

Vehicle Routing Planning Using a Hybrid Approach of Spectral Clustering and Nearest Neighbour Algorithms

Natapat Areerakulkan^{1*}, Lanlalit Lhaochot², and Sivarak Kijwattanaphokin³

College of Logistics and Supply Chain, Rajabhat Suan Sunandha University^{1}*

Faculty of Liberal Arts, Rajamangala University of Technology Rattanakosin²

Newgen Energy Solution Co., Ltd³

* Corresponding Author E-mail: natapat.ar@ssru.ac.th, Lanlalit.l@rmutr.ac.th², sivarak@ngesth.com³

Abstract

Vehicle routing is a critical problem in logistics and supply chain management, particularly in urban environments where delivery points are spatially complex and demand constraints must be satisfied. This study proposes a hybrid approach combining spectral clustering and the Nearest Neighbour (NN) algorithm to improve routing efficiency. The dataset consists of 51 delivery points located in Bangkok, Thailand, with demand information derived from historical data. Spectral clustering is first applied to group delivery points into four clusters based on spatial proximity, effectively reducing problem complexity, after which the NN algorithm is used to construct delivery routes within each cluster. The results show that the proposed method successfully generates four feasible routes, all of which satisfy the vehicle capacity constraint of 120 units. The total travel distance is reduced from 320 km to 268 km, representing an improvement of approximately 16.25%. This reduction leads to an estimated cost saving of 1,560 THB per trip, or approximately 468,000 THB annually (assuming 300 operating days). Additionally, the approach contributes to environmental sustainability by reducing carbon emissions by approximately 3.28 tons of CO₂ per year. Overall, the proposed hybrid method demonstrates strong potential for improving routing efficiency, reducing operational costs, and supporting sustainable logistics practices, with future work focusing on integrating advanced optimization techniques and real-time data.

Keywords: Vehicle Routing Problem, Spectral Clustering, Nearest Neighbour Algorithm, Logistics Optimization, Sustainable Transportation.

Introduction

Transportation routing is one of the key issues in logistics and supply chain management, with the goal of determining the most efficient transportation routes under constraints such as time, cost, vehicle capacity, and customer needs. This issue is commonly referred to as a vehicle routing problem, which is a highly complex problem and classified as NP-hard. This makes it difficult to find the optimal solution using traditional methods, especially as the size of the problem increases (Toth & Vigo, 2014; Konovalenko & Hvattum, 2024). In the past, the solution to routing problems often relied on mathematical methods and heuristic algorithms, such as the Clarke-Wright Savings Algorithm or Genetic Algorithm (Konstantakopoulos et al., 2022). They can provide good answers in a limited time but also

have limitations in dealing with big data and uncertain environments (Laporte, 2009). However, with the development in data technology and computing power in the present day, the concepts of machine learning play an important role in solving this routing problem. Machine learning allows systems to learn patterns from historical data, such as traffic patterns, ordering behavior, or delivery times, to improve routing decisions more efficiently. In particular, advanced techniques such as deep learning and reinforcement learning can be used in sequential learning and dynamic decision-making, which are suitable for rapidly changing routing problems (Kool et al., 2019; Bai et al., 2023).

Spectral clustering is a common machine learning method for solving routing problems. It groups data points that are similar by using the structure of graphs and the eigenvalues of Laplacian matrices. This technique is ideal for grouping customers or points of delivery in a spatial manner. In the field of transportation, spectral clustering can be used to find complex spatial structures, like groups of customers with nonlinear geometric dispersions, that other grouping methods, like K-means clustering, might not be able to handle. As a result, service zones are more accurate, leading to a reduction in total transportation distances and delivery times (Shi & Malik, 2000; Von Luxburg, 2007). Especially in big data and uncertain environments such as large cities or smart transportation systems, it can evidently lead to cost reductions, increase efficiency, and improve service quality in a sustainable manner (Waller & Fawcett, 2013; Shahbazian et al., 2024). Therefore, transportation routing using machine learning algorithms is a high-potential approach to the development of modern logistics systems. In addition, machine learning techniques are commonly based on a combination of both data-driven learning and mathematical optimization. This makes them able to rapidly respond to the fast-paced environment of the supply chain in the digital age, which is crucial for improving efficiency, reducing costs, and enhancing customer satisfaction (León Villalba & González La Rotta, 2022).

Research Objectives

This research aims to study the application of machine learning techniques and optimization methods to solve transportation routing problems. The vehicle routing problem is a complex issue that has many limitations, particularly regarding vehicle capacity. In this study, the focus is on the use of spectral clustering to group customers according to their spatial proximity to help reduce the complexity of the problem before planning the transportation route for each group. In addition, the research aims to understand the process through the preparation of geographic location data for the delivery point and warehouse. The next step is to calculate distances between points and create a distance matrix, which is an important basis for route planning. Then we apply simplified routing methods such as Nearest Neighbor to the results of grouping to create appropriate transportation routes within defined constraints. This can increase service efficiency, and support data-driven decision-making, which can be effectively applied in the context of modern supply chains.

Materials and Methods

This study focuses on vehicle routing planning for a logistics company with its depot located in Bangkok, Thailand. The delivery network consists of 51 drop points distributed across urban areas with irregular spatial patterns as shown in figure 1. The company operates four delivery trucks, each with a maximum capacity of 120 units.

Data Collection and Preparation

The dataset used in this study includes geographic and demand information for each delivery point. Location data are represented using latitude and longitude coordinates, while demand values are calculated as the average daily demand derived from historical records. A sample of the dataset structure, including location and demand attributes, is presented in Table 1. To support route planning, the geographic coordinates were used to compute pairwise distances between all points, forming a distance matrix. This matrix serves as the foundation for both clustering and routing processes.

Clustering Method

To reduce the complexity of the vehicle routing problem, spectral clustering was applied to group delivery points based on spatial proximity. This method leverages graph-based similarity and Laplacian matrix decomposition to identify natural groupings within non-linear spatial distributions. The optimal number of clusters was determined prior to routing (four clusters in this study), allowing each cluster to represent a service zone corresponding to a delivery route.

Route Construction

After clustering, routes within each cluster were constructed using the Nearest Neighbour (NN) algorithm. Starting from the depot, the algorithm iteratively selects the closest unvisited delivery point until all points in the cluster are visited and then returns to the depot. This approach provides a simple and computationally efficient heuristic for route generation.

Feasibility Check

For each generated route, the total demand of all assigned drop points was calculated and compared against the vehicle capacity constraint (120 units). Routes exceeding capacity were considered infeasible and would require adjustment (if applicable).

Workflow Summary

The overall methodology consists of the following steps:

1. Group delivery points using spectral clustering.
2. Visualize clusters to validate grouping effectiveness.
3. Generate routes for each cluster using the nearest neighbour method.
4. Evaluate route feasibility based on vehicle capacity constraints.

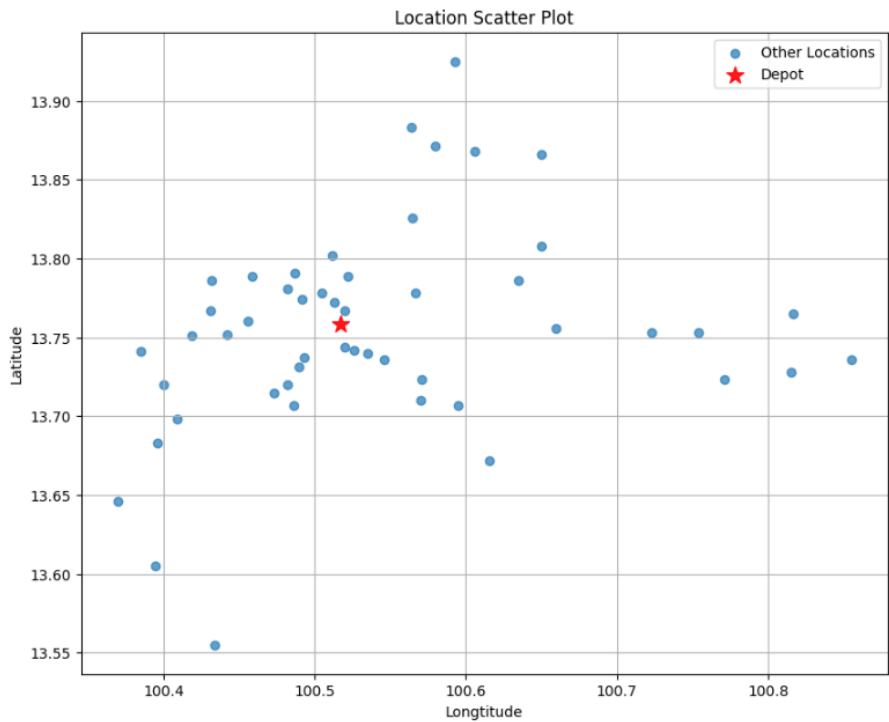


Figure 1: 51 drop points scatter plot (Areerakulkan, 2025)

Table 1: First five rows of input data (Areerakulkan, 2025)

	Name	District	Latitude	Longitude	Demand
0	DP1	Din Daeng	13.778	100.567	9
1	Depot	Dusit	13.758	100.517	13
2	DP2	Thon Buri	13.731	100.490	9
3	DP3	Thon Buri	13.737	100.493	6
4	DP4	Taling Chan	13.760	100.456	9

Results

The vehicle routing planning process began with grouping the 51 delivery points using the spectral clustering method. The optimal number of clusters was determined to be four based on the elbow method. The clustering results are presented in Figure 2, which illustrates the spatial distribution of delivery points divided into four distinct clusters.

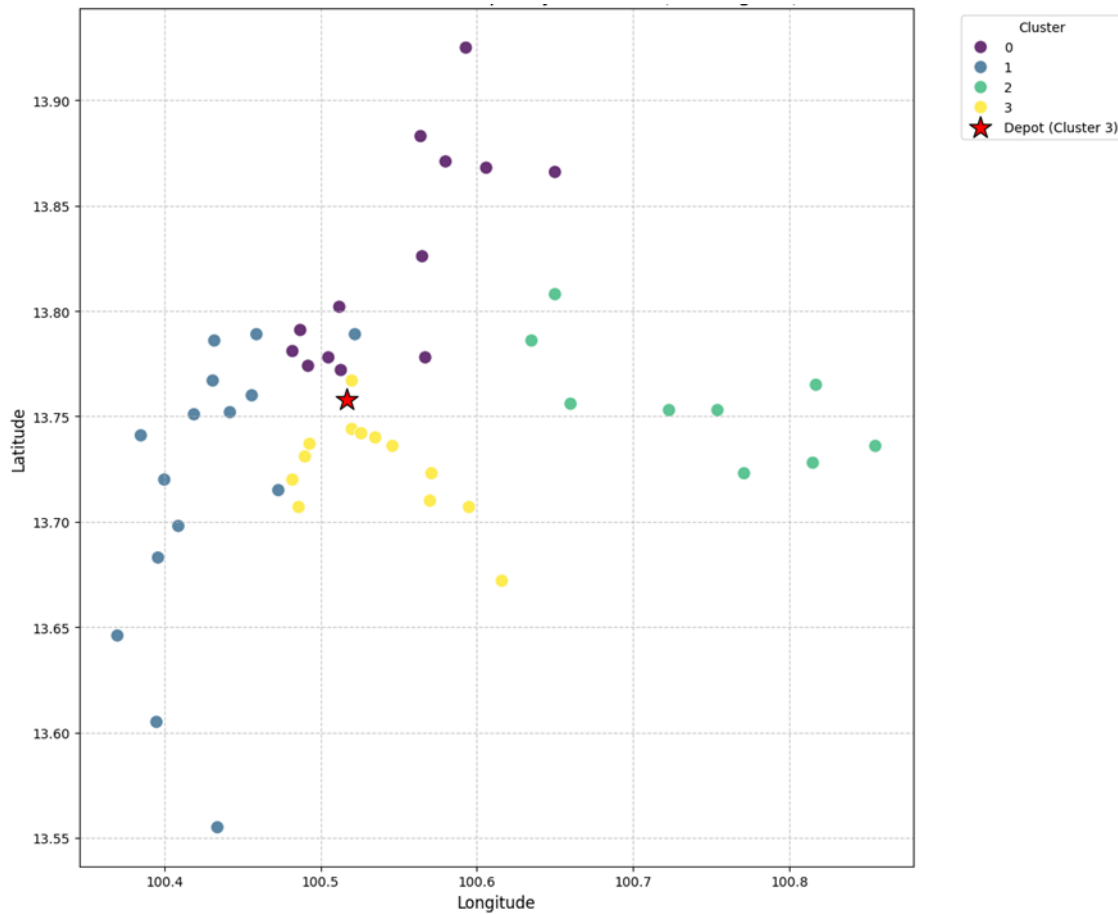


Figure 2: Spectral clustering into 4 clusters (Areerakulkan, 2025)

As shown in Figure 2, delivery points are grouped based on spatial proximity, resulting in compact and well-separated clusters. This clustering structure helps reduce problem complexity and ensures that each group can be served within a localized area. The depot is positioned near the center of the service region, supporting efficient route connectivity.

Following the clustering process, route construction was performed for each cluster using the Nearest Neighbour (NN) algorithm. The routing results are illustrated in Figure 3, where each cluster is represented by a distinct route connecting all assigned delivery points.

As shown in Figure 3, each route starts and ends at the depot and follows a continuous path that connects nearby delivery points within the same cluster. The NN algorithm successfully generates logical routing sequences with minimal backtracking and no overlap between clusters.

A total of four routes were obtained, corresponding to the four clusters identified in Figure 2. Each route includes only the delivery points assigned to its cluster, ensuring clear separation of service zones and improving operational efficiency.

Furthermore, the total demand for each route was evaluated against the vehicle capacity constraint of 120 units. The results confirm that all routes satisfy the capacity limitation, indicating that the proposed approach effectively balances load distribution among vehicles.

Overall, the combination of spectral clustering (Figure 2) and nearest neighbour routing (Figure 3) provides an efficient and feasible solution for the vehicle routing problem. The clustering step reduces spatial complexity, while the routing step generates practical delivery paths suitable for real-world logistics operations.

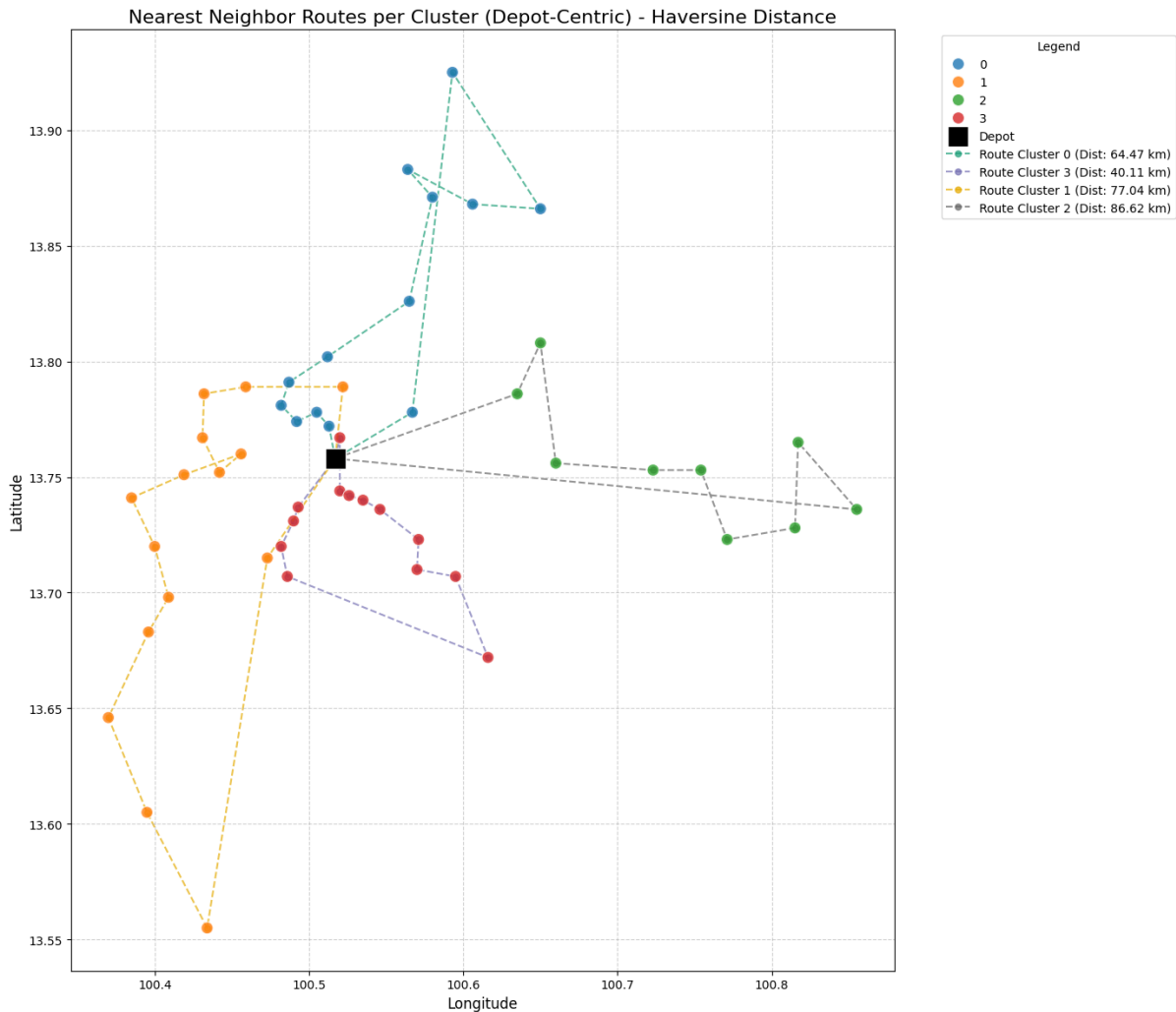


Figure 3: Transportation route planning for each cluster (Areerakulkan, 2025)

Route Descriptions

Route 1:

Depot → DP13 → DP5 → DP15 → DP14 → DP17 → DP16 → DP45 → DP24 → DP38 → DP40 → DP41 → DP48 → DP1 → Depot

Total demand: 119 units; Total distance: 64.41 km.

This route covers delivery points located in a relatively dense cluster near the central area. The sequence shows smooth progression between nearby points, minimizing travel distance and avoiding unnecessary backtracking.

Route 2:

Depot → DP23 → DP25 → DP27 → DP29 → DP22 → DP4 → DP20 → DP35 → DP32 → DP31 → DP28 → DP47 → DP49 → DP50 → DP9 → Depot

Total demand: 95 units; Total distance: 40.11 km.

This route includes a larger number of delivery points distributed across a wider area. Despite the broader coverage, the NN algorithm maintains an efficient visiting order by prioritizing the nearest unvisited locations.

Route 3:

Depot → DP33 → DP42 → DP34 → DP36 → DP39 → DP30 → DP43 → DP46 → DP44 → Depot

Total demand: 67 units; Total distance: 86.62 km.

This route serves a smaller cluster with fewer delivery points. The route is relatively short and compact, indicating high spatial concentration within this cluster.

Route 4:

Depot → DP10 → DP6 → DP8 → DP12 → DP11 → DP26 → DP19 → DP37 → DP21 → DP7 → DP2 → DP3 → Depot

Total demand: 106 units; Total distance: 77.04 km.

This route covers delivery points that are moderately dispersed. The routing path reflects a structured traversal pattern that connects outer points efficiently before returning to the depot.

Capacity Feasibility

The total demand for each route was evaluated against the vehicle capacity constraint of 120 units. The results indicate that all routes satisfy the capacity requirement, confirming that the clustering and routing approach produces feasible solutions without exceeding vehicle limitations.

Discussion

The results demonstrate that combining spectral clustering with the Nearest Neighbour (NN) algorithm is an effective approach for solving the vehicle routing problem in an urban setting. Spectral clustering successfully groups delivery points based on spatial proximity, even under irregular distributions, thereby reducing problem complexity by dividing it into smaller service zones.

The NN algorithm generates feasible routes within each cluster, producing logical delivery sequences with minimal backtracking. Although it does not guarantee optimal solutions, its computational simplicity makes it suitable for practical applications.

Importantly, the total demand of each route remains within the vehicle capacity constraint, indicating that the proposed method is operationally feasible. The clustering process also contributes to a relatively balanced distribution of workload among routes, despite not explicitly considering demand during clustering.

However, the approach has limitations. The NN algorithm may yield suboptimal routes due to its greedy nature, and the model assumes static conditions without accounting for real-world factors such as traffic or time constraints. Future work could incorporate demand-aware clustering and more advanced optimization techniques to improve solution quality.

Overall, the proposed hybrid method effectively reduces routing complexity and generates feasible delivery plans, demonstrating strong potential for real-world logistics applications.

Conclusion

This study proposes a hybrid approach combining spectral clustering and the Nearest Neighbour (NN) algorithm to solve the vehicle routing problem in an urban logistics context. The results demonstrate that spectral clustering effectively groups delivery points into geographically compact clusters, thereby reducing problem complexity. The NN algorithm subsequently generates feasible and well-structured routes within each cluster.

The findings show that the total travel distance was reduced from 320 km to 268 km, representing an improvement of approximately 16.25%, or a reduction of 52 km per routing cycle.

In terms of economic impact, assuming an average transportation cost of 30 THB per kilometer, the reduction results in a cost saving of approximately 1,560 THB per trip. Considering daily operations (e.g., 300 working days per year), the total annual cost saving is approximately 468,000 THB per year.

From an environmental perspective, using an emission factor of 0.21 kg CO₂ per kilometer, the 52 km reduction corresponds to approximately 10.92 kg CO₂ reduction per trip. On an annual basis (300 operating days), this results in a reduction of approximately 3,276 kg CO₂ per year (or 3.28 tons CO₂ per year).

In addition, all routes satisfy the vehicle capacity constraint, confirming the practical feasibility of the solution. The combination of clustering and heuristic routing provides a good balance between computational efficiency and solution quality, making it suitable for real-world implementation.

Overall, the proposed method not only improves routing efficiency but also delivers significant economic and environmental benefits. Future work may focus on integrating advanced optimization techniques, demand balancing, and real-time data to further enhance routing performance and sustainability.

Acknowledgments: None.

Funding Statement: None.

Conflict of Interest: None.

Ethics Approval: None.

References

- [1] Areerakulkan, N. (2025). *Data analytics and machine learning for supply chain management* [Teaching material]. Rajabhat Suan Sunandha University.
- [2] Bai, R., Chen, X., Chen, Z.-L., Cui, T., Gong, S., He, W., Jiang, X., Jin, H., Jin, J., Kendall, G., Li, J., Lu, Z., Ren, J., Weng, P., Xue, N., & Zhang, H. (2023). Analytics and machine learning in vehicle routing research. *International Journal of Production Research*, 61(1), 4–30. <https://doi.org/10.1080/00207543.2021.2013566>
- [3] Kool, W., van Hoof, H., & Welling, M. (2019). Attention, learn to solve routing problems! *International Conference on Learning Representations (ICLR)*. <https://doi.org/10.48550/arXiv.1803.08475>
- [4] Konovalenko, A., & Hvattum, L. M. (2024). Optimizing a dynamic vehicle routing problem with deep reinforcement learning: Analyzing state-space components. *Logistics*, 8(4), Article 96. <https://doi.org/10.3390/logistics8040096>
- [5] Konstantakopoulos, G. D., Gayialis, S. P., & Kechagias, E. P. (2022). Vehicle routing problem and related algorithms for logistics distribution: A literature review and classification. *Operational Research*, 22, 2033–2062. <https://doi.org/10.1007/s12351-020-00600-7>
- [6] Laporte, G. (2009). Fifty years of vehicle routing. *Transportation Science*, 43(4), 408–416. <https://doi.org/10.1287/trsc.1090.0301>
- [7] León Villalba, A. F., & González La Rotta, E. C. (2022). Clustering and heuristics algorithm for the vehicle routing problem with time windows. *International Journal of Industrial Engineering Computations*, 13(2), 165–184. <https://doi.org/10.5267/j.ijiec.2021.12.002>
- [8] Shahbazian, R., Di Puglia Pugliese, L., Guerriero, F., & Macrina, G. (2024). Integrating machine learning into vehicle routing problem: Methods and applications. *IEEE Access*, 12, 93087–93115. <https://doi.org/10.1109/ACCESS.2024.3422479>
- [9] Shi, J., & Malik, J. (2000). Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8), 888–905. <https://doi.org/10.1109/34.868688>
- [10] Toth, P., & Vigo, D. (2014). *Vehicle routing: Problems, methods, and applications* (2nd ed.). Society for Industrial and Applied Mathematics.
- [11] Von Luxburg, U. (2007). A tutorial on spectral clustering. *Statistics and Computing*, 17, 395–416. <https://doi.org/10.1007/s11222-007-9033-z>
- [12] Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. *Journal of Business Logistics*, 34(2), 77–84. <https://doi.org/10.1111/jbl.12010>