network as a hybrid between the radial and ring systems.



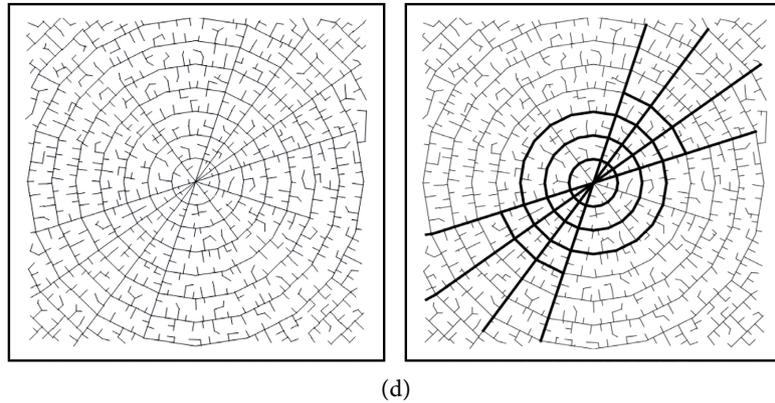


Figure 1. (a) The radial network: no corridors (left), with corridors (on the right with thick lines); (b) The grid network: no corridors (left), with corridors (on the right with thick lines); (c) The ring network: no corridors (left), with corridors (on the right with thick lines); (d) The ring-radial network: no corridors (left), with corridors (on the right with thick lines).

3.2. Characterizing Road Networks

Rather than merely describing the shape of different road networks, we characterize them using centrality and centralization measures from social network analysis and connectivity measures from planar network analysis, as defined in **Table 1**. Centrality indices such as degree, betweenness, and closeness are node-level measures showing the centrality position of a node within a network, while our purpose is to create network-level indices using node-level indices as input. Below, we will first define node-level centrality indices and then use them to define network-level indices.

For a network or graph $G(V, E)$, with nodes V and edges E , the degree centrality of a node, $d(v)$, indicates the total number of direct connections (*i.e.*, the number of directly adjacent nodes) each node has in a network. The betweenness centrality of a node, $b(v)$, refers to the sum of the ratio of the number of shortest paths that pass through a node to the total number of shortest paths among node pairs in the network, as expressed in Equation (1):

$$b(v) = \sum_{v \neq i, j} \frac{p(v)_{ij}}{p_{ij}} \quad (1)$$

where $p(v)_{ij}$ is the number of the shortest paths between node-pair i and j which passes through node v , and p_{ij} is all shortest paths between node-pair i and j in a network [34] [35].

The closeness centrality of a node, $c(v)$, is the reciprocal of the farness, which is defined as the total steps (topological lengths) between a node and all other nodes in a network [34] [35], as expressed in Equation (2):

$$c(v) = \frac{1}{\sum_u d(v, u)} \quad (2)$$

where $d(v, u)$ is the topological length between nodes v and u .

A centralization index, as seen in Equation (3), is a network-level index created from a centrality index using the sum of the absolute deviations of the centrality in the graph, as seen in the numerator. This sum is normalized by the maximum theoretical value of the sum of deviations, as seen in the denominator [36] [37].

$$\text{Centralization} = \frac{\sum_u |\text{max. centrality} - \text{centrality}(u)|}{\text{maximum theoretical value}} \quad (3)$$

When a centrality measure takes on the value of the degree, betweenness, or closeness centrality, various centralization measures are of no unit and between 0 and 1, as seen in **Table 1**.

A centralization index measures the distribution of a centrality index within a network. A high value indicates an uneven distribution with a small number of dominant nodes playing a centralizing role. The dominant nodes often occupy a central position in the network with higher centralities than other nodes. In contrast, a low value means a relatively even distribution of centrality among nodes, and no node stands out over the others.

Table 1. Definitions of network centralization and planar network connectivity.

Index	Definition	Comment
Degree	The number of connections a node has with directly adjacent neighbors	A higher value means a node has more connections
Degree centralization	Ratio of the sum of absolute degree deviation to the maximum theoretical value (0 to 1)	A high value indicates uneven distribution of degrees; a small number of nodes are dominant
Betweenness centralization	Ratio of the sum of absolute betweenness deviations to the maximum theoretical value (0 to 1)	A high value indicates that a few top nodes in the network keep the network connected
Closeness centralization	Ratio of the sum of absolute closeness deviations to the maximum theoretical value (0 to 1)	A high value indicates the network is closely connected and integrated
Algebraic connectivity	The second smallest eigenvalue of the Laplacian matrix of a network	A higher value indicates the network is more connected and integrated
Number of cycles	Number of closed circuits in a network	A higher value indicates a more connected network
α	Ratio of the number of closed circuits to the maximum possible	A higher value indicates a more connected network
β	The number of edges over the number of nodes in a network; edges per node	A higher value indicates more roads per node
γ	Ratio of the number of edges over the maximum number possible	A higher value indicates a more connected network

While centralization measures indicate the distribution of centrality within a network, connectivity measures indicate the connectedness among parts within a network. In spatial analysis, where the networks are road networks, planar network indices are often used to characterize the topological structure of transport networks [7]. A few straightforward measures are used here. The alpha index and the gamma index are the ratios of the existing circuits or edges, respectively, to the maximum possible. The beta index is the number of edges per node. A Laplacian matrix is used to express a graph and show the difference between its adjacency and degree matrices. The second smallest eigenvalue of the Laplacian matrix is called Algebraic connectivity. For connectivity indices listed in **Table 1**, a higher value means a more connected and integrated network.

3.3. Performing Location-Allocation Models

The location-allocation model we use is a stylized classroom teaching model which involves setting up four electronic recycling item collection centers (facilities) to serve 49 local communities (demand points) in the study area. Specifically, for graph $G(V, E)$, a subset of vertices $I \in V$ is defined as the demand points. The problem can be expressed as

$$\text{Min}Z = \sum_{i \in I} \sum_{j \in V} d_{ij} w_{ij} \quad (4)$$

$$\sum_{j \in V} w_{ij} = 1 \quad i \in I \quad (5)$$

Subject to

$$\sum_{i \in I} q_i w_{ij} \leq Q_{y_j} \quad j \in V \quad (6)$$

$$\sum_{j \in V} y_j = p \quad j \in V \quad (7)$$

$$x_{ij}, y_j \in \{0, 1\}, w_{ij} \in \{0, 1\} \quad i \in I, j \in V \quad (8)$$

where Z is the total distance-based travel cost; d_{ij} is the distance on the network between demand vertex i and supply vertex j ; w_{ij} is an allocating variable equal to 1 if demand at i is allocated to j or 0 if not; q_i is the quantity of demand at i ; y_j is a locating variable equal to 1 if a facility is located at j or 0 if not; Q_{y_j} is the facility capacity at j ; and p is the number of facilities. Equation (5) guarantees each demand point is allocated to only one facility. Equation (6) makes sure the total demand allocated to j is within the capacity at j ; Equation (7) sets the total number of facilities to p . Equations from (4) through (8) represent the typical set-up of a capacitated p -median problem, as presented in [11] [12].

3.4. Controlling for Confounding Factors

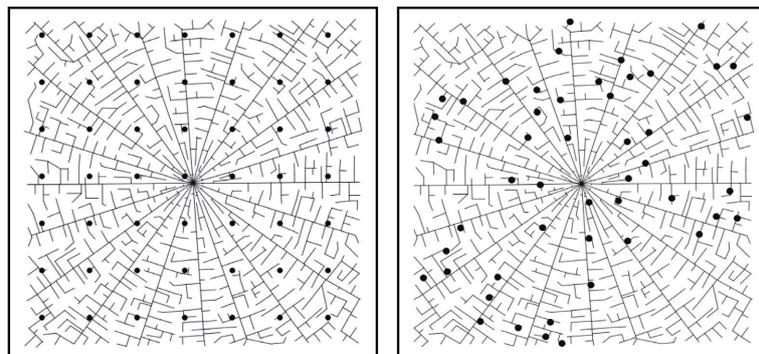
Since our purpose is to identify the impacts of road networks on facility location modeling, possible confounding factors need to be controlled for. As discussed earlier, a basic distinction in the location-allocation problem is between non-capacitated and capacitated models [38]. Since location-allocation modeling is ex-

plicitly spatial, it is significantly impacted by information related to locations and distribution [39]. Xie and Levinson [40] and Rodrigue [7] recognize roads of different qualities and importances in shaping traffic flows. Zhao *et al.* [31] incorporate such information in location-allocation modeling. All these indicate possible confounding factors. To control for these factors, we adopt four types of control factors in the modeling, each with two levels, as shown in **Table 2**.

Table 2. Confounding factors are to be controlled for modeling.

	A: capacity	B: road types	C: demand locations	D: demand quantities
Level 1	Non-capacitated	All regular roads	Regularly distributed	Constant at all sites
Level 2	Capacitated	With corridors	Randomly distributed	Varying at all sites

Implementing control factor A means that we run non-capacitated models as well as capacitated models. For non-capacitated models, the constraint value in Equation (6) is set at a positive infinity. As for capacitated models, facility capacity constraints in Equation (6) are randomly set for facilities, but the total capacity from all facilities equals the total demand. To implement control factor B, maps in **Figure 1** are used to distinguish networks with and without corridor roads. We designate corridors for roads close to the center of each road network. In running models with corridors, nodes on corridors have a 50% higher probability of being selected in routing. To incorporate control factor C, maps in **Figure 2** are used. While maps on the left show demand locations in a regular pattern, maps on the right show demand locations at randomly chosen locations. The 49 demand points are chosen since a 7-by-7 lattice of demand points leads to nearly equal spacing between all demand points, contributing to a regularly distributed distribution, as shown in the maps on the left of **Figure 2**. Finally, implementing control factor D involves setting up the demand quantity for demand points either as a constant or as random values. In either case, the total demand from 49 demand points is the same.



(a)

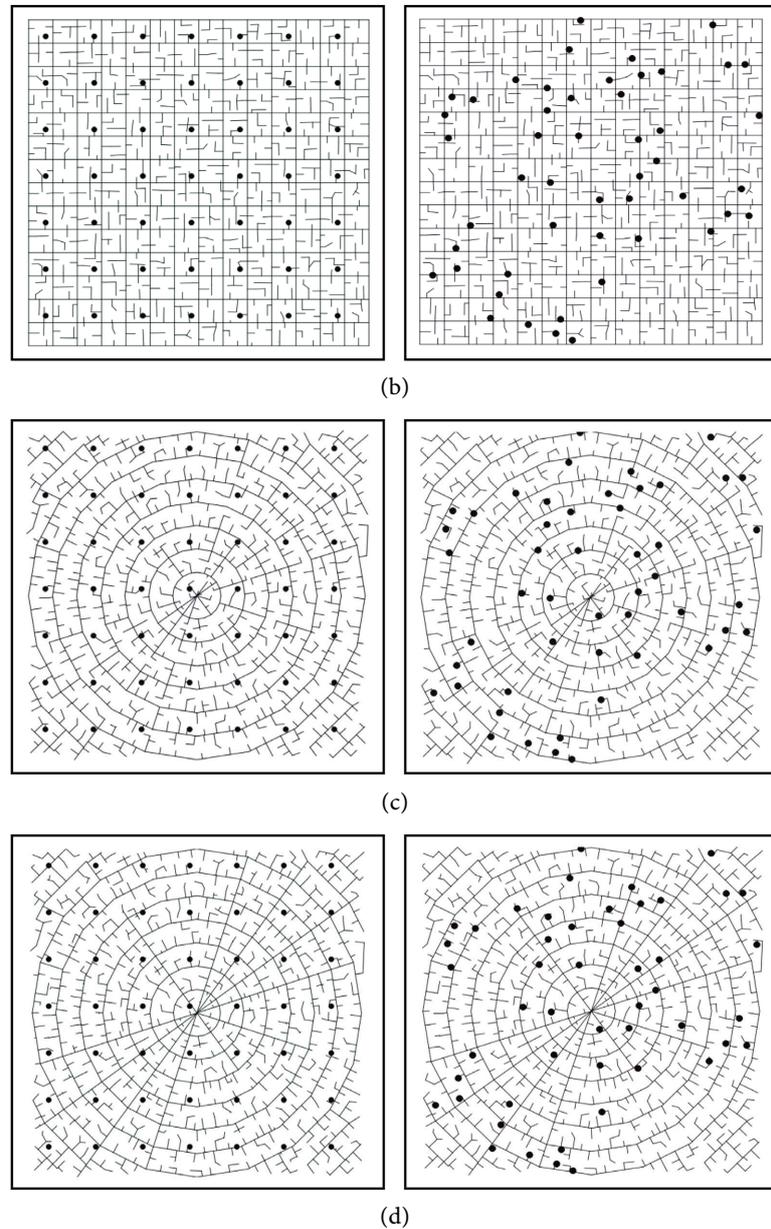


Figure 2. (a) The radial network with regularly (left) and randomly (right) distributed demand points; (b) The grid network with regularly (left) and randomly (right) distributed demand points; (c) The ring network with regularly (left) and randomly (right) distributed demand points; (d) The ring-radial network with regularly (left) and randomly (right) distributed demand points.

3.5. Spatial Analysis

In analyzing the modeling result, we will first conduct a spatial analysis by examining how facilities are connected with demand points geographically. Location-allocation modeling is inherently spatial. The spatial pattern helps reveal the nature of the solutions and, most importantly, the nature of the road networks that give rise to those solutions. In analyzing spatial patterns of the solutions associated with different road networks, attention is focused on the pattern that is the result

of and/or consistent with the network topological characteristics.

3.6. Five-Way ANOVA

We use a heuristic routing algorithm that generates local solutions with certain randomness. To avoid basing our analysis on a random result, we obtain a large solution sample by running model variants in a large number of iterations. Specifically, one model may involve the radial network with level 1 of the four confounding factors. Another model may use the grid network with level 1 of control factors A and B and level 2 of control factors C and D. With four road networks and four control factors each having two levels, there is a total of sixty-four model types (4×2^2). While detailed combinations of control factors and road networks will be discussed later, we run each model type 100 iterations and use the 6400 results in a five-way analysis of variance (ANOVA) to ascertain the role of networks when all factors are controlled for.

We run models using R and NetLogo. R has the advantage of running optimization models fast, while NetLogo allows for real-time visualizations, which provide quick feedback to revise and improve the model. The routing is performed using an algorithm akin to the tabu algorithm [41]. Following Hakimi [42], a node can be both a demand and a facility location.

4. Results

4.1. Topological Characteristics of Networks

Before discussing the modeling result, we present the topological characteristics of the four road networks in **Table 3** as a background for interpreting the results later. The four networks are essentially of the same size in terms of the number of nodes and edges. The radial network has the highest degree centrality and the most uneven distribution of degree and betweenness centralities, with high degree and betweenness centralizations. While the ring network seems to have the second largest uneven centralization, its centralization in closeness is the least, indicating a weak connectivity. The grid and ring-radial networks have generally low centrality and distributed centrality (low centralization). As for connectivity, it is nearly the opposite, with the grid network being the most connected, followed by the ring-radial network. The least connected is the radial network, followed by the ring network.

The topological characteristics of different networks can be largely attributed to their network structure. The high centrality of the radial network is mainly due to its long radial lines extending from the network center outward. However, there are no ring roads that integrate these radial lines into closed circuits. Instead, these radial lines only connect to each other at the network center. The high centrality, uneven centrality distribution, and poor connectivity are anticipated results. The ring network has many ring roads. However, there is seldom any long radial line that connects with ring roads, forming closed circuits, resulting in its poor connectivity. Its high degree of centrality is largely due to a few short radial lines con-

necting at the network center. Interestingly, combining the radial and ring networks into the ring-radial network turns their disadvantages into advantages. Now, the long radial lines and ring roads crisscross each other, forming many closed circuits, which significantly increases connectivity. Finally, the many closed circuits in the grid network are largely responsible for its distributed centrality (or relatively even degrees across the network) and the highest connectivity.

Table 3. Topological attributes of the four prototypes of road networks.

Index	Radial	Grid	Ring	Ring-radial
Number of nodes	1522	1541	1505	1565
Number of edges	1521	1736	1541	1638
Max. degree	20	4	12	8
Degree centralization	0.012	0.001	0.007	0.004
Betweenness centralization	0.933	0.178	0.200	0.132
Closeness centralization	0.521	0.052	0.030	0.053
Algebraic connectivity	0.000 723	0.003 293	0.000 561	0.000 807
Number of cycles	0	196	33	70
α	0.000	0.064	0.012	0.023
β	0.999	1.126	1.024	1.047
γ	0.334	0.376	0.342	0.349

4.2. Maps Showing Connections from a Single Solution

Since we run each model type 100 times and each model contains four optimal facility locations, this will create 400 facility-demand connections. To understand the pattern from a map showing 400 connections, we first show and explain maps that contain solutions from a single model (**Figure 3**). Every map in **Figure 3** contains four facilities from one model, each serving a number of demand locations. Maps on the left use straight desire-lines as typically seen in location-allocation modeling to connect supply and demand points. While desire-lines show explicit facility and demand locations, they do not reveal the actual road paths through which supply and demand points are connected. Maps on the right show these road paths.

Maps in **Figure 3** provide clues to the unique facility-demand connecting patterns associated with different road systems. For example, the radial network tends to have service facilities at the center of the network serving demand sites through long radial roads. On the other hand, the grid network shows clusters of short and localized connections with short road paths. As for the ring network, facilities tend to serve demand points through long ring roads. A similar pattern is also seen in the ring-radial network, but with shorter road paths than the model using the ring network.

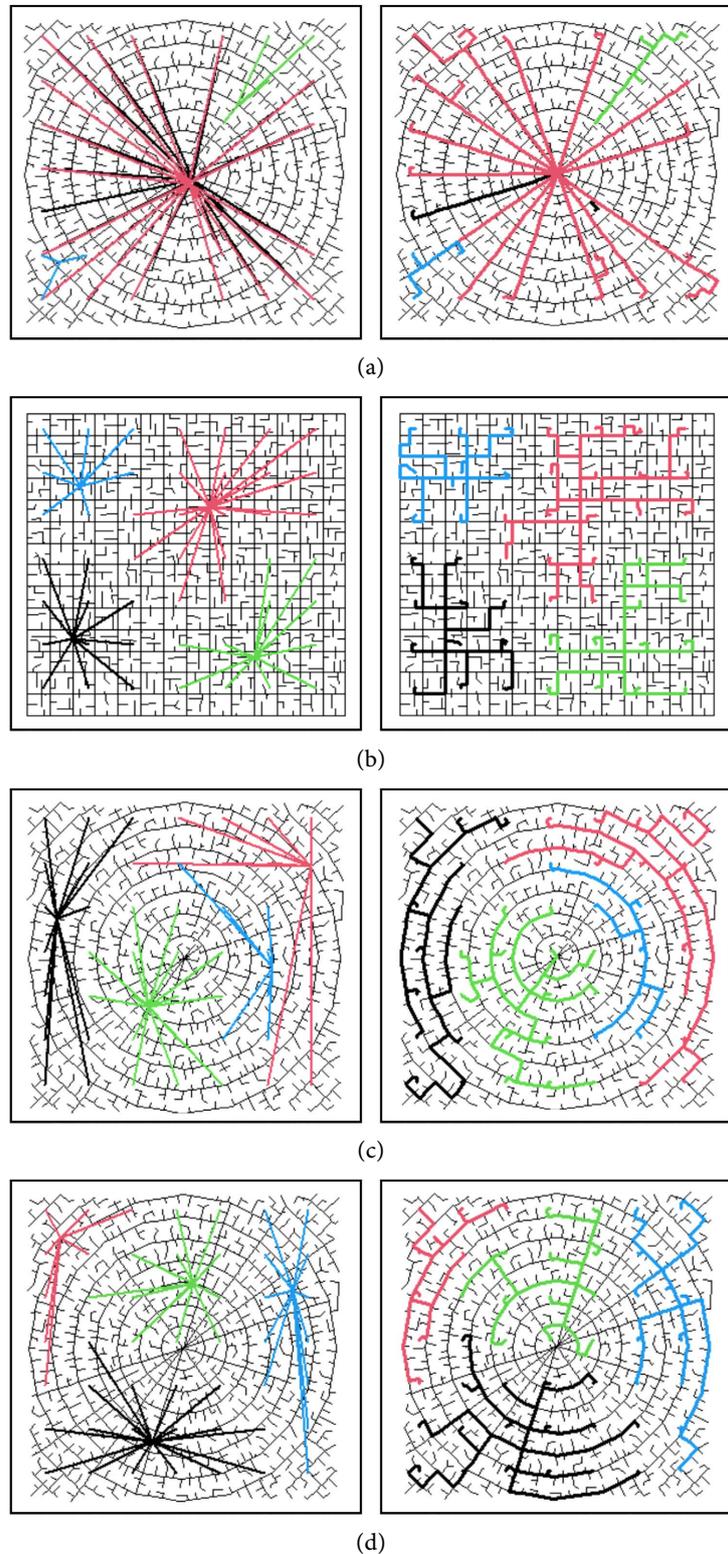
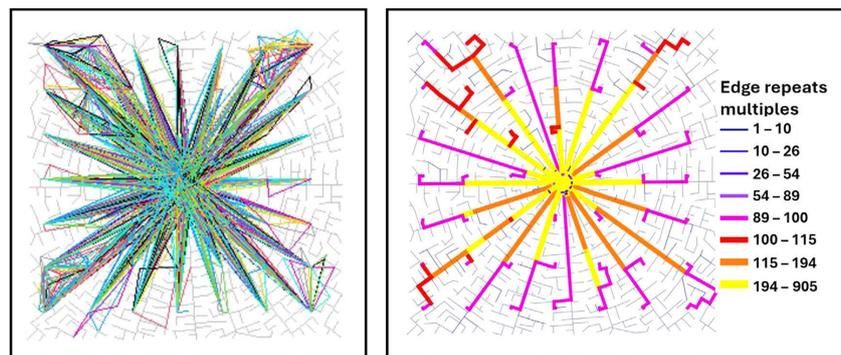


Figure 3. (a) Solutions from one model using the radial network: desire-lines (left) and connecting paths (right); (b) Solutions from one model using the grid network: desire-lines (left) and connecting paths (right); (c) Solutions from one model using the ring network: desire-lines (left) and connecting paths (right); (d) Solutions from one model using the ring-radial network: desire-lines (left) and connecting paths (right).

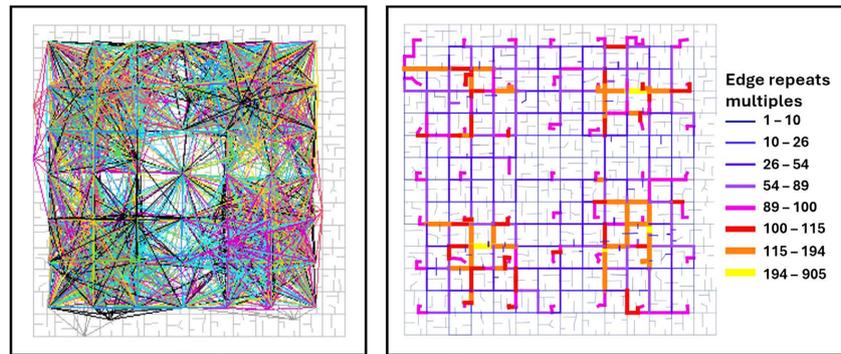
4.3. Models with Regularly Distributed Demand Locations

Now we present maps containing 100 runs of a model to illustrate their spatial patterns (**Figure 4**). These are capacitated models (Factor A level 2) using regularly distributed demand locations (Factor C level 1; maps on the left in **Figure 1**) without distinction of corridors from other roads (Factor B level 1; maps on the left in **Figure 2**), and with an equal demand quantity in all demand sites (Factor D level 1).

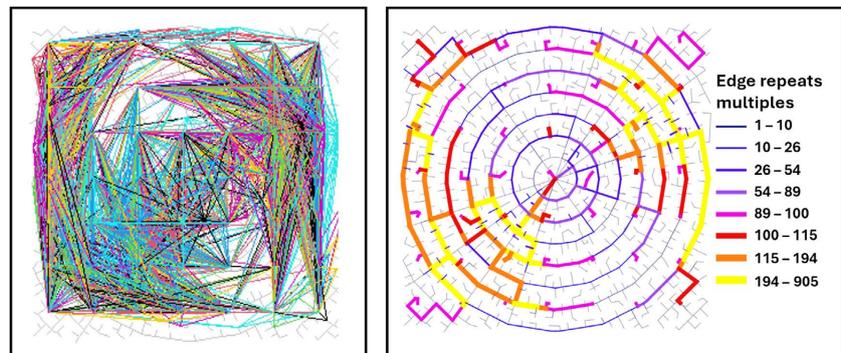
In maps in **Figure 4**, desire-lines (maps on the left) for the radial and ring networks tend to connect facility demand points over long distances, along the radial roads in the former, and along the ring roads in the latter. For the grid network, desire-lines are often short and localized lines. As for the ring-radial network, there are both long and short, localized desire-lines.



(a)



(b)



(c)

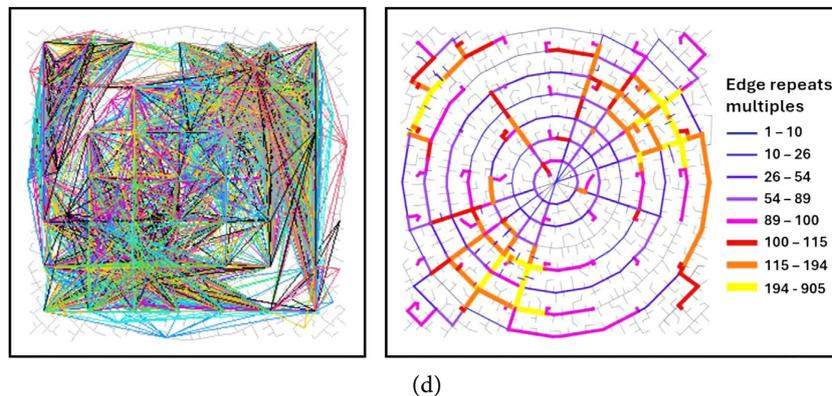


Figure 4. (a) Solution desire-lines (left) and edge repeat multiples (right) from 100 iterations of capacitated models on the radial network with regularly distributed demand locations without corridors and with constant demand at all demand locations; (b) Solution desire-lines (left) and edge repeat multiples (right) from 100 iterations of capacitated models on the grid network with regularly distributed demand locations without corridors and with a constant demand at all demand locations; (c) Solution desire-lines (left) and edge repeat multiples (right) from 100 iterations of capacitated models on the ring network with regularly distributed demand locations without corridors and with a constant demand at all demand locations; (d) Solution desire-lines (left) and edge repeat multiples (right) from 100 iterations of capacitated models on the ring-radial network with regularly distributed demand locations without corridors and with a constant demand at all demand locations.

Table 4 summarizes characteristics of solutions from models using different networks. The *total travelled distance* is the sum of the solution distance over 100 models. When divided by 100, it gives the *mean solution distance*. This is the average minimum distance and measures the optimality associated with different networks. In a model, a road path can be selected in routing many times if it is in a critical location, or only one or two times if it is in a marginal location. Regardless of how often they are selected in routing, the roads that are used at least once are added as the *travelled path distance*. For example, suppose roads a, b, and c have lengths 5, 6, 7 (km), respectively, and road a is used two times and road b 3 times, but road c is never set foot on. In this case, the travelled path distance is 11, and the total travelled distance is 28.

Table 4. Characteristics of capacitated models with regularly distributed demand locations.

Net-work	Mean solution distance (km)	Total travelled distance (km)	Travelled path distance (km)	Network path repeats multiple	Cumulative numbers of edges travelled	Numbers of edges travelled	Network edge repeats multiple	Max. edge repeat multiple	Average connecting path length (km)
Radial	2736.4	273 642.5	2834.5	97	69,890	836	834	905	55.9
Grid	2112.0	211 204.0	6534.2	32	51,322	1480	35	212	43.1
Ring	2760.9	276 090.6	4263.4	65	81,412	1236	66	450	56.3
Ring-radial	2355.8	235 578.9	4774.2	49	68,937	1345	51	309	48.1

The *network path repeats multiple* is the ratio of the total travelled distance to the travelled path distance. This multiple shows how often a road is travelled. A value of 97 (for the radial network) means that each unit length of a road is used 97 times on average. Similarly, a network edge can be selected in routing 0 multiple times. Summing up this value over all edges produces the *cumulative number of edges travelled*. The *number of edges travelled* is the number of edges that are used in routing at least once. The ratio of the two values is the *network edge repeats multiple*, indicating how often an edge is used on average in modeling. In the above example of roads or edges a, b, c, the number of edges being travelled on is 2 (edges a and b). The cumulative number of edges being travelled on is 5. The network edge repeats multiple is 2.5. The edge repeats multiple can also be calculated for each edge. The edge that is selected most times during routing has the *maximum edge repeats multiple*. Finally, the *average connecting path length* is the total travelled distance divided by 4900 (Since the number of demand sites is 49, there are 4900 facility-demand paths by which these sites are served in 100 model iterations). This tells the average distance by which a facility serves one demand location.

As seen in **Table 4**, the grid network produces the best model in terms of the shortest solution distance (2112 km), followed by the ring-radial network (2356 km). The worst result occurs for the ring-radial (2761 km) network, followed by the radial network (2736 km). **Table 4** also shows that both the radial and ring networks have higher network path and edge repeat multiples (97 and 65, respectively) than the grid and ring-radial networks (32 and 49, respectively). The maximum edge repeats multiples for the radial and ring networks are much higher than those for the grid and ring-radial networks, too.

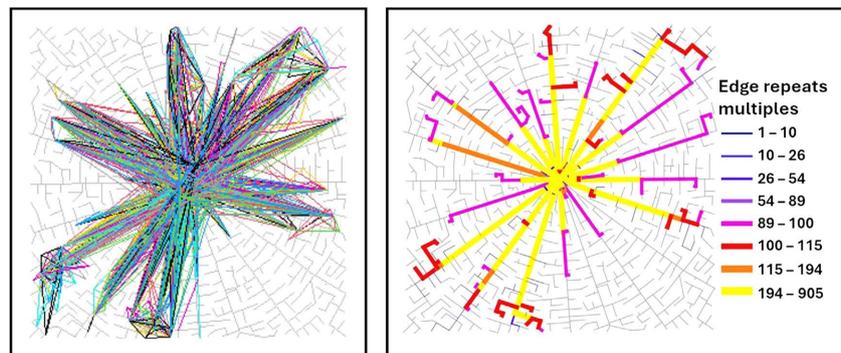
Maps on the right in **Figure 4** show spatial patterns of the edge repeat multiples. The radial and ring networks tend to have long stretches of roads with high edge repeat multiples. On the other hand, the grid network has high edge repeats multiples only in small local areas. As for the ring-radial network, it does have high edge repeat multiples along radial and ring roads, but the lengths of these are significantly less than those for either the radial or the ring network. There seems to be a loose relationship between the edge repeat multiples and their importance in modeling. For example, the grid network has one-third of the edges with edge repeats multiples at 50 or fewer. These edges account for 50% of the total travelled distance in the model. For the other three networks, the corresponding values are 25% and 37% for the ring-radial network, 17% and 37% for the ring network, and 16% and 18% for the radial network. It seems that for a given edge repeat multiple, more edges are used and longer distances are travelled by well-connected networks than poorly connected ones. Well-connected networks also have shorter average connecting path distances than poorly connected ones. As shown in **Table 4**, the shortest average connecting path distance is from the grid network, followed by the ring-radial network, and the longest is from the ring-radial network, followed by the radial network. In general, spatial patterns of supply-demand con-

necting paths, as shown in **Figure 4**, are consistent with information revealed in **Table 4**.

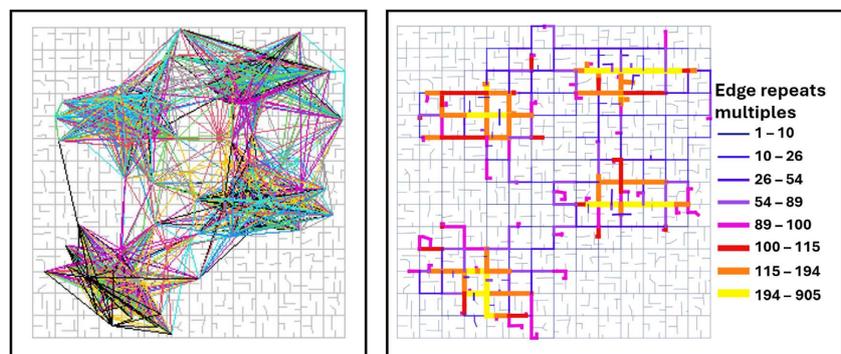
4.4. Models with Randomly Distributed Demand Locations

Maps in **Figure 5** show results from models using randomly distributed demand locations (Control factor C level 2 or maps on the right in **Figure 2**) and keeping other conditions the same as previously presented models. Specifically, Factors A, B, and D remain at level 2, 1, and 1, but Factor C is set at level 2. The spatial connection patterns are similar to those in maps in **Figure 4**. Briefly, for the radial and ring networks, desire-lines tend to have long connections, along the radial roads for the former and the ring roads for the latter. Their connecting paths follow long radial or ring roads with high edge repeat multiples. The grid network has mostly short desire-lines and connecting paths with high edge repeat multiples occurring only in local areas. The ring-radial network has a mix of long and short desire-lines, but its roads with high edge repeat multiples are much shorter than those for the radial and ring networks.

One noticeable difference is that desire-lines and connecting paths on maps in **Figure 5** are aligned in a northeast-southwestern direction consistent with randomly distributed demand point locations, rather than spreading apart and stretching across the entire network as seen on maps in **Figure 4**. This explains the shorter optimal solutions in models using randomly distributed demand locations, as will be discussed below.



(a)



(b)

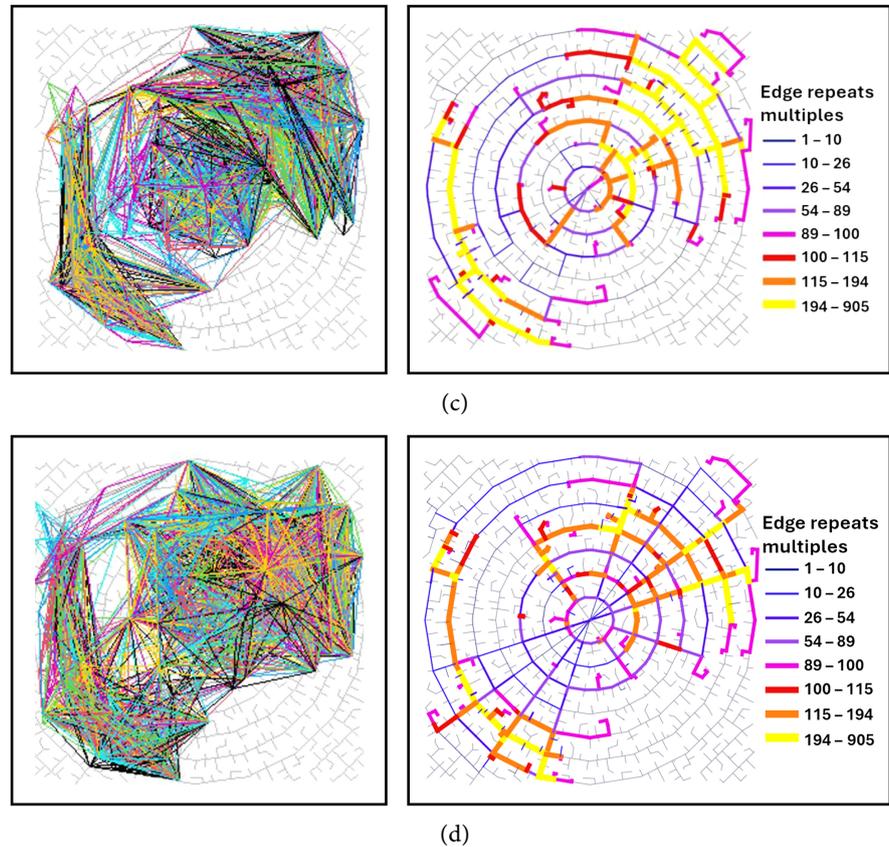


Figure 5. (a) Solution desire-lines (left) and connecting paths and frequencies (right) from 100 iterations of capacitated models on the radial network with randomly distributed demand locations without corridors and with a constant demand quantity at all demand locations; (b) Solution desire-lines (left) and connecting paths and frequencies (right) from 100 iterations of capacitated models on the grid network with randomly distributed demand locations without corridors and with a constant demand quantity at all demand locations; (c) Solution desire-lines (left) and connecting paths and frequencies (right) from 100 iterations of capacitated models on the ring network with randomly distributed demand locations without corridors and with a constant demand quantity at all demand locations; (d) Solution desire-lines (left) and connecting paths and frequencies (right) from 100 iterations of capacitated models on the ring-radial network with randomly distributed demand locations without corridors and with a constant demand quantity at all demand locations.

Table 5 summarizes characteristics of model solutions using the four networks. The overall pattern is consistent with the models being characterized in **Table 4**. Specifically, the grid network has the most optimal solution with the minimum solution distance, followed by the ring-radial network. The worst result comes from the ring network, followed by the radial network. In addition, the network path and edge repeat multiples are the highest for the radial and ring networks, and the corresponding values for the grid and ring-radial networks are significantly lower. This point is reinforced by the maximum edge repeats multiple for the four networks. Finally, the average connecting path lengths of the four networks follow the same order as seen for models with regularly distributed demand

locations (Table 4). As mentioned earlier, the lower solution distances than models reported in Table 4 are the result of using random demand locations that are closer to each other than the regularly distributed demand locations.

Table 5. Characteristics of the capacitated models with random demand points and corridors.

Net-work	Mean solution distance (km)	Total distance travelled (km)	Path distance travelled (km)	Network path repeats multiple	Cumulative numbers of edges travelled	Numbers of edges travelled	Network edge repeats multiple	Max. edge repeat multiple	Average connecting path length (km)
Radial	2488.4	248 838.9	2315.1	107	63,884	691	92	595	50.8
Grid	1703.1	170 311.4	4257.7	40	43,596	1010	43	309	34.8
Ring	2532.4	253 238.7	3664.0	69	73,064	1061	69	602	51.7
Ring-radial	2036.5	203 649.8	4392.6	46	59,842	1247	48	427	41.6

The different performances of the four networks can be largely attributed to their network structure, which ultimately determines network topological characteristics. The long radial or ring roads in the radial or ring networks do not criss-cross roads in other directions to form sufficient closed circuits. In such a structure, the long roads essentially act as constraints, forcing the routing repeatedly through limited numbers of edges or paths in critical positions, leading to high edge and path repeats multiples. In contrast, closed circuits in the grids and ring-radial networks provide flexibility through shortcuts, branch roads, and direct paths to access nearby nodes.

The results presented so far incorporate both levels of control factor C while keeping other control factors at the same level. These controls are visualized in Figure 6, where control factor C takes both “regular” and “random” levels while other control factors remain at one level. Ranges of the solution values are depicted by black points forming vertical lines, and the corresponding average values are shown as the blue horizontal lines. In both levels of control factor C, the lowest average solution distance occurs in the grid network, followed by the ring-radial network, and the highest average solution value is from the ring network, followed by the radial network.

Models incorporating both levels of control factor C demonstrate the impact of the demand point locations as a confounding factor. Specifically, when switching from regularly to randomly distributed demand locations, the solution distances are lower, and this effect is consistent across all networks. This is equivalent to linear models where the inclusion of a factor variable changes the intercept of the model. A critical point is that while the blue lines in Figure 6 for models with random demand points are lower than those for models with regular demand points, are these differences statistically significant? In other words, do these differences occur by chance from random samples, or are there actually such differences in the statistical population? This question will be addressed in the next section.

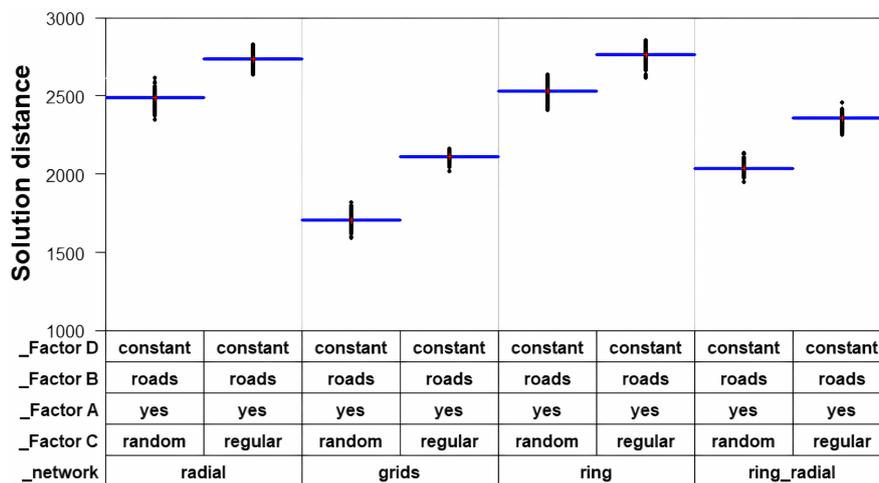


Figure 6. Distributions and the means of solution distances from models using different networks with both levels of control factor C, while keeping other control factors at the same level.

A careful comparison of the mean solution distances at both levels of factor C reveals that the value for the radial network in **Table 5** is 248 (km) or 9% lower than that in **Table 4**. The corresponding values for other networks are 409 (km) or 19% lower for the grid network, 229 (km) or 8% lower for the ring network, and 319 (km) or 14% lower for the ring-radial network. Modeling on the same set of new demand locations leads to different reductions in solution distance. This implies the interaction effects between networks and control factor C. The magnitude of change is larger for better-connected networks than for poorly connected networks. Using the same linear model analogy as mentioned above, this is equivalent to varying intercept changes from different levels of a (network) factor variable. These issues will be further explored below.

5. Discussion

5.1. Solution Means of All Models

As stated earlier, we incorporate four control factors in modeling, and each confounding factor has two levels. This leads to a total of 16 combinations of the control factor levels, as shown in the first five columns of **Table 6**. Modeling with each combination of factors on four networks amounts to a total of 64 model variants. The last four columns give the 64 mean solution values, each averaging over 100 model runs. In the table, each mean value corresponds to a combination of control factor levels and networks.

Figure 7 graphs the 64 mean values in **Table 6** as red crosses. Each mean value lies within a range of solution distance values indicated by the black vertical lines. Each mean solution and the range of solution values correspond to a combination of control factor levels and a network. As shown in **Table 2**, factor B contains level 1 “all regular roads” and level 2 “with corridors”. Similarly, factor D contains level 1 “constant at all sites” and level 2 “varying at all sites”. As can be seen, under each

combination of control factors, the grid network generates the lowest distance, followed by the ring-radial network; and the ring network produces the worst result, followed by the radial network.

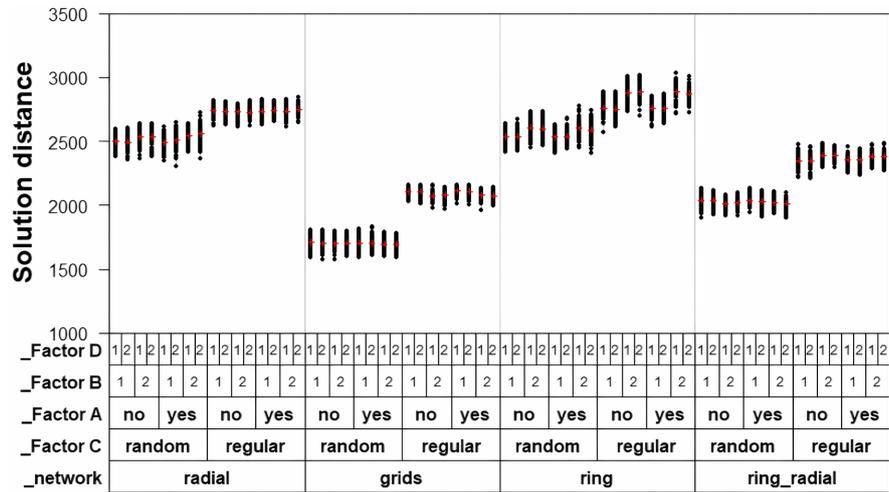


Figure 7. Mean solution distances from road networks under all control factors: factor B: 1 = all regular roads; 2 = with corridors; control factor D: 1 = constant at all sites; 2 = varying at all sites.

Table 6. The mean model solution distances (km) on four networks with all control factors.

Type	Factor A	Factor B	Factor C	Factor D	Radial	Grid	Ring	Ring-radial
1	Non-capacitated	No corridors	Regular	Varying demand	2733.3	2108.5	2753.9	2349.4
2			Locations	Constant demand	2738.0	2106.3	2761.8	2349.6
3	Capacitated	No corridors		Varying demand	2745.9	2109.4	2756.2	2353.8
4				Constant demand	2736.4	2112.0	2760.9	2355.8
5	Non-capacitated	With corridors		Varying demand	2725.6	2079.9	2887.1	2391.9
6				Constant demand	2729.8	2074.4	2879.3	2393.2
7	Capacitated	With corridors		Varying demand	2753.5	2073.5	2872.9	2381.9
8				Constant demand	2735.2	2077.0	2886.5	2383.2
9	Non-capacitated	No corridors	Random	Varying demand	2494.1	1701.2	2536.9	2038.6
10			Locations	Constant demand	2498.0	1711.2	2538.8	2038.6
11	Capacitated	No corridors		Varying demand	2506.8	1702.0	2535.9	2030.2
12				Constant demand	2488.4	1703.1	2532.4	2036.5
13	Non-capacitated	With corridors		Varying demand	2535.8	1704.2	2597.7	2017.6
14				Constant demand	2532.7	1701.4	2606.4	2009.6
15	Capacitated	With corridors		Varying demand	2557.9	1696.1	2588.2	2012.3
16				Constant demand	2542.1	1697.6	2601.2	2019.6

5.2. Confounding Effects of Control Factors

As discussed previously when comparing **Table 4** and **Table 5**, models with random demand locations (factor C level 2) improve over those with regular demand locations (factor C level 1) by reducing the average optimal distance by 248 (km) or 9% for the radial network, 409 (km) or 19% for the grid network, 229 (km) 8% for the ring network, and 319 (km) or 14% for the ring-radial network. The same results can be obtained from **Table 6** by subtracting values in Type 12 from those in Type 4. These are the effects of controlling for both levels of factor C while setting factor A at level 2, and factors B and D at level 1. Similarly, taking the absolute values of differences between Types 1 and 9 in **Table 6** gives the effects of controlling for both levels of factor C while keeping factor A at level 1, and factors B and D at level 1. Doing the same between Types 2 and 10, 3 and 11, 5 and 13, 6 and 14, 7 and 15, and 8 and 16 generates the effects of controlling both levels of factor C while keeping factors A, B, and D at other levels. Averaging these values produces values in row 3 of **Table 7** (outside the parentheses). These are the average effects of controlling for factor C using different networks. Other values in **Table 7** are obtained by finding (absolute) differences between Types 1 and 3, 5 and 7, 9 and 11, 13 and 15, 2 and 4, 6 and 8, 10 and 12, 14 and 16 for the effect of factor A; between Types 1 and 5, 2 and 6, 3 and 7, 4 and 8, 9 and 13, 10 and 14, 11 and 15, and 12 and 16 for the effect of factor B; and between Types 1 and 2, 3 and 4, 5 and 6, 7 and 8, 9 and 10, 11 and 12, 13 and 14, and 15 and 16 for the effect of factor D.

Table 7. Confounding effects of control factors (%)*.

Control factor	Radial	Grid	Ring	Ring-radial	Row Mean
A	0.48	0.25	0.21	0.32	0.32
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
B	1.02	0.96	3.46	1.27	1.68
	(0.91)	(0.78)	(3.46)	(1.06)	(1.55)
C	7.95	18.66	8.94	14.53	12.52
	(7.95)	(18.66)	(8.94)	(14.53)	(12.52)
D	0.37	0.20	0.28	0.18	0.25
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Column mean	2.46	5.01	3.22	4.07	
	(2.22)	(4.86)	(3.10)	(3.90)	

*Values outside parentheses are descriptive statistics from the result sample; values inside parentheses incorporate significance tests (p -value = 0.01) from the analysis of variance (see below).

Values outside parentheses in **Table 7** are descriptive statistics obtained by getting the between-level differences of a factor and then adding them up, as described above. Factor C is the most influential. Models with randomly distributed demand locations improve results by 7.95% on average over regularly distributed demand locations for the radial network, 18.66% for the grid network, 8.94% for the ring network, and 14.53% for the ring-radial network. The average over the four networks is 12.52%. Values inside parentheses incorporate significance tests concerning whether the pairwise differences are statistically significant. In cases where they are, values are the same as those outside parentheses, as is the case for factor C. The large impact of control factor C is the main reason for the result section focusing on models that control for both levels of factor C.

Compared with factor C, the confounding effects for control factors A, B, and D are small, especially for factors A and D, where all descriptive values are below 1%. In significance tests, all these small pairwise differences turn out to be insignificant, leading to 0 values inside parentheses. In other words, capacitated and non-capacitated models do not produce significantly different results, as in the case for models using constant and varying demand quantities. Models with and without corridors produce results that, on average, are different by 1.32% for all networks based on descriptive statistics. The significance tests show that some pairwise differences are statistically significant while others are not. In the latter case, the pairwise differences are 0. This leads to values inside the parentheses not being the same as those outside. Details of the significance tests on factor B can be found in **Table 12** and the relevant discussion below. Finally, looking at the column averages, the impact of the cofounding factors is the largest on the grid network at 5.01% or 4.86%, and the smallest for the radial network at 2.46% or 2.22%.

5.3. Five-Way ANOVA

A five-way analysis of variance is performed using the 6400 solution distances (64 model variants from four networks, each with 16 combinations of four control factors, each model variant running 100 iterations) as the dependent variable and network types and control factors as independent variables. In ANOVA, one type of statistical test is the test of the main effects. This refers to the test of statistical differences in the mean values between levels of a variable. In the context of this research, four road networks are at different levels of the variable network. Out of the 6400 solution values, each road network has 1600 values. The main effect test analyzes whether the mean values are significantly different between at least one pair of networks (e.g., radial vs. grid or ring vs. ring-radial, etc.) without controlling for confounding factors. For a confounding factor, the main effect test analyzes whether the mean values between the two levels of a factor (3200 values each) are significantly different. The results for the main effect tests are shown in **Table 8**. Networks as a factor have the largest mean-square value and the F value, rendering the mean values among different networks statistically significant. Control

factors B and C are also statistically significant, implying significant differences between levels of these factors. The main effects for factors A and D are not statistically significant, a result consistent with descriptive analysis in the context of **Table 7**, where different levels of the two factors cause less than 1% of differences. In fact, the five-way post-hoc tests on the analysis of variance show that two levels of factor A do not generate statistically significant differences under any combinations of factors B, C, and networks. This is also true for factor D.

Table 8. Main effects of analysis of variance.

Factor	Degrees of freedom	Sum of squares	Mean square	F value	Pr (>F)
Network	3	677,594,985	225,864,995	1.09E+05	<2e-16
A	1	606	606	2.93E-01	0.588 52
B	1	1,054,262	1,054,262	5.09E+02	<2e-16
C	1	145,298,628	145,298,628	7.02E+04	<2e-16
D	1	38	38	1.80E-02	0.891 91

Table 9. Post-hoc tests for the main effects of networks, factor B, and factor C*.

Level comparison	Mean difference	Lower bound	Upper bound
Network: Radial vs. Grid	730.98	726.84	735.11
Network: Ring vs. Grid	796.14	792.00	800.28
Network: Ring_radial vs. Grid	300.26	296.12	304.39
Network: Ring vs. Radial	65.17	61.03	69.30
Network: Ring_radial vs. Radial	-430.72	-434.86	-426.58
Network: Ring_radial vs. Ring	-495.89	-500.02	-491.75
B: roads vs corridors	-25.67	-27.89	-23.44
C: regular vs random	301.35	299.12	303.58

*All mean differences are statistically significant at the 0.001 level.

The main effect test only shows whether there is a significant difference between at least two levels of a variable. The post-hoc test shows which mean value is higher than the other. **Table 9** shows the post-hoc tests for the main effects of networks and control factors B and C. The values in column 2 are the differences in means between levels of different factors. In the case of networks, the mean solution value from models using the radial network is 730.98 (km) higher than that of models using the grid networks. Similarly, the mean distance from the ring

network is 796.14 (km) higher than that from the grid network, and the mean distance from the ring-radial network is 300.26 (km) more than that from the grid network. However, the mean distance from the ring-radial network is 430.72 (km) less than the radial network, and 495.89 (km) less than the ring network. All the differences are statistically significant at the 0.01 level. The mean solution values between different networks follow the order of grid < ring-radial < radial < ring (column 2). The 95% confidence intervals show no overlap between pairwise networks (columns 3 and 4). For factor B, the average solution value from models without corridors is 26 (km) lower than that for models with corridors. As for factor B, the mean solution value from models with regularly distributed demand locations is 301 (km) higher than that for models with randomly distributed demand locations.

Table 10. The post-hoc tests on five-way interaction effects of different networks*.

A: capacitated	B: with corridors	C: random points	D: varying quantity	Radial vs. Grid	Ring vs. Grid	Ring_radial vs. Grid	Ring vs. Radial	Ring_radial vs. Radial	Ring_radial vs. Ring
No	No	No	Yes	624.8	645.4	240.9	(20.6)	-383.9	-404.5
			No	631.8	655.5	243.3	(23.8)	-388.4	-412.2
			Yes	636.5	646.8	244.4	(10.4)	-392.1	-402.4
			No	624.4	648.9	243.7	(24.5)	-380.6	-405.1
Yes	Yes	No	Yes	645.7	807.3	312.1	161.6	-333.7	-495.2
			No	655.5	804.9	318.8	149.4	-336.7	-486.1
			Yes	680.0	799.4	308.4	119.4	-371.5	-490.9
			No	658.1	809.5	306.2	151.4	-351.9	-503.3
No	No	Yes	Yes	792.9	835.7	337.3	42.8	-455.6	-498.4
			No	786.8	827.6	327.5	40.8	-459.3	-500.1
			Yes	804.8	833.9	328.2	29.1	-476.5	-505.6
			No	785.3	829.3	333.4	44.0	-451.9	-495.9
Yes	Yes	Yes	Yes	831.7	893.5	313.4	61.8	-518.2	-580.1
			No	831.3	905.0	308.2	73.7	-523.1	-596.8
			Yes	861.8	892.1	316.2	30.3	-545.6	-575.9
			No	844.5	903.6	322.0	59.1	-522.5	-581.6

*The values show the pairwise solution mean difference between levels of different networks under various combinations of control factors. All mean differences are statistically significant at the 0.001 level, except that values in parentheses are not significant at the 0.05 level.

The above post-hoc tests are only for different levels of the same variable without controlling for other factors. **Table 10** contains the post-hoc tests for five-way interaction effects, which control for all factors under each network. The first four columns contain the same information as in columns 2 to 5 of **Table 6**, which are various combinations of the four control factors, although labeled differently to conserve the column space. The last 6 columns show differences in the mean solution values between pairwise networks. To interpret these values, take the first row which compares the mean values from different networks with a combination of capacitated models (factor A level 2) with no corridors (factor B level 1) serving regularly distributed demand locations (factor C level 1) with a constant demand (factor D level 1). Column 5 shows that the mean value from the radial network is 624.8 (km) higher than that from the grid network; column 6 shows the mean value from the ring network is 645.4 (km) higher than that from the grid network; column 10 shows the mean value from the ring-radial network is 404.5 (km) lower than that from the ring network, etc. Most differences of mean values in **Table 10** are statistically significant at the 0.001 level, except for a few cases (the first 4 values in the Ring vs Radial column) where the differences of means are not significant at the 0.05 level. In general, the mean distance values for different networks show an order of grid < ring-radial < radial < ring. The four exceptions, as pointed out earlier, occur between the radial and ring networks where the mean distances are not significantly different from each other.

Table 11 shows the post-hoc tests on different levels of factor C under the five-way interaction. The values in the table are differences of the mean distance values between models with regularly and randomly distributed locations under various combinations of control factors.

Table 11. The post-hoc tests on different levels of factor C under five-way interactions*.

Factor A	Factor B	Factor D	Radial	Grid	Ring	Ring_radial
Non-capacitated	No corridors	Varying demand	239.2	407.3	217	310.8
		Constant demand	240.1	395.1	223	311
Varying demand		239.1	407.4	220.4	323.6	
Constant demand		248	408.9	228.5	319.3	
Non-capacitated	With corridors	Varying demand	189.8	375.7	289.5	374.3
		Constant demand	197.1	372.9	272.8	383.6
Varying demand		195.5	377.4	284.7	369.6	
Constant demand		193.1	379.4	285.3	363.6	
Capacitated						

*The values show the mean differences between levels of factor (B) under various combinations of other control factors and networks. All mean differences are statistically significant at the 0.01 level.

A, B, D and different networks. For example, as discussed when comparing **Table 4** and **Table 5**, the average solution value from models with regular demand points is 248 (km) higher than models with random demand sites for the radial network, 409 (km) higher for the grid network, etc. All differences are positive and statistically significant at the 0.001 level, indicating models with regularly distributed demand locations are consistently less efficient than those with randomly distributed demand locations. This leads to the same values inside and outside parentheses for factor C in **Table 7**. As mentioned earlier, depending on networks, changing from regularly distributed to randomly distributed demand locations leads to different amounts of model improvement changes, where better-connected networks (e.g., the grid and ring-radial networks) improve more than the poorly connected networks (the radial and ring networks). There are interactions between networks and the demand locations.

Table 8 shows that the main effect of control factor B is statistically significant, and the post-hoc tests on the main effects in **Table 9** show that models without corridors generate a lower average solution distance than models with corridors. However, the post-hoc tests on factor B under the five-way interaction show this is only true for models using the ring network, as shown in **Table 12**. For other networks, the results are mixed, involving positive, negative, or insignificant differences in mean values. Incorporating the insignificant mean differences in calculating the magnitudes of effects of factor B leads to values in parentheses in **Table 7** for factor B. Thus, control factor B does not generate consistent results in location-allocation modeling. Although this result may be of significance in studies about interactions between networks and road types, it is outside the scope of this research.

Table 12. The post-hoc tests on five-way interaction effects of different levels of factor B*.

Factor A	Factor C	Factor D	Radial	Grid	Ring	Ring_radial
Non-capacitated	Regular	Varying demand	(7.7)	28.7	-133.2	-42.5
		Constant demand	(8.2)	31.9	-117.4	-43.6
		Varying demand	(-7.6)	35.9	-116.6	-28.1
		Constant demand	(1.3)	35.0	-125.6	-27.5
Capacitated	Random	Varying demand	-41.7	(-2.9)	-60.8	21.0
		Constant demand	-34.7	(9.7)	-67.6	29.0
		Varying demand	-51.2	(5.9)	-52.3	(17.9)
		Constant demand	-53.7	(5.5)	-68.8	(16.9)

*The values show the mean differences between levels of factor B under various combinations of other control factors and networks. The mean differences are statistically significant at the 0.01 level. Mean differences in parentheses are not statistically significant.

6. Summary and Concluding Remarks

This research investigates the impact of the road network topological structure on facility location modeling. We create the radial, the grid, the ring, and ring-radial networks and characterize them with centrality and centralization indices from social network analysis and connectivity indices from planar network analysis. We perform p -median location-allocation models on different road networks, incorporating four confounding factors: the service capacity, the road type, the demand locations, and the demand quantity. On each network and with every combination of control factor levels, we run a model 100 times. We use the resulting large sample to perform spatial and statistical analyses. The spatial pattern of facility-demand connecting paths for the radial and ring networks shows long distances along the radial or ring roads. The connecting paths for the grid network show short and localized patterns. The ring-radial network contains a mix of long and short local connection patterns. The analysis of variance shows that under any combination of the four control factors, the grid network produces the best result with the lowest optimal solution distance, followed by the ring-radial network. The worst outcome is produced by the ring network, followed by the radial network.

We attribute the results to the topological characteristics of the networks. The best result by the grid network is mostly due to its highest network connectivity and distributed centrality. The worst result from the ring network can be attributed to its low connectivity. The high centrality but low connectivity of the radial leads to a result comparable to that of the ring network. It is interesting that while the radial and ring networks have low connectivity by themselves, blending the two results in a ring-radial network with significantly higher connectivity than the radial and ring networks alone. This makes the ring-radial networks have the second-best performance in modeling.

This research has important policy implications. Building roads has long been used to improve the conditions for local development. This study suggests that improving transportation is more than just building more roads. For communities that intend to attract business investment, efforts should be made to improve the network structure. For communities with limited means, a smart design to improve the road network structure may get results that are just as desirable. As shown in this study, the grid road system is the most connected due to a large number of closed circuits contained within the system. However, building such a well-connected road network may not only be expensive, but also time-consuming. On the other hand, as seen in the example of the ring-radial networks, combining features of poorly connected road systems may result in a road system that performs much better than the original sources. This may serve as the second-best approach to assist the local needs in improving transportation accessibility and connectivity at a lower budget.

While insights from this study are of both academic and policy implications, they are based on simulations using stylized road systems. The roads in reality are

complex and diverse, and cannot possibly be represented by the four road systems designed for this study. Our stylized road systems focus on the connection patterns of nodes and edges, but neglect nodes of various importance, such as the feeder, hub, and gateway. Thus, the road systems used in the study do not reflect the variations of characteristics of topological structure at the local, regional, or national scales. In addition, the methods used in the study have not been tested against the practical needs of the private and public institutions of the real world. These limitations point to the needs and directions for continuing improvement in future investigations.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- [1] Eiselt, H.A. and Sandblom, C.-L. (2004) Decision Analysis, Location Models, and Scheduling Problems. Springer-Verlag.
- [2] Klose, A. and Drexl, A. (2005) Facility Location Models for Distribution System Design. *European Journal of Operational Research*, **162**, 4-29. <https://doi.org/10.1016/j.ejor.2003.10.031>
- [3] Carling, K., Han, M. and Håkansson, J. (2012) Does Euclidean Distance Work Well When the p -Median Model Is Applied in Rural Areas? *Annals of Operations Research*, **201**, 83-97. <https://doi.org/10.1007/s10479-012-1214-2>
- [4] Neema, M.N. and Ohgai, A. (2010) Multi-Objective Location Modeling of Urban Parks and Open Spaces: Continuous Optimization. *Computers, Environment and Urban Systems*, **34**, 359-376. <https://doi.org/10.1016/j.compenvurbsys.2010.03.001>
- [5] ReVelle, C.S. and Eiselt, H.A. (2005) Location Analysis: A Synthesis and Survey. *European Journal of Operational Research*, **165**, 1-19. <https://doi.org/10.1016/j.ejor.2003.11.032>
- [6] Streck, S. (2010) The Fermat Problem.
- [7] Rodrigue, J.-P. (2020) The Geography of Transport Systems. Routledge.
- [8] Kulin, H.W. and Kuenne, R.E. (1962) An Efficient Algorithm for the Numerical Solution of the Generalized Weber Problem in Spatial Economics. *Journal of Regional Science*, **4**, 21-33. <https://doi.org/10.1111/j.1467-9787.1962.tb00902.x>
- [9] Erlenkotter, D. (1978) A Dual-Based Procedure for Uncapacitated Facility Location. *Operations Research*, **26**, 992-1009. <https://doi.org/10.1287/opre.26.6.992>
- [10] Teitz, M.B. and Bart, P. (1968) Heuristic Methods for Estimating the Generalized Vertex Median of a Weighted Graph. *Operations Research*, **16**, 955-961. <https://doi.org/10.1287/opre.16.5.955>
- [11] Moon, I.D. and Chaudhry, S.S. (1984) An Analysis of Network Location Problems with Distance Constraints. *Management Science*, **30**, 290-307. <https://doi.org/10.1287/mnsc.30.3.290>
- [12] Daskin, M.S. and Maass, K.L. (2015) The p -Median Problem. In: Laporte, G., Nickel, S. and da Gama, F.S., Eds., *Location Science*, Springer International Publishing, 21-45. https://doi.org/10.1007/978-3-319-13111-5_2
- [13] Lei, T.L. (2021) Integrating GIS and Location Modeling: A Relational Approach. *Transactions in GIS*, **25**, 1693-1715. <https://doi.org/10.1111/tgis.12804>

- [14] Roozkhosh, P. and Motahari Farimani, N. (2022) Designing a New Model for the Hub Location-Allocation Problem with Considering Tardiness Time and Cost Uncertainty. *International Journal of Management Science and Engineering Management*, **18**, 36-50. <https://doi.org/10.1080/17509653.2022.2089261>
- [15] Li, H., Zhang, B. and Ge, X. (2022) Modeling Emergency Logistics Location-Allocation Problem with Uncertain Parameters. *Systems*, **10**, Article No. 51. <https://doi.org/10.3390/systems10020051>
- [16] Saif, A. and Delage, E. (2021) Data-Driven Distributionally Robust Capacitated Facility Location Problem. *European Journal of Operational Research*, **291**, 995-1007. <https://doi.org/10.1016/j.ejor.2020.09.026>
- [17] Li, H., Yu, D., Zhang, Y. and Yuan, Y. (2025) A Two-Stage Robust Optimization Model for Emergency Service Facilities Location-Allocation Problem under Demand Uncertainty and Sustainable Development. *Scientific Reports*, **15**, Article No. 2895. <https://doi.org/10.1038/s41598-025-86129-1>
- [18] Çelik, S. and Ok, Ş. (2024) Electric Vehicle Charging Stations: Model, Algorithm, Simulation, Location, and Capacity Planning. *Heliyon*, **10**, e29153. <https://doi.org/10.1016/j.heliyon.2024.e29153>
- [19] Han, J., Zhang, J., Zeng, B. and Mao, M. (2021) Optimizing Dynamic Facility Location-Allocation for Agricultural Machinery Maintenance Using Benders Decomposition. *Omega*, **105**, Article ID: 102498. <https://doi.org/10.1016/j.omega.2021.102498>
- [20] Sun, X., Yu, H. and Solvang, W.D. (2020) Solving the Location Problem of Printers in a University Campus Using p -Median Location Model and Anylogic Simulation. In: Wang, Y., *et al.*, Eds., *Advanced Manufacturing and Automation IX*, Springer, 577-584. https://doi.org/10.1007/978-981-15-2341-0_72
- [21] Bartkowski, B., Beckmann, M., Drechsler, M., Kaim, A., Liebelt, V., Müller, B., *et al.* (2020) Aligning Agent-Based Modeling with Multi-Objective Land-Use Allocation: Identification of Policy Gaps and Feasible Pathways to Biophysically Optimal Landscapes. *Frontiers in Environmental Science*, **8**, Article No. 103. <https://doi.org/10.3389/fenvs.2020.00103>
- [22] Zhao, J. and Zhang, F. (2025) Agent-Based Simulation System for Optimising Resource Allocation in Production Process. *IET Collaborative Intelligent Manufacturing*, **7**, e70020. <https://doi.org/10.1049/cim2.70020>
- [23] Chang, K., Pan, Y. and Chen, H. (2024) Shelter Location-Allocation Problem for Disaster Evacuation Planning: A Simulation Optimization Approach. *Computers & Operations Research*, **171**, Article ID: 106784. <https://doi.org/10.1016/j.cor.2024.106784>
- [24] Chen, Y., Yao, H., Weng, S. and Tai, Y. (2022) An Analysis of the Optimal Facility Location of Tourism Industry in Plain Region by Utilizing GIS. *Sage Open*, **12**, 1-14. <https://doi.org/10.1177/21582440221095020>
- [25] Abd El Karim, A. and Awawdeh, M.M. (2020) Integrating GIS Accessibility and Location-Allocation Models with Multicriteria Decision Analysis for Evaluating Quality of Life in Buraidah City, KSA. *Sustainability*, **12**, Article No. 1412. <https://doi.org/10.3390/su12041412>
- [26] Vafaeinejad, A., Bolouri, S., Alesheikh, A.A., Panahi, M. and Lee, C. (2020) The Capacitated Location-Allocation Problem Using the VAOMP (Vector Assignment Ordered Median Problem) Unified Approach in GIS (Geospatial Information System). *Applied Sciences*, **10**, Article No. 8505. <https://doi.org/10.3390/app10238505>
- [27] Perreur, J. and Thisse, J. (1974) Central Metrics and Optimal Location. *Journal of Regional Science*, **14**, 411-421. <https://doi.org/10.1111/j.1467-9787.1974.tb00463.x>
- [28] Schilling, D.A., Rosing, K.E. and ReVelle, C.S. (2000) Network Distance Characteris-

- tics that Affect Computational Effort in p -Median Location Problems. *European Journal of Operational Research*, **127**, 525-536. [https://doi.org/10.1016/s0377-2217\(99\)00336-7](https://doi.org/10.1016/s0377-2217(99)00336-7)
- [29] Carling, K., Han, M., Håkansson, J. and Rebreyend, P. (2014) Distance Measure and the p -Median Problem in Rural Areas. *Annals of Operations Research*, **226**, 89-99. <https://doi.org/10.1007/s10479-014-1677-4>
- [30] Peeters, D. and Thomas, I. (1995) The Effect of Spatial Structure on p -Median Results. *Transportation Science*, **29**, 366-373. <https://doi.org/10.1287/trsc.29.4.366>
- [31] Zhao, X., Rebreyend, P. and Håkansson, J. (2019) How Does the Complexity of a Road Network Affect Optimal Facility Locations? *Journal of Traffic and Logistics Engineering*, **7**, 64-71. <https://doi.org/10.18178/jtle.7.2.64-71>
- [32] Wahlisch, M. (2010) Modeling the Network Topology. In: Wehrle, K., Günes, M. and Gross, J., Eds., *Modeling and Tools for Network Simulation*, Springer, 471-486.
- [33] Steinbach, S.F. (2018) Alternative Transport. Network Types Overview—Alternative Transport.
- [34] Wasserman, S. and Faust, K. (1994) Social Network Analysis. Cambridge University Press. <https://doi.org/10.1017/cbo9780511815478>
- [35] Borgatti, S.P., Everett, M.G. and Johnson, J.C. (2013) Analyzing Social Networks. Sage.
- [36] Freeman, L.C. (1978) Centrality in Social Networks Conceptual Clarification. *Social Networks*, **1**, 215-239. [https://doi.org/10.1016/0378-8733\(78\)90021-7](https://doi.org/10.1016/0378-8733(78)90021-7)
- [37] Butts, C.T. (2008) Social Network Analysis with SNA. *Journal of Statistical Software*, **24**, 1-51. <https://doi.org/10.18637/jss.v024.i06>
- [38] Verter, V. (2011) Uncapacitated and Capacitated Facility Location Problems. In: Eiselt, H.A. and Marianov, V., Eds., *Foundations of Location Analysis*, Springer Science and Business Media, 25-37. https://doi.org/10.1007/978-1-4419-7572-0_2
- [39] Kim, D., Chun, Y. and Griffith, D.A. (2024) Impacts of Spatial Imputation on Location-Allocation Problem Solutions. *Spatial Statistics*, **59**, Article ID: 100810. <https://doi.org/10.1016/j.spasta.2024.100810>
- [40] Xie, F. and Levinson, D. (2007) Measuring the Structure of Road Networks. *Geographical Analysis*, **39**, 336-356. <https://doi.org/10.1111/j.1538-4632.2007.00707.x>
- [41] Arostegui, M.A., Kadipasaoglu, S.N. and Khumawala, B.M. (2006) An Empirical Comparison of Tabu Search, Simulated Annealing, and Genetic Algorithms for Facilities Location Problems. *International Journal of Production Economics*, **103**, 742-754. <https://doi.org/10.1016/j.ijpe.2005.08.010>
- [42] Hakimi, S.L. (1965) Optimum Distribution of Switching Centers in a Communication Network and Some Related Graph Theoretic Problems. *Operations Research*, **13**, 462-475. <https://doi.org/10.1287/opre.13.3.462>