



# Enhancing e-grocery order fulfillment: improving product availability, cost, and emissions in last-mile delivery

Banu Y. Ekren<sup>1,2</sup> · Sara Perotti<sup>3</sup> · Laura Foresti<sup>3</sup> · Lorenzo Pratavia<sup>1</sup>

Accepted: 9 December 2023  
© Crown 2024

## Abstract

This paper studies e-grocery order fulfillment policies by leveraging both customer and e-grocery-based data. Through the utilization of historical purchase data, product popularity trends, and delivery patterns, allocation strategies are informed to optimize performance metrics such as fill rate, carbon emissions, and cost per order. The study aims to conduct a sensitivity analysis to identify key drivers influencing these performance metrics. The results highlight that fulfillment policies optimized with the utilization of the mentioned data metrics demonstrate superior performance compared to policies not informed by data. These findings underscore the critical role of integrating data-driven models in e-grocery order fulfillment. Based on the outcomes, a grocery allocation policy, considering both proximity and product availability, emerges as promising for simultaneous improvements in several performance metrics. The study recommends that e-grocery companies leverage customer data to design and optimize delivery-oriented policies and strategies. To ensure adaptability to new trends or changes in delivery patterns, continual evaluation and improvement of e-grocery fulfillment policies are emphasized.

**Keywords** E-grocery · Last-mile · Order delivery · E-commerce · Fulfilment

## List of symbols

$D = (1, 2, 3)$	Depots
$t = 1, \dots, 500$	Products in the assortment
$(a\%, b\%, c\%)$	Percentage of products in class A, B, and C, respectively
$(L_A, L_B, L_C)$	Ranges for generation of $pa_{it}$ in L scenario for class A, B, and C

---

✉ Banu Y. Ekren  
banu.yetkinekren@cranfield.ac.uk; banu.ekren@yasar.edu.tr

<sup>1</sup> School of Management, Cranfield University, Bedfordshire MK43 0AL, UK

<sup>2</sup> Department of Industrial Engineering, Yasar University, Bornova, Izmir, Turkey

<sup>3</sup> Politecnico di Milano, Department of Management, Economics and Industrial Engineering Via Lambruschini, 4/B, 20156 Milano, Italy

$(M_A, M_B, M_C)$	Ranges for generation of $pa_{it}$ in M scenario for class A, B, and C
$(H_A, H_B, H_C)$	Ranges for generation of $pa_{it}$ in H scenario for class A, B, and C
$(x\%, y\%, z\%)$	Percentage of products in class X, Y, and Z
$(X\%, Y\%, Z\%)$	Cumulative probability of ordering products in class X, Y, and Z
C	Delivery cost km
$e_{van}$	Emissions per km traveled by van
$e_{car}$	Emissions per km traveled by car
$d_{local}$	Average distance between the customer's home and a local store
XYZ/ABC	Degree of overlapping between XYZ and ABC classes
K	Geographical dispersion factor
$d\%$	Delivery trip inefficiency
$g\%$	Probability of going complementary shopping
$po_t$	Customer's probability of ordering product $t$ in the next
$pa_{it}$	Historical probability of product availability for product $t$ in depot $i \in D$
$d_i$	Distance between depot $i \in D$ and the customer
$d_{ij}$	Distance between depot $i \in D$ and $j \in D$
$\alpha$	Importance of improving product availability over reducing travel distance
$AS_i$	Availability score for depot $i \in D$
$CS_i$	Combined score for depot $i \in D$
$MAS_{iz}$	Merged availability score for depot $i \in D$ and depot $z \in D$
$MCS_{iz}$	Merged combined score for depot $i \in D$ and depot $z \in D$
$s_{ru}$	Availability setting in the simulation run $ru$
$\Lambda_{ru}$	Set of allocated depots for the fulfillment of the order at order time $ru$
$N_{ru}$	Basket size in the simulation run $ru$
$B_{ru}$	Set of $N_{ru}$ distinct products ordered in the simulation run $ru$
$AS_{it_{ru}}$	Binary variable that represents whether product $t$ is available in depot $i \in \Lambda_{ru}$ (1) or not (0) in the simulation run $ru$
$pick\_in\_farthest_{ru}$	Binary variable that represents whether one or more items of the customer order are picked in the farthest allocated depot (1) or not (0) in the simulation run $ru$
$pick\_in\_middle_{ru}$	Binary variable that represents whether one or more items of the customer order are picked in the middle allocated depot (1) or not (0) in the simulation run $ru$
$go\_buy_{ru}$	Binary variable that represents whether the customer does complementary shopping (1) or not (0) in the simulation run $ru$
RU	Number of runs in the Monte Carlo simulation
RE	Number of replications of the Monte Carlo simulation

## 1 Introduction

The digital revolution has reshaped consumer behaviors and preferences, propelling the growth of e-commerce and online purchasing [8], PwC [39, 40]. Consumers are increasingly turning to online platforms to fulfill their shopping needs, drawn by the convenience of browsing a vast array of products, making purchases from the comfort of their homes, and having items delivered directly to their doorsteps [55]. Within this landscape, the e-grocery sector has gained substantial traction, capitalizing on the changing consumer dynamics [30]. E-grocery combines the convenience of online shopping with the necessity of grocery items, offering consumers a convenient way to acquire essential goods without the need to visit physical stores.

E-grocery platforms offer a plethora of benefits to consumers [10]. The ability to order groceries online provides a level of convenience that is particularly appealing to busy urban residents, individuals with limited mobility, and those seeking to minimize time spent on routine shopping activities. Also, a notable advantage lies in the typically broader product range offered by online fulfillment centers compared to local stores. E-grocery shoppers can enjoy a wide array of choices when making their purchases online. Furthermore, e-grocery aligns with the growing demand for environmental sustainability, as online orders can be consolidated, reducing the number of trips to physical stores and potentially lowering overall carbon emissions [36]. However, as the e-grocery industry continues to evolve, several challenges related to product availability, order fulfillment costs, and environmental impact need to be addressed [32, 34]. These challenges highlight the need for continuous innovation and improvement within the industry. Maintaining a diverse range of available products, optimizing the logistics and delivery processes to manage costs effectively, and finding ways to further reduce the carbon footprint of e-grocery operations remain critical areas of focus. Within this study, we treat these three focal areas as integral sustainable practices. We consider sustainability from three vital perspectives – social, environmental, and economic. Social sustainability focuses on satisfying customers and fostering loyalty. Environmental sustainability involves reducing the environmental impact, particularly carbon emissions. Economic sustainability aims at cost-efficiency and profitability.

In the e-grocery context, finding available products that meet the preferences of online shoppers is of paramount importance. The digital shelf must mirror the physical store's inventory, offering customers a seamless and accurate selection experience [31]. However, e-grocery customers face the risk of encountering two types of stock-outs: pre-purchase out-of-stock and post-purchase out-of-stock. While pre-purchase stock-outs occur when customers are informed about a product's unavailability before placing an order, post-purchase stock-outs can arise due to delays between order placement and fulfillment or inaccuracies in product availability information. These instances of stock-outs can result in customer dissatisfaction, missed sales opportunities, and a potential decline in customer loyalty [15]. Moreover, these implications highlight the importance of considering social sustainability solutions as well.

The delivery phase further compounds the challenges faced by e-grocery platforms. To minimize last-mile delivery costs and environmental impact, online customers are often assigned to specific grocery stores or fulfillment centers based on their proximity [22, 35, 53]. While this strategy reduces delivery distances, it may lead to scenarios where products selected by customers during online shopping become unavailable at the time of delivery. To address this issue, e-grocery platforms frequently employ substitution strategies, providing alternative products when initially selected items are out of stock [31]. Achieving a balance between product availability, delivery efficiency, and customer satisfaction remains a critical concern.

This paper aims to bridge the existing research gap by exploring e-grocery order fulfillment policies that leverage customer and e-grocery-based data. The central objectives include enhancing product availability, promoting sustainability, and optimizing cost efficiency in last-mile delivery. To achieve these goals, we develop and compare different order fulfillment center allocation policies for online customers, evaluating their performance based on pre-defined objectives.

To guide the study, we address the following research questions:

- RQ1: How e-grocery fulfillment policies affect this industry's performance from sustainable delivery perspectives (i.e., product availability, cost-efficiency and carbon emission)?
- RQ2: How customer and e-grocery relevant data can be utilized effectively for managing e-grocery fulfillment?

In addressing RQ1, our aim is to investigate the influence of e-grocery fulfillment policies on the industry's performance, with a particular focus on product availability, cost-efficiency, and carbon emissions in delivery, all of which are key aspects associated with sustainability. To achieve this, we devise e-grocery fulfillment policies that strategically allocate resources, such as e-grocery stores, depots, or fulfillment centers, for processing online orders. These policies are designed based on insights derived from customer and e-grocery-related data, as discussed in RQ2. We then proceed to simulate these policies and assess their performance using key performance metrics, which include: fill rate (product availability), carbon emissions, and cost per order.

In RQ2, the effective utilization of customer and e-grocery-based data is explored to improve developed fulfillment policies, ensuring they are optimized for performance. This involves leveraging historical purchase data, product popularity trends, and delivery patterns to inform allocation strategies. We conduct a sensitivity analysis to ascertain how performance metrics are influenced by input data. We utilize an experimental design and apply ANOVA to identify factors significantly impacting performance metrics. Through the application of ANOVA, the intention is to mitigate potential biases stemming from the utilization of hypothetical data.

The organization of this paper is as follows: In Sect. 2, an extensive review of the existing literature is presented, encapsulating cutting-edge research on

challenges and solution methodologies in the realm of e-grocery. Special emphasis is placed on delineating issues concerning product availability and the strategies associated with allocating fulfillment centers for order processing. Moving to Sect. 3, the structure of the problem is expounded upon, elucidating the pivotal constituents and considerations intrinsic to the domain of e-grocery fulfillment. Section 4 delves into the intricacies of our chosen simulation modeling approach, illuminating how it is harnessed to address the problem at hand. The culmination of our analysis is found in Sect. 5, where we furnish the outcomes of our simulations, spotlighting the efficacy and performance of the bespoke e-grocery fulfillment policies developed in this study. Finally, Sect. 6 serves as a comprehensive denouement, encapsulating the synthesis of our efforts and encapsulating the salient revelations derived from this research endeavor.

## 2 Literature review

The objective of this section is to present an exhaustive panorama of the latest advances in research concerning e-grocery predicaments, particularly pertaining to the issues of product availability and the allocation of depots for fulfillment operations. The analytical survey is rooted in the examination of 39 pertinent papers, sourced from the Scopus database, and employs a semi-systematic review methodology. Through the synthesis of insights gleaned from this amalgam of sources, this section furnishes valuable vantages into the existing academic corpus, establishing a bedrock upon which subsequent sections may build for deeper exploration and scrutiny.

### 2.1 Product availability relevant works

As discussed previously, within the realm of e-commerce, the unavailability of a desired product can precipitate dissatisfaction among online customers, potentially prompting them to resort to visiting nearby physical grocery stores to secure the elusive item(s). This transition to in-person shopping, particularly if accomplished via private vehicular transportation, can inadvertently amplify carbon emissions [36]. Thus, the orchestration of product availability assumes a dual significance: not only is it pivotal for ensuring customer contentment, but it also bears weighty implications for the broader spectrum of environmental sustainability.

Hoang and Breugelmans [14] categorize strategies aimed at tackling the intricate challenge of bolstering customer contentment in the face of product unavailability into two distinct groupings: those concentrated on preventing stock-outs and those centered on mitigating the repercussions of stock-outs on customer satisfaction. The former classification encompasses investigations into assortment optimization where Kim and Lennon [19] and Rodriguez Garcia et al. [44] delve into assortment optimization strategies, exploring how the careful selection and arrangement of products can contribute to maintaining product availability in grocery settings. Seidel [47] and Siawsolit and Gaukler [48] focus on the multi-centered fulfillment paradigm,

investigating how distributing fulfillment across multiple centers can enhance product availability in the grocery supply chain. Rodriguez Garcia et al. [44], Lamballais Tessensohn et al. [21], and Milella et al. [33] explore technology-driven interventions, examining how technologies can be leveraged to optimize inventory management and improve product availability. Ulrich et al. [51], Siawsolit and Gaukler [48], and Liu [24] investigate strategies involving the utilization of delivery windows, examining how efficient scheduling and coordination of deliveries contribute to maintaining consistent product availability. Wang and Yang [54] and Ulrich et al. [51] focus on enhancing the prediction of demand, exploring how accurate forecasting can positively impact inventory management and contribute to sustained product availability. And, Marques et al. [28] delve into the application of lean principles, exploring how the principles of lean inventory management can be employed to streamline processes and ensure continuous product availability in grocery operations.

The latter focus on strategies such as substitutions. Seidel [47], Hoang and Breugelmans [14], Kim and Lennon [19], and Breugelmans et al. [5] explore the role of substitutions as a strategy to manage product unavailability in e-grocery settings. Breugelmans et al. [5] investigate the impact of virtual rearrangements of online shelves, studying how the presentation of products can be strategically altered to address stock-outs. Breugelmans et al. [5], Kim and Lennon [19], and Pizzi and Scarpi [38] focus on pre-purchase stockout alerts, examining how providing customers with timely information about product availability influences purchasing decisions. Kim and Lennon [19], Kumar et al. [20], Breugelmans et al. [5], and Peinkofer et al. [37] discuss supplementary availability cues, investigating various methods to signal product availability to online shoppers. Breugelmans et al. [5], Jinzhong and Jian [17], Bhargava et al. [4], and Kim and Lennon [19] explore the use of financial incentives as a strategy to manage and mitigate the impact of product unavailability. Kim and Lennon [19] and Pizzi and Scarpi [38] delve into strategic communication methodologies, studying how effective communication can positively influence customer perceptions and responses to product unavailability.

Nevertheless, despite the considerable body of scholarly work in this field, a gap remains. Notably lacking is an examination of e-grocery fulfillment policies that intricately assign dedicated e-grocery stores to individual customers while concurrently leveraging data from both customers and the e-grocery platform. It is within this gap that our study takes on the challenge of not only exploring the creation of such policies but also rigorously assessing their effectiveness across a range of objectives. The following section offers a detailed exposition of these policies in all their facets.

## 2.2 E-grocery depot allocation policies in order fulfillment related works

The strategic allocation of fulfillment centers stands as a pivotal determinant for effectuating efficient and satisfactory online order completion. With a specialized focus on storing, packaging, and dispatching products from e-commerce platforms, fulfillment centers hold sway over the quality of the consumer

experience. Conventionally, in the retail landscape, the allocation of these centers' hinges on cost efficiency, often prioritizing the proximity of the center to the customer to meet immediate demands. This approach typically seeks to optimize cost structures and delivery timelines [1, 2, 11, 25]. However, it is vital to strike a delicate balance between cost optimization and customer satisfaction, as designating a distant center may entail longer delivery times or higher shipping costs. While the rationale of opting for the nearest fulfillment center for cost efficiency appears logical, it may not invariably constitute the optimal decision-making approach. Ideally, a holistic strategy should account for diverse objectives and trade-offs to ensure both customer satisfaction and cost-effectiveness [27].

Existing research mostly employed mathematical modeling to determine optimal fulfillment center assignment strategies, often emphasizing single objectives. Common objectives encompass minimizing last-mile delivery expenses [1, 42, 50, 57]. Additional factors, such as shipment volumes [16, 56], dimensional attributes [6, 16, 27], and the costs associated with storage and handling [11, 18, 23, 25–27, 52], have also been factored in. Yet, these studies predominantly revolved around optimizing single objectives, without accounting for the intricate interplay and trade-offs among multiple objectives.

Scarce research explores the dual realms of product availability and environmental sustainability in home delivery contexts. While certain studies touched on the influence of delivery times on consumer choices [2, 3, 6, 41, 46]), few delved into ensuring adequate product availability during online shopping. Typically, product availability has been treated as a constraint rather than an optimization objective within the existing body of literature. A handful of studies [27, 41] have explored the repercussions of allocation policies on customer satisfaction. Notably, Mahar et al. [27] explored this within the realm of in-store shoppers, while Razmi et al. [41] addressed customer satisfaction in a broader context.

To bridge this research gap, it becomes important to address both product availability and economic and environmental sustainability considerations in the context of home delivery operations. Product availability is pivotal within the domain of e-commerce, as customers anticipate finding the items they seek while shopping online. As previously mentioned, product unavailability may lead to customer dissatisfaction and potential revenue loss. Furthermore, customers might opt to visit physical stores to obtain missing items when they become aware of unavailability just before the scheduled delivery, which could potentially lead to an increase in carbon emissions if they choose to travel by car [9]. This research aims to fill this gap by examining allocation policies for online fulfillment centers that achieve a dual purpose: enhancing product availability for customers and simultaneously reducing delivery costs and carbon emissions for e-grocery businesses. By effectively balancing these three goals, the study strives to present a holistic framework for enhancing the efficiency of online order fulfillment. This approach aims to enhance overall customer satisfaction while also promoting environmental sustainability.

### 3 Problem statement

In the case of e-grocery shopping, customers typically employ the company's mobile app or website to place their orders, which are subsequently fulfilled by dedicated fulfillment centers responsible for delivering items to specified addresses. During the creation of their shopping baskets, customers often verify product availability through the e-grocery's online platforms. If a desired product isn't accessible at the chosen store, customers may consider switching to another e-grocery provider. Central to this process are challenges stemming from unavailable products that customers intend to purchase, materializing as either pre-purchase out-of-stock or post-purchase out-of-stock scenarios.

This research endeavors to augment product availability for e-grocery customers by crafting and executing fulfillment center dedication policies harnessing both customer and e-grocery relevant data. One avenue for achieving this may involve modifying the allocation of e-depots based on customer buying patterns, with a specific focus on their frequently chosen items. Through the dedication of stores stocked with high-demand products, the likelihood of encountering post-purchase out-of-stock situations can be diminished, potentially boosting customer satisfaction. This implies that the fulfillment center tasked with fulfilling a customer's orders might vary depending on the products being sought, thereby aligning with evolving customer preferences. By enacting a personalized and efficient fulfillment process tailored to immediate needs, the probability of meeting customer expectations and heightening satisfaction is amplified (i.e., social sustainability).

Moreover, this study aims to delve into how these e-grocery dedication policies can concurrently curtail average delivery costs and minimize carbon emissions for the e-grocery entity (i.e., economic and environmental sustainability). To realize this goal, the investigation will delve into pertinent data to be tracked from the e-grocery system, encompassing customer purchasing habits, product demand, delivery distances, inventory levels, and carbon emissions. Through the analysis and application of this data, the study will formulate potential e-grocery dedication policies that grapple with these potentially conflicting objectives, striving to enhance product availability, slash delivery costs, and mitigate environmental impact. The Monte Carlo simulation methodology is employed to replicate and assess the performance of the devised policies across diverse scenarios. This approach permits the examination and evaluation of the policies under assorted conditions, shedding light on their potential advantages and trade-offs. The specifics of the simulation modeling approach, encompassing the data employed, model structure, and performance metrics, will be elaborated upon in Sect. 4 of this study.

### 4 Simulation modelling and utilization of e-grocery data

Effectively harnessing and accurately managing data holds paramount importance in the efficient management of an e-grocery system and in making well-informed choices. The data sourced from the e-grocery system can offer invaluable insights

into customer purchasing behaviors, product demands, delivery distances, and carbon emissions. This data can be utilized to optimize the distribution of fulfillment centers to customers, yielding improved product availability, diminished delivery expenses, and decreased carbon emissions. Moreover, data tracking empowers the e-grocery system to promptly respond to shifts in customer preferences and market dynamics. Through consistent data monitoring and analysis, the system can adapt and refine its operations to align with evolving customer needs and sustain its competitiveness in the market.

Within the scope of this study, the importance of data tracking and its accessibility in the administration of e-grocery systems is emphasized as a critical factor. Additionally, the research aims to pinpoint the specific categories of valuable data to be monitored within the e-grocery system, utilizing this data to establish and evaluate effective e-grocery delivery policies. Adopting a data-driven approach holds the potential to streamline the management of e-grocery systems, facilitate well-informed decision-making, and enhance the overall customer experience.

To enhance product availability, reduce delivery expenses, and mitigate carbon emissions within the e-grocery industry, precise data tracking originating from the e-grocery system is indispensable. Two essential data parameters that warrant meticulous tracking encompass:

- Probability of fulfilling product  $t$  in e-grocery  $i$  ( $pa_{it}$ ): This parameter indicates probability of a specific product being accessible in a given e-grocery store. Monitoring and dissecting this probability affords the ability to pinpoint e-grocery outlets with higher probabilities of fulfilling customer orders.
- Probability of purchasing product  $t$  in the next order for the e-grocer ( $po_t$ ): This parameter illustrates the chances of a customer purchasing a particular product in their forthcoming order. Tracing this probability offers deeper insights into customer preferences and requirements, which can guide decisions regarding inventory management and the allocation of fulfillment centers.

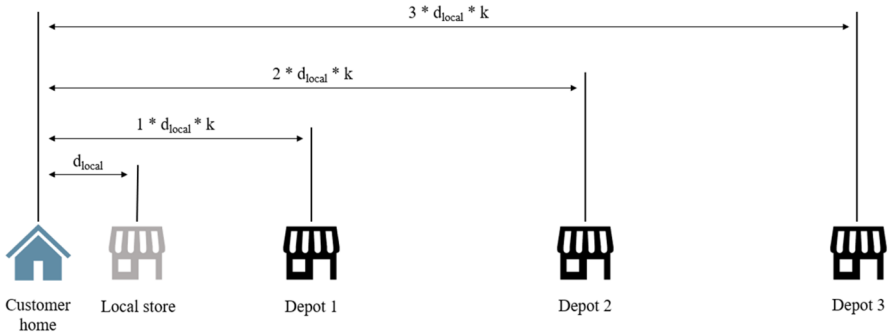
By leveraging probability values derived from historical e-grocery data, it becomes possible to formulate and implement a robust e-grocery assignment policy. For example, a customer's profile can be taken into account, along with cross-referencing the availability of frequently requested items across different fulfillment centers. This approach empowers the e-grocery app to highlight and prioritize the store with the highest probability of stocking the desired products.

The central objective of this research is to create and evaluate specialized allocation policies for e-grocery, incorporating these data-driven methodologies. To assess the effectiveness and impact of these policies, simulation techniques are employed. The simulation utilized in this study is based on the following assumptions and parameters within the e-grocery system:

- **No substitution:** The simulation excludes any form of replacements for unavailable products. Should a product be unavailable, it is not substituted with an alternative.

- **Order repetition:** The simulation involves 500 iterations of recurring orders for a single customer, with each order duplicated independently 50 times.
- **Product Variety:** The e-grocery system provides a diverse assortment of 500 product types that customers can select from when placing their orders.
- **No Minimum Order Requirement:** Customers are not constrained by any minimum order size; they are free to place orders of any magnitude.
- **Customer and Product Classification:** Customer-preferred items are categorized into *XYZ* classes using a Pareto analysis of their order frequency. In parallel, e-grocery products are assigned to *ABC* classes based on their inventory management. In this classification, *X* and *A* denote highly demanded and selling classes, *Y* and *B* indicate moderately demanded and selling classes, and *Z* and *C* signify lowly demanded and selling classes, correspondingly.
- **Randomized values for “ $pa_{it}$ ”:** Stochastic values for the probability of fulfilling a product ( $pa_{it}$ ) are allocated utilizing predefined random distributions. These values encapsulate the likelihood of each product’s availability across the e-grocery depots.
- **Consideration of Visiting Physical Stores:** If the e-grocery system fails to fulfill one or more products, customers retain the option to journey to the nearest physical store to procure those items. The probability of this occurrence is taken into account.
- **Candidate e-grocery depots:** The selection encompasses three potential e-grocery depots for order fulfillment: D1, D2, and D3. D1 is located in closest proximity to the customer, D2 maintains a moderate distance, and D3 is situated at the greatest distance from the customer’s locale.
- **Pre-Determined Depot Allocation:** The assignment of depots to customers is established prior to order placement and remains constant throughout the process of assembling the shopping basket.
- **Distance Calculation Mechanism:** The computation of each depot’s distance (D1, D2, and D3) from the customer’s position entails multiplying the geographical dispersion factor ( $k$ ) with the distance between the local grocery store and the customer’s location ( $d_{local}$ ). These calculated distances facilitate depot assignments to customers. Figure 1 provides a visual representation of this context.
- **Optimized Delivery Logistics:** The conveyance of products is conducted using vans, with delivery routes optimized to accommodate additional orders, thereby maximizing operational efficiency.
- **Handling Multiple Depot Assignments:** In scenarios where more than one depot is assigned for order fulfillment, all items are consolidated within the nearest depot and dispatched in a singular delivery.
- **Key Performance Indicators:** The assessment of system performance is based on the fill rate of customer orders, delivery expenses, and carbon emissions per basket.

The simulation methodology employs the Monte Carlo technique to generate random data, and an experimental design is implemented for sensitivity analysis. This approach empowers the study to evaluate how various parameters influence the performance of the system.



**Fig. 1** Representation of the distances between the customer’s delivery address, the local store, and the depots

### 4.1 Context parameters, design factors, input data

In this section, we outline the parameters, design factors, and input data used in the simulation model.

#### 4.1.1 Context parameters

Within the simulation, certain parameters are designated as predefined variables that form the foundation of the simulation models. These variables retain unchanging values across the simulation runs and are denoted as fixed variables or context parameters. The ensuing explanations outline the meanings attributed to these parameters:

- $(a, b, c)$ : We designate  $a$ ,  $b$ , and  $c$  as 20%, 30%, and 50%, respectively, indicating that 20%, 30%, and 50% of the e-grocery products belong to classes  $A$ ,  $B$ , and  $C$ , respectively.
- $(L_A, L_B, L_C)$ : This represents the low uniform probability scenario for the availability of class  $A$ ,  $B$ , and  $C$  products, respectively. For example, we set  $L_A = [70\%, 90\%]$ ,  $L_B = [70\%, 90\%]$ , and  $L_C = [70\%, 90\%]$ . This means that at the start of the simulation, a random availability probability for each product class is generated from a uniform distribution of [70%, 90%].
- $(M_A, M_B, M_C)$ : This corresponds to the moderate uniform probability scenario for the availability of class  $A$ ,  $B$ , and  $C$  products, respectively. For instance, we set  $M_A = [95\%, 100\%]$ ,  $M_B = [90\%, 95\%]$ , and  $M_C = [80\%, 90\%]$ . This means that at the beginning of the simulation, a random availability probability for each product class is created from uniform distributions of [95%, 100%], [90%, 95%], and [80%, 90%], respectively.
- $(H_A, H_B, H_C)$ : This represents the high uniform probability scenario for the availability of class  $A$ ,  $B$ , and  $C$  products, respectively. For example, we set  $H_A = [99\%, 100\%]$ ,  $H_B = [85\%, 95\%]$ , and  $H_C = [70\%, 85\%]$ . This means that at

the start of the simulation, a random availability probability for each product class is generated from uniform distributions of [99%, 100%], [85%, 95%], and [70%, 85%], respectively.

- $(x, y, z)$ : These denote the percentages of products belonging to classes  $X$ ,  $Y$ , and  $Z$ , respectively, which are defined based on customer ordering frequency within the entire product assortment. We set  $x$ ,  $y$ , and  $z$  as 20%, 30%, and 50%, respectively. It is important to note that classes  $X$ ,  $Y$ , and  $Z$  do not represent the same product groups as classes  $A$ ,  $B$ , and  $C$  in e-grocery. While the latter is defined based on inventory value, the former is defined based on customer preferences.
- $(P_X, P_Y, P_Z)$ : These represent the probabilities of ordering from class  $X$ ,  $Y$ , and  $Z$ , respectively, for an online customer. Here, we set  $P_X$ ,  $P_Y$ , and  $P_Z$  to 80%, 15%, and 5%, respectively. For example, a customer orders from class  $X$ ,  $Y$ , and  $Z$  with probabilities of 0.80, 0.15, and 0.05, respectively.
- $c$ : This denotes the delivery cost per kilometer, taking into account driver and fuel-related expenses. The value of  $c$  is set to 1.29 €/km [52].
- $e_{\text{van}}$ : This represents the amount of emissions in CO<sub>2</sub> equivalent per kilometer traveled during delivery. We assume a diesel commercial van, and  $e_{\text{van}}$  is set to 0.23156 kgCO<sub>2</sub>e/km [7].
- $e_{\text{car}}$ : This signifies the amount of emissions in CO<sub>2</sub> equivalent per kilometer traveled by the customer's car.  $e_{\text{car}}$  is set to 0.17067 kgCO<sub>2</sub>e/km [7].
- $d_{\text{local}}$ : This represents the distance between the customer's location and their local store in an urban context. For the purpose of this simulation,  $d_{\text{local}}$  is set to 2 km, based on the study by Siragusa and Tumino [49].

#### 4.1.2 Design factors

In this section, we clarify the design factors that have been considered in the experimental design. It's important to note that, in order to address RQ2, we have adopted a full factorial design approach. This approach aims to minimize potential biases that could arise from the use of hypothetical data. Specifically, we have incorporated two levels for each of the four design factors, which are delineated as follows:

- **XYZ/ABC relations:** In this design factor, we consider two levels: a high level and a low level. In the high-level, there is a perfect overlap between customer preferences ( $X$ ,  $Y$ ,  $Z$  classes) and the inventory value ( $A$ ,  $B$ ,  $C$  classes) of the grocery store. High level means that the types of products the customer's order ( $X$ ,  $Y$ ,  $Z$  classes) align perfectly with the inventory classes ( $A$ ,  $B$ ,  $C$  classes) of the grocery store. However, in the low-level scenario, there is an imperfect overlap between customer preferences and inventory classes. Specifically, the customer's most frequently ordered  $X$  and  $Y$  classes align with the grocery store's  $C$  class products, while the customer's  $Z$  class products align with  $A$  and  $B$  class products at a ratio of 40% and 60%, respectively. This mismatch between customer preferences and inventory classes introduces a discrepancy in the product assortment available to customers. Table 1 shows those relations.
- **$k$ —geographical dispersion factor:** As elaborated in Sect. 3, the simulation accounts for the geographical distribution of customers and fulfillment depots by

**Table 1** Levels of *XYZ/ABC* factor levels

	A	B	C
Low			
X	0	0	100%
Y	0	0	100%
Z	40%	60%	0
High			
X	100%	0	0
Y	0	100%	0
Z	0	0	100%

introducing a multiplier factor, denoted as  $k$ , as visualized in Fig. 1. The magnitude of  $k$  defines the separation between two points, where a higher  $k$  value signifies an increased distance. Precisely, the distance between a customer and the local physical store is represented as  $d_{local}$  in Fig. 1. Within the experimental design, two distinct levels of  $k$  are under scrutiny, accommodating diverse geographical dispersion scenarios for depots:

*Low level:*  $k = 1$ , indicating low geographical dispersion of the depots.

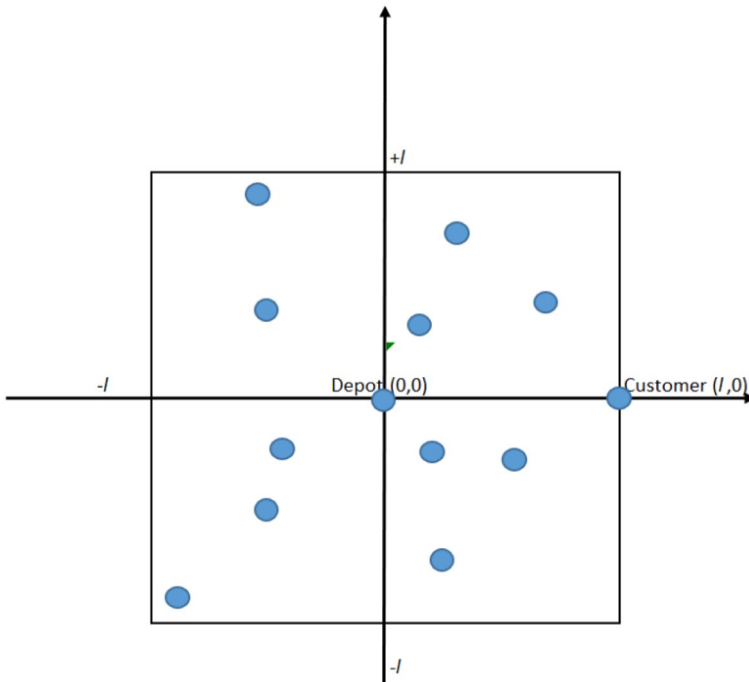
*High level:*  $k = 5$ , indicating high geographical dispersion of the depots.

- $d\%$ : delivery trip inefficiency:** To estimate the cost and carbon emissions of delivery in a more realistic manner, the experimental design incorporates the concept of a complete tour, where a customer’s order is delivered as part of a larger set of orders for multiple customers. This means that multiple orders are delivered together, optimizing the delivery route and reducing the overall distance traveled. To account for this, the delivery distance for a single basket is calculated by (1), taking into consideration the combined distance of all the orders included in the tour.

$$customer\ point - depot\ location \times d\% \tag{1}$$

where the first expression represents the distance between the customer point and the distribution depot, and  $d\%$  denotes a measure of delivery trip inefficiency. The value of  $d\%$  reflects the degree of inefficiency in the delivery trip, with a higher percentage indicating a greater level of inefficiency and resulting in a longer travel distance for the delivery tour assigned to the customer point. In the experimental design, two levels of  $d\%$  are taken into account to accommodate different levels of delivery trip inefficiency.

In the high-level determination of  $d\%$ , the experimental design takes into account that, on average, there are twelve orders to be delivered in a tour. All the delivery addresses are located within a square area, as depicted in Fig. 2. The customer point is located at a distance of  $l$  from the central depot. To find the average



**Fig. 2** Random generation of 11 delivery addresses in an area  $2l \times 2l$

tour travel distance, a traveling salesman optimization solution is employed. To accomplish this, eleven delivery addresses are randomly generated within the square area, as shown in Fig. 2. The location of the customer point remains fixed and is  $l$  units away from the depot. The traveling salesman optimization procedure is then applied to determine the optimal tour length that begins at the depot, visits all twelve delivery addresses, and concludes at the starting depot. The tour length is divided by twelve to estimate the distance covered by a single household within that tour. Subsequently,  $d\%$  is calculated by dividing the distance portion by  $l$ , representing the distance between the customer point and the distribution depot. By repeating this experiment 50 times, a reasonably accurate estimation of  $d\%$  is obtained, which is determined to be 53.5%. Consequently, the high level of  $d\%$  is set at 53.5% in the experimental design.

- *High level:*  $d\% = 53.5\%$
- *Low level:*  $d\% = 16.7\%$

In the low-level calculation of  $d\%$ , the experimental design assumes that in the journey to reach the customer's delivery address, all the other eleven delivery addresses are arranged in a straight line, as depicted in Fig. 3. Consequently, the optimal tour length amounts to  $2l$ , considering the round trip. Dividing this value

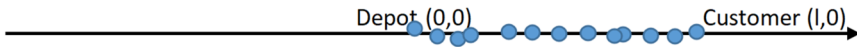


Fig. 3 Possible locations of delivery addresses in the low-level situation

by the number of customers (i.e., 12), we determine that the distance associated with the delivery of a single basket corresponds to  $(2/12) \times 100 = 16.7\%$ .

- **$g\%$ : probability of a customer engaging in complementary shopping:** This factor represents the likelihood of a customer visiting their local grocery store to purchase any unavailable items when one or more products from their order are not in stock. The experimental design considers different levels for this factor, which are outlined as follows:

- *Low level:*  $g\% = 20\%$
- *High level:*  $g\% = 80\%$

#### 4.1.3 Historical input data

While formulating depot allocation policies, our approach relies on the utilization of historical data. This encompasses the likelihood of product availability at a specific depot during the fulfillment stage, as well as a customer's past purchasing history categorized according to product types. In a practical context, such data can be conveniently traced and retrieved from e-grocery application software. Subsequently, we will delineate the two specific categories of data that necessitate monitoring and recording:

1. *Probability of Product Availability during Fulfillment:* This dataset offers valuable insights into the probability of a product being accessible at a particular depot during the process of fulfilling customer orders. This information aids in assessing the readiness and feasibility of obtaining products from different depots.
2. *Customer Purchase History Segmented by Product Type:* This dataset records the historical buying patterns of customers, organized according to distinct product types. It facilitates the scrutiny of customer preferences, thereby enabling well-informed choices concerning product assortment and allocation tactics across different depots.

*Likelihood of product availability during fulfillment* The probability of product availability during the fulfillment stage reflects the frequency with which products can be successfully obtained for fulfilling customer orders. As outlined in Sect. 4, this probability is represented as  $pa_{it}$ , indicating the likelihood of satisfying a customer's request for product  $t$  from e-grocery  $i$ . Conversely, the complementary probability,  $1 - pa_{it}$ , signifies the chance of product  $t$  being unavailable at depot  $i$  during the delivery phase. As elucidated in Sect. 4.1.1, the simulation process generates a random probability of product availability from uniform distributions as specified in

**Table 2** Ranges for  $pa_{it}$  generation in simulation

	Class A	Class B	Class C
Scenario <i>L</i>	[70%,90%]	[70%,90%]	[70%,90%]
Scenario <i>M</i>	[95%,100%]	[90%,95%]	[80%,90%]
Scenario <i>H</i>	[99%,100%]	[85%,95%]	[70%,85%]

Table 2. These distribution ranges are established according to the product classes: *A*, *B*, and *C*. For instance, referring to Table 2, in the scenario of moderate availability, a random availability probability is generated within the range of [0.8, 0.9] for each product falling under class *C*.

Within the simulation model, a random probability value is generated for the parameter “ $pa_{it}$ ” from the defined ranges outlined in Table 2. Once this value is generated, it remains constant throughout the entire duration of the simulation run until its completion.

*Customer buying patterns per product category* The customer’s purchasing history categorized by product type is a representation of the likelihood that product “*i*” will be added to the customer’s upcoming basket. This probability is denoted as “ $po_{it}$ ,” as defined in Sect. 4 and is assumed to be derived from the customer’s past buying behavior. Various methodologies can be explored to ascertain this probability [12]. One of the direct and meaningful approaches involves calculating it based on the historical frequency of the product’s purchase by the customer.

According to Rivière et al. [42], product popularity tends to follow a Pareto distribution. Within this distribution, the pool of items is divided into three classes: *X*, *Y*, and *Z*, categorized according to their ranked frequencies. Specifically, in the customer’s order, 20% of the items fall into class *X* and are bought 80% of the time, while classes *Y* and *Z* encompass 30% and 50% of the items, respectively, with purchase rates of 15% and 5%. For simplicity, it’s assumed that within each class, the likelihood of purchasing any particular product is uniform.

## 4.2 Depot allocation policies

In this section, we present the developed depot allocation policies:

### 4.2.1 Policy C: Proximity-based single depot assignment

This basic and widely adopted approach for allocating a fulfillment depot based on proximity involves designating *D1* as the primary fulfillment center for the e-grocery customer.

### 4.3 Policy A: Maximum availability depot assignment

Under Policy A, this approach entails allocating a depot to the customer based on the depot’s highest product availability score. The availability score ( $AS_i$ ) for each depot “*i*” within the set of depots “*D*” is computed using Eq. (2):

$$AS_i = \sum_t p_{o_t} \times p_{a_{it}}, \quad \forall i \in D \tag{2}$$

where (2) computes the score by multiplying the probability of the customer in question placing an order for product “*t*” and the probability of depot “*t*” fulfilling that specific product “*t*” in the e-grocery context. In this approach, the depot “*t*” that holds the maximum  $AS_i$  value is designated to cater to the customer’s requirements.

### 4.3.1 Policy S: Maximum combined score depot assignment

This strategy revolves around assigning a depot based on the highest combined score, as illustrated in Eq. (3):

$$CS_i = \alpha \times \frac{AS_i - \min_i(AS_i)}{\max_i(AS_i) - \min_i(AS_i)} + (1 - \alpha) \times \frac{d_i - \min_i(d_i)}{\max_i(d_i) - \min_i(d_i)} \tag{3}$$

Given Eq. (3), the process commences by calculating the availability score ( $AS_i$ ) for each depot “*t*” using Eq. (2). Subsequently, the combined score ( $CS_i$ ) for depot “*t*” is ascertained by taking the weighted summation of the normalized  $AS_i$  and the normalized distance ( $d_i$ ) between the depot and the customer. The process of normalization employs linear scaling. In this context, “ $d_i$ ” signifies the geographical distance in kilometers between depot “*t*” and the customer’s location.

The coefficient “ $\alpha$ ” is a weighting factor, ranging between 0 and 1. This coefficient captures the balance between enhancing product availability and minimizing travel distances. When  $\alpha=0$ , the primary emphasis is placed on optimizing delivery costs and reducing emissions. Conversely, when  $\alpha=1$ , the top priority is granted to ensuring product availability. In the simulation models,  $\alpha$  is set at 0.5, thus ensuring that equal significance is attributed to both product availability and distance considerations. Following the computation of  $CS_i$  values, the depot holding the highest  $CS_i$  score is selected to fulfill the customer’s order.

### 4.3.2 Policy 2C: Proximity-based dual depot assignment

Implemented as Policy 2C, this approach involves designating the two nearest fulfillment depots, namely D1 and D2, to handle the customer’s delivery. This strategy offers visibility into the inventories of both D1 and D2 to the customer. In instances where the ordered products are present at D1, a consolidation takes place at this depot to facilitate a singular delivery. However, in cases where the products are absent at D1, no consolidation occurs, and the items are dispatched individually from D2.

### 4.3.3 Policy 2A: Dual high availability depot assignment

Policy 2A entails assigning two depots characterized by the highest combined availability scores for the products in the customer’s order. The merged

availability score, denoted  $aMAS_{iz}$ , is computed for the pairing of two depots, “ $i$ ” and “ $z$ ,” chosen from the depots in the set  $D$ , as outlined in Eq. (4):

$$S_{iz} = \sum_t po_t \times ((1 - (1 - pa_{it})) \times (1 - pa_{zt})) \quad \forall i, z \in D, i \neq z \quad (4)$$

It’s important to note that in (4), the variables have specific meanings: “ $pa_{it}$ ” corresponds to the probability of fulfilling product “ $t$ ” in e-grocery “ $i$ ,” “ $pa_{zt}$ ” signifies the probability of fulfilling product “ $t$ ” in e-grocery “ $z$ ,” and “ $po_t$ ” represents the probability of the customer in question ordering product “ $t$ ”. Taking into account the highest value of  $MAS_{iz}$ , the two depots, “ $i$ ” and “ $z$ ,” which exhibit the greatest combined availability score, are designated to serve the customer.

#### 4.3.4 Policy 2S: Dual high combined score depot assignment

Under Policy 2S, the focus is on designating two depots with the highest merged combined scores. These scores result from considering both product availability and minimizing distance. The computation procedure is detailed in Eq. (5):

$$MCS_{iz} = \alpha \times \frac{MAS_{iz} - \min_{iz}(MAS_{iz})}{\max_{iz}(MAS_{iz}) - \min_{iz}(MAS_{iz})} + (1 - \alpha) \times \frac{d_i \times d\% + d_{iz} - \min_{iz}(d_i \times d\% + d_{iz})}{\max_{iz}(d_i \times d\% + d_{iz}) - \min_{iz}(d_i \times d\% + d_{iz})}, \quad \forall i, z \in D, i \neq z \quad (5)$$

As per Eq. (5), the process begins by calculating the merged availability score ( $MAS_{iz}$ ) using Eq. (4). Subsequently, the merged combined score ( $MCS_{iz}$ ) for depots “ $i$ ” and “ $z$ ” within the set “ $D$ ” is determined. This computation involves the weighted summation of the normalized  $MAS_{iz}$  and the normalized travel distance required for order fulfillment ( $d_i \times d\% + d_{iz}$ ). Both of these quantities are normalized using linear scaling, with the weighting factor “ $\alpha$ ” set to 0.5. It’s worth noting that “ $d_{iz}$ ” represents the distance between depot “ $i$ ” and depot “ $z$ ,” while “ $d\%$ ” signifies the inefficiency factor of the delivery trip.

Ultimately, the two allocated depots are depots “ $i$ ” and “ $z$ ” within the set “ $D$ ” that possess the highest  $MAS_{iz}$  score.

#### 4.3.5 Policy 3C: Triad of closest depots assignment

Operating under Policy 3C, this strategy involves allocating the three depots that exhibit the closest proximity to the customer’s delivery address for fulfilling their order. As a result, the customer’s order is handled by Depot 1, Depot 2, and Depot 3.

To assess the effectiveness of each policy, we conduct simulations within the e-grocery system and evaluate their performance based on the subsequent metrics:

- *Fill rate*: This metric gauges the proportion of ordered products that are available for the customer's order, reflecting how effectively the requested items are fulfilled.
- *Delivery cost per basket*: This metric quantifies the cost linked with delivering each individual customer basket, considering variables like travel distance and delivery inefficiencies.
- *Carbon emission amount per basket*: This metric computes the environmental impact of delivering each customer basket, quantifying it in terms of carbon emissions.

In Sect. 4.3, we offer a comprehensive explanation of the employed simulation model and the methodology employed for calculating these performance metrics.

#### 4.4 Simulation model

This section offers an in-depth overview of the simulation framework and the configurations employed in the simulation process. Initially, we establish "*ru*" as the replication number, indicating the frequency at which the simulation is replicated.

##### 4.4.1 Product availability configurations

To comprehensively explore the impact of product availability on e-grocery assignment policies, we take into account that each depot can fall under one of three distinct product availability scenarios (*L/M/H*), elaborated in Sect. 4.1.3.1. To thoroughly assess the efficacy of the formulated policies and appraise their performance across a spectrum of product availability levels prevalent in the e-grocery domain, we define six distinct product availability configurations as outlined in Table 3. This method enables us to scrutinize policy performance across varying availability conditions and ascertain their applicability in specific scenarios. The allocation of different depots to diverse product availability scenarios facilitates an all-encompassing assessment of policy efficiency and offers insights into their performance across different availability configurations.

As indicated in Table 3, the initial product availability configuration entails investigating the situation where Depot 1, Depot 2, and Depot 3 possess product availability levels denoted as *L* (low), *M* (medium), and *H* (high) correspondingly. It's crucial to acknowledge that the precise probability ranges associated with these

**Table 3** Product availability settings for depots

	Setting 1	Setting 2	Setting 3	Setting 4	Setting 5	Setting 6
Depot 1	<i>L</i>	<i>L</i>	<i>M</i>	<i>H</i>	<i>M</i>	<i>H</i>
Depot 2	<i>M</i>	<i>H</i>	<i>L</i>	<i>L</i>	<i>H</i>	<i>M</i>
Depot 3	<i>H</i>	<i>M</i>	<i>H</i>	<i>M</i>	<i>L</i>	<i>L</i>

availability levels are established in alignment with the predefined product classes, as elucidated in Table 2.

#### 4.4.2 Depot allocation

In the simulation run  $ru$ , the customer is assigned one or multiple depots in accordance with the chosen policy. The collection of depots allocated for that specific simulation run, designated as  $A_{ru}$ , signifies the depots designated to cater to the fulfillment of the customer's order.

#### 4.4.3 Basket generation

The generation process for the customer's order (basket) adheres to a methodology akin to that of Rivière et al. [42], encompassing the subsequent two steps:

*Step 1: Basket Size Generation:* The initial step involves generating the size of the customer's basket. Drawing inspiration from observations in physical grocery channels [29], where the distribution of basket sizes follows a truncated negative binomial distribution, a similar assumption is extended to the realm of e-grocery while disregarding any minimum order constraints. To establish the average basket size, literature presents divergent findings due to variations in grocers and contexts. In this simulation model, we draw on the insights of Yuan et al. [58], who scrutinized a Walmart Online Grocery database case. Their outcomes are adapted into a truncated negative binomial distribution, yielding an average basket size of 21.4 unique items, visually depicted in Fig. 4. This probability distribution is subsequently employed to generate the basket size " $N_{ru}$ " for each individual simulation run, " $ru$ ".

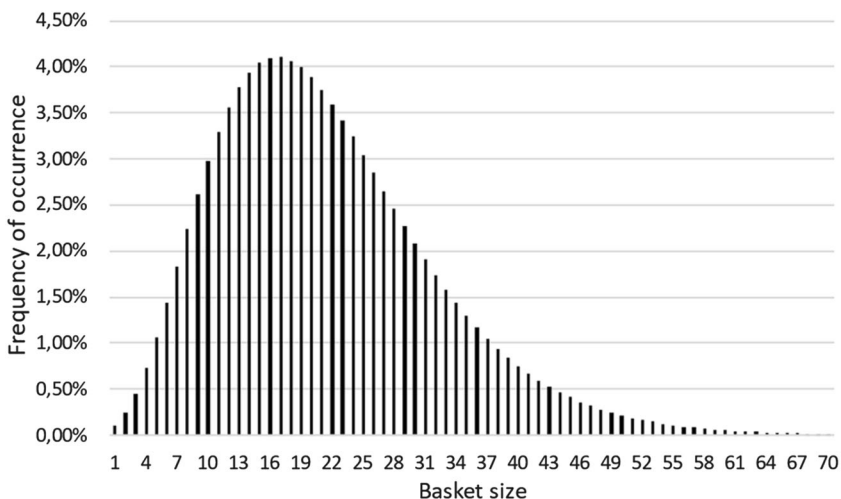


Fig. 4 Basket size frequency distribution for basket generation in the simulation

*Step 2: Product Type Generation in the Basket:* Generation of product types in the basket is the next step. With a total assortment of 500 products, each product has a probability ( $p_{ot}$ ) of being included in the customer's basket. The number of product types generated corresponds to the number of basket sizes ( $N_{ru}$ ), assuming that one product is ordered per product type.

#### 4.4.4 Product availability status generation

Subsequent to basket size determination, the process proceeds to generate the product types to be included in the basket. In consideration of a comprehensive product assortment encompassing 500 items, each individual product holds a distinct probability ( $p_{oi}$ ) of being chosen for inclusion in the customer's basket. The count of generated product types aligns with the determined basket size ( $N_{ru}$ ), assuming a one-to-one correspondence between product types and the quantity of items in the basket.

#### 4.4.5 Computation of Performance Metrics

The evaluation of the online grocery system's performance metrics is accomplished through the following methodology:

**Fill rate** is computed by (6):

$$Fill\ rate_{ru} = \frac{\sum_{t \in B_{ru}} \max_{i \in \Lambda_{ru}} (AS_{it_{ru}})}{N_{ru}} \tag{6}$$

As per Eq. (6), the fill rate for the particular order placed in simulation run "ru" is determined by calculating the proportion of products within the order, denoted as " $B_{ru}$ ," that are accessible within the assigned depots " $i$ " belonging to the set " $\Lambda_{ru}$ ." This result is divided by the order size " $N_{ru}$ ." The term  $\max_{i \in \Lambda_{ru}} (AS_{it_{ru}})$  signifies the highest availability status of product " $i$ " across all allocated depots within " $\Lambda_{ru}$ ." In cases where product " $i$ " is available in at least one of the allocated depots,  $\max_{i \in \Lambda_{ru}} (AS_{it_{ru}})$  is equivalent to 1; conversely, it is set to 0 if the product is unavailable across all assigned depots.

**Delivery cost per basket (DCB)** is computed by (7):

$$DCB_{ru} = c \times (d_i \times d\% + pick\_in\_middle_{ru} \times d_{ij} + pick\_in\_farthest_{ru} \times d_{iz}) \tag{7}$$

where  $i, j, z \in \Lambda_{ru}$  and  $i < j < z$ . The delivery cost per basket in the simulation run ru ( $DCB_{ru}$ ) is computed by multiplying the delivery cost per kilometer ( $c$ ) by the total distance traveled. When two or more depots are designated for order fulfillment, the ordered products are consolidated at the nearest depot. As a result, the traveled distance comprises two elements: the distance covered for home delivery and the distance associated with transshipment between depots. The home delivery distance is determined by multiplying the distance " $d_i$ " between the nearest assigned depot " $i$ "  $\in \Lambda_{ru}$  and the customer by the delivery tour inefficiency " $d\%$ ." This factor accounts for any inefficiencies in the delivery route. If an intermediate depot " $j$ " is

allocated for the customer order and at least one product is selected from there, the binary variable “ $pick\_in\_middle_{ru}$ ” is set to 1; otherwise, it’s set to 0. Consequently, the transshipment distance “ $d_{ij}$ ” between the closest depot “ $i$ ” and the intermediate depot “ $j$ ” is also incorporated. Similarly, if the farthest depot “ $z$ ” is assigned, and at least one product is selected from there, the binary variable “ $pick\_in\_farthest_{ru}$ ” is set to 1; otherwise, it’s set to 0. Consequently, the travel distance “ $d_{iz}$ ” between the closest assigned depot “ $i$ ” and the farthest depot “ $z$ ” is also factored into the calculation of the delivery cost per basket.

Emission amount per basket (EB) is computed by (8):

$$EB_{ru} = e_{van} \times (d_i \times d\% + pick\_in\_middle_{ru} \times d_{ij} + pick\_in\_farthest_{ru} \times d_{iz}) + go\_buy_{ru} \times e_{car} \times 2 \times d_{local} \quad (8)$$

The emissions per basket in simulation run “ $ru$ ” are determined by two distinct components: emissions associated with the delivery process and emissions linked to complementary shopping endeavors. The computation of emissions tied to the delivery process involves multiplying the emissions per kilometer for a van, denoted as “ $e_{van}$ ,” by the total travel distance. The calculation of the travel distance parallels the method employed for determining the delivery cost per basket. It incorporates both the home delivery distance and any transshipment distances between depots.

Incorporating emissions from complementary shopping, we introduce a binary variable “ $go\_buy_{ru}$ .” When “ $go\_buy_{ru}=1$ ,” it signifies that the customer engages in complementary shopping. Conversely, when “ $go\_buy_{ru}=0$ ,” no emissions are attributed to this activity. In simulation run “ $ru$ ,” there exists a probability “ $g\%$ ” that “ $go\_buy_{ru}=1$ .” In cases where the customer undertakes complementary shopping (i.e., “ $go\_buy_{ru}=1$ ”), it’s assumed they travel to a local store by car, completing a round trip. The computation of emissions tied to complementary shopping entails multiplying the emissions per kilometer for an average car, denoted as “ $e_{car}$ ,” by a factor of two (round trip), and subsequently by the average distance separating the customer from a local grocery store, represented as “ $d$ ”.

#### 4.5 Monte Carlo simulation and replications

To thoroughly evaluate the performance of each policy, a Monte Carlo simulation is conducted utilizing MS Excel. This simulation encompasses 500 repetitive orders to ensure robust and comprehensive assessment. Within each Monte Carlo replication, the establishment of steady-state conditions is achieved by calculating the progressive mean of every performance metric across each run.

The warm-up period, which signifies the initial transient phase, is identified through visual analysis of performance metric graphs, following the methodology outlined by Hoad et al. [13]. Subsequently, the outcomes from the warm-up runs are excluded from the analysis. The results for the replication are then computed by averaging the remaining runs. This process mirrors the approach adopted by Sandıkçı and Sabuncuoğlu [45].

The Monte Carlo simulation is replicated 50 times in total. To acquire the final outcomes for each performance metric, averages of the results from the

**Table 4** Levels of the design factors

Factors	Low level (0)			High level (1)				
XYZ/ABC		<i>A</i>	<i>B</i>	<i>C</i>		<i>A</i>	<i>B</i>	<i>C</i>
	<i>X</i>	0	0	100%	<i>X</i>	100%	0	0
	<i>Y</i>	0	0	100%	<i>Y</i>	0	100%	0
	<i>Z</i>	40%	60%	0	<i>Z</i>	0	0	100%
k	1			5				
d%	16.7%			53.5%				
g%	20%			80%				

replications are computed, accompanied by 95% confidence intervals. This comprehensive methodology offers a dependable and accurate assessment of the performance exhibited by each policy.

#### 4.6 Simulation flow charts

To visually show the algorithmic sequence of the simulation models, a flow chart is presented, drawing inspiration from the model outlined in Appendix A. The purpose of this flow chart is to offer a graphical representation of the step-by-step progression followed in each simulation experiment. This visualization enhances comprehension of the algorithmic workflow.

This flow chart is executed for each simulation experiment, representing various combinations of factor levels and policies. Table 4 furnishes a concise summary of the four design factors, each spanning two levels. Consequently, a total of 16 ( $= 2^4$ ) distinct simulation experiments are conducted. When accounting for the inclusion of seven diverse assignment policies, the cumulative number of experiments reaches 112 ( $= 16 \times 7$ ).

### 5 Results and discussion

In this section, we compile a comprehensive summary encompassing the outcomes and ensuing discussions. Furthermore, we delve into the verification and validation processes associated with the models utilized in the study.

## 5.1 Verification and validation of the model

A significant element of any simulation study is the verification and validation of the simulation model [43]. We explain the application procedures of them in this section.

### 5.1.1 Verification

The simulation model undergoes a comprehensive verification process to ensure its accuracy. This involves conducting thorough computational checks to validate the correctness of the model's algorithms and calculations. Additionally, the simulation results are carefully examined by manipulating input variables to confirm that the allocated depot(s) align with the selected allocation policy. For example, when implementing policy C, it is expected that D1 is always allocated, whereas policy A assigns an equal probability to each depot. By verifying these allocation patterns, we can ensure that the model consistently follows the intended policies. Furthermore, the model's correctness is evaluated by observing how the allocated depot(s) change when the parameter  $\alpha$  is adjusted for policies S and 2S. According to the conceptual model, as  $\alpha$  transitions from 0 to 1, the distribution of allocated depots should shift from C/2C to A/2A. This analysis validates whether the model accurately reflects the expected behavior.

Once the verification process is successfully completed, the model enters the validation phase. During validation, the model's outputs are compared with real-world data or established knowledge to confirm its reliability and fidelity to the actual system being simulated. By undertaking these verification and validation procedures, we can confidently assert the accuracy and reliability of the simulation model. This enables us to utilize the model effectively for studying and analyzing e-grocery systems, knowing that it faithfully represents the allocation policies and exhibits the expected behavior.

### 5.1.2 Validation

The validation of the model is conducted in the absence of a real system, utilizing the expertise of two experienced individuals proficient in simulation modeling and e-grocery implementations. These experts carefully review the critical performance metrics, such as cost per basket, considering both the available literature and their own extensive knowledge in the field. Through this validation process, the model's outputs are evaluated for their accuracy and alignment with established benchmarks.

After successfully completing the validation phase, the model proceeds to the experimentation stage. A total of 50 separate and independent replications are conducted, aimed at ensuring the robustness and credibility of the obtained results. These replications are pivotal in establishing the reliability of the findings.

The ensuing sub-sections meticulously dissect and analyze the outcomes of these experiments. This analysis provides a holistic understanding of the performance and efficiency exhibited by the e-grocery system within a multitude of scenarios. The diverse scenarios under consideration enable the exploration of how the

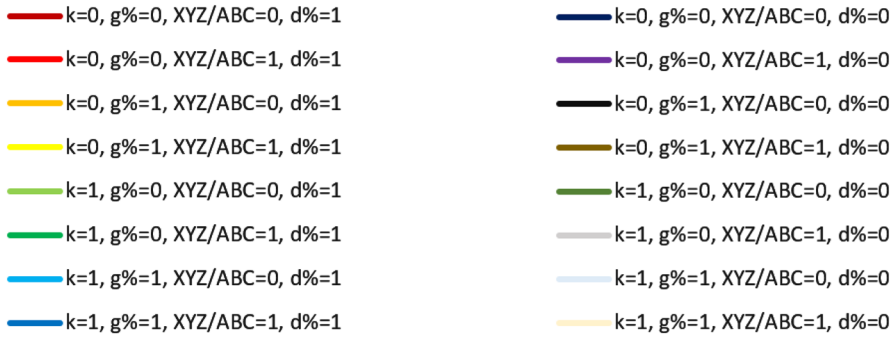


Fig. 5 Description of each experiment

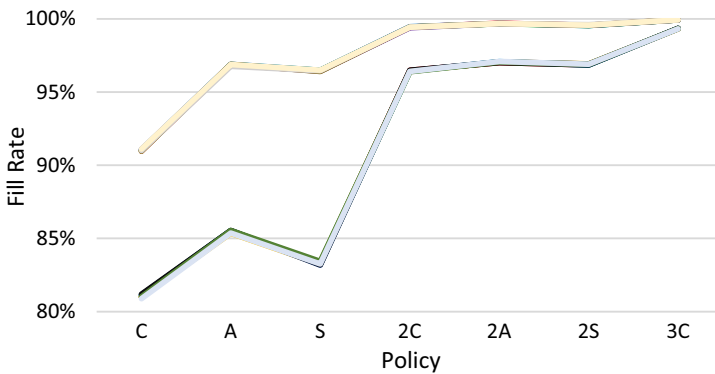


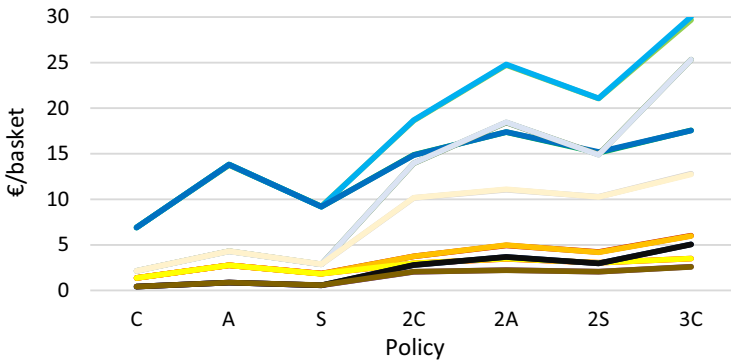
Fig. 6 Policy versus fill rate

system operates under varying conditions, shedding light on its adaptability and effectiveness.

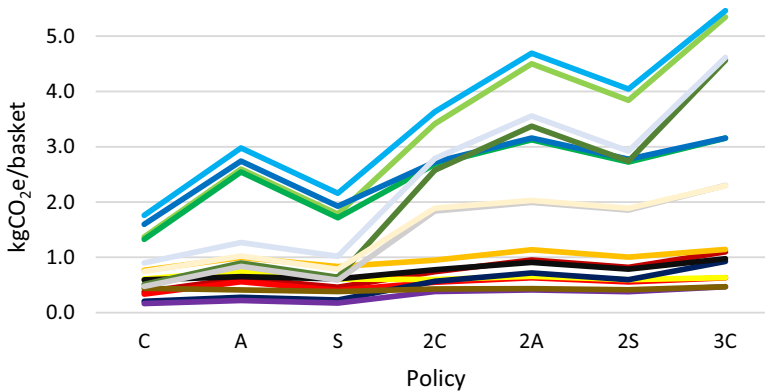
### 5.2 Simulation results

Figures 5, 6, 7 and 8 encapsulate the simulation results for the diverse combinations outlined in Table 3. Figure 5 visually encapsulates each experimental configuration, while Figs. 6, 7 and 8 offer graphical representations elucidating the interplay between assignment policy and three pivotal metrics: fill rate, delivery cost per basket, and emissions per basket. Our observations yield the following key insights:

- The Policy C: Proximity-Based Single Depot Assignment yields the lowest fill rate output, whereas the Policy 3C: Triad of Closest Depots Assignment results in the highest fill rate.
- Notably, in scenarios where there is substantial alignment between customer preferences and grocery inventory classes (high XYZ/ABC level), the fill rate values demonstrate a propensity to be higher. In contrast, scenarios character-



**Fig. 7** Policy versus delivery cost per basket



**Fig. 8** Policy versus emissions per basket

ized by a limited alignment (low XYZ/ABC level) tend to yield lower fill rate values. This highlights the significance of meticulous inventory management strategies tailored to the unique preferences of e-grocery customers. Such tailored approaches hold the potential to substantially enhance fill rate performance.

- In terms of delivery cost per basket, the Policy C: Proximity-Based Single Depot Assignment consistently generates the lowest cost, aligning with expectations. Additionally, Policy S: Maximum Combined Score Depot Assignment, which amalgamates both product availability and distance metrics, consistently produces relatively favorable cost and carbon emissions values.
- A comparison between the fill rate performances of the “S” and “C” policies reveals the superior performance of the “S” policy. As such, adopting the “S” policy holds the potential to enhance all three-performance metrics concurrently.
- The trends observed in carbon emissions align closely with those of cost performance.

- For e-grocery enterprises targeting a fill rate performance exceeding 95%, the Policy 2S: Dual High Combined Score Depot Assignment emerges as a viable option. Notably, this policy, which leverages both depot availability and distance metrics, not only achieves impressive fill rate performance but also delivers relatively favorable cost outcomes.

Given a multi-objective perspective, the “S” and “2S” policies appear promising. Subsequently, their performances are further scrutinized by varying the  $\alpha$  value— a representation of the weight attributed to product availability in combined score calculations. The results of this assessment are succinctly summarized in Sect. 5.2.1.

In summary, the comprehensive analysis underscores the potential advantages associated with the “S” and “2S” policies. These policies exhibit the capability to concurrently enhance multiple performance metrics, rendering them valuable choices for e-grocery enterprises aiming to optimize their operational outcomes.

### 5.2.1 Policy S under different $\alpha$ scenarios

Within this section, our focus centers on the experiments executed under the ambit of the “S” policy, with a deliberate manipulation of the  $\alpha$  values. We endeavor to assess the ramifications of altering  $\alpha$  values, specifically exploring the values of 0, 0.25, 0.5, 0.75, and 1. The outcomes stemming from these experiments are meticulously detailed and showcased in Figs. 9 through 12. Figure 9 contributes a comprehensive depiction of the experiments carried out within the scope of the “S” policy, visually outlining the array of  $\alpha$  values subjected to testing.

Figure 10 shows  $\alpha$  values vs. fill rate relationship. The graphical representation in Fig. 10 provides insight into the interplay between varying  $\alpha$  values and the corresponding fill rate outcomes. It is discernible that as the  $\alpha$  value escalates, there is a discernible enhancement in the fill rate performance metric. This effect is particularly pronounced within experiments characterized by a low overlap between XYZ/ABC factors in comparison to those with a high overlap. These findings underscore that in situations where the XYZ/ABC overlap is limited, assigning heightened

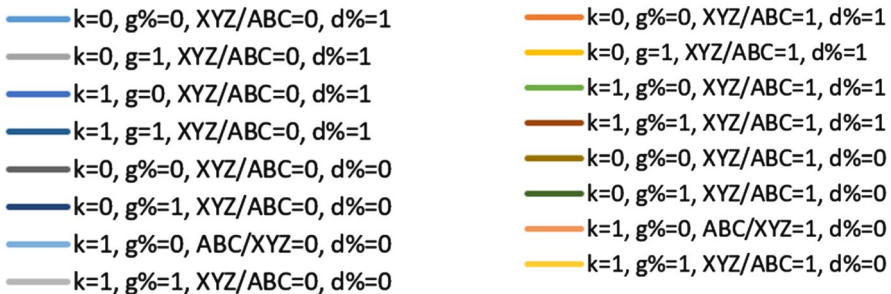
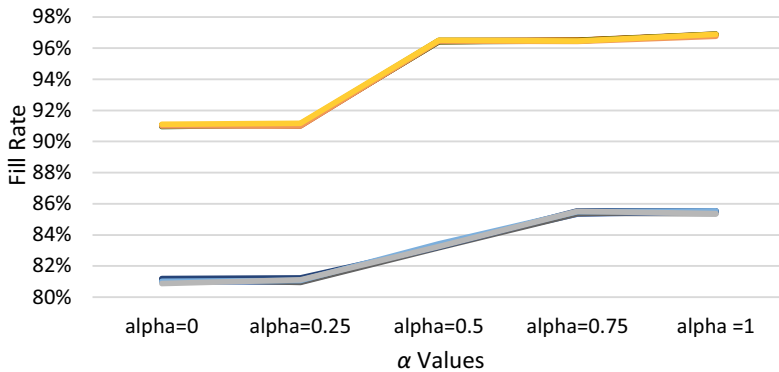
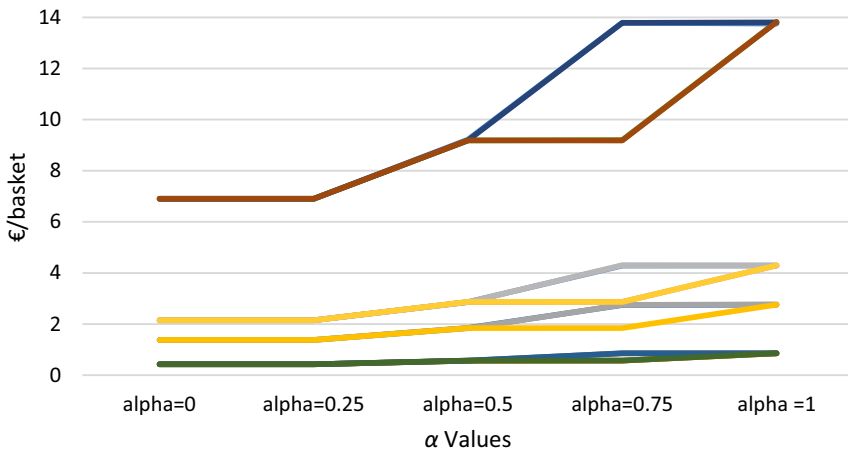


Fig. 9 Description of each experiment of Policy S



**Fig. 10**  $\alpha$  versus fill rate for S policy



**Fig. 11**  $\alpha$  versus delivery cost per basket for S policy

significance to the product availability status within depots becomes pivotal for achieving superior fill rates.

The graphical representation in Fig. 11 sheds light on the relationship between varying  $\alpha$  values and the resultant delivery cost per basket. A prevailing trend is apparent, wherein an augmentation of the  $\alpha$  value is linked to an escalation in the cost per basket. This observation can be rationalized by the fact that as the  $\alpha$  value increases, the significance placed on the distance between the fulfillment center and the customer's location diminishes. As a result, this shift in emphasis triggers a surge in delivery costs, given that other factors, such as product availability, gain prominence in the decision-making process governing allocation.

Figure 12 showcases the relationship between  $\alpha$  and emissions per basket. Notably, as  $\alpha$  increases, there is a corresponding increase in the emissions per basket. This outcome can be attributed to the reduced importance of the distance between

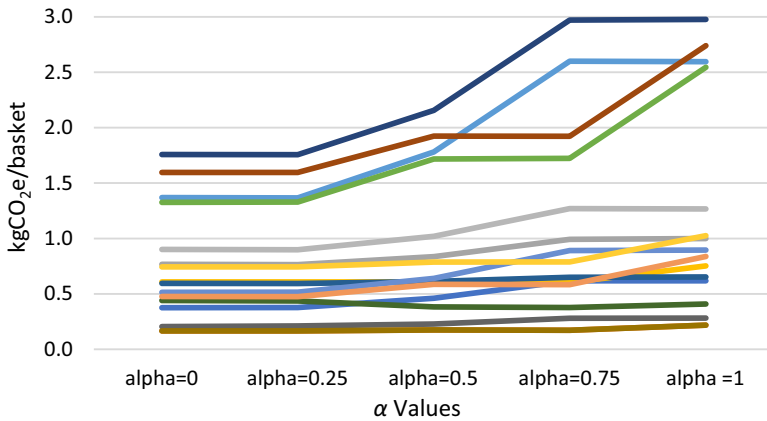


Fig. 12  $\alpha$  versus emissions per basket for the policy S

the fulfillment center and the customer point when  $\alpha$  is higher. Consequently, this leads to higher delivery costs and, subsequently, an elevated emission rate.

Considering Figs. 10, 11 and 12 collectively, it can be deduced that, in terms of a multi-objective approach, selecting  $\alpha=0.75$  may offer the most advantageous outcome. This choice substantially improves the fill rate performance while only moderately increasing the cost and carbon emissions.

The same analysis and conclusions can be applied to the 2S policy, and the corresponding results are presented in the subsequent section.

### 5.2.2 Policy 2S under different $\alpha$ scenarios

In this section, we examine the experiments performed with the 2S policy, utilizing five distinct  $\alpha$  values: 0, 0.25, 0.5, 0.75, and 1. The outcomes of these experiments are illustrated in Figs. 13, 14 and 15. It should be noted that the description of the experiments conducted with the 2S policy remains the same as those conducted with the S policy, as depicted in Fig. 9.

Figure 13 depicts the relationship between  $\alpha$  and fill rate in the 2S policy. It can be observed that the fill rate in the 2S policy is less responsive to changes in  $\alpha$  compared to the S policy. This can be attributed to the fact that the 2S policy assigns two depots to each customer point, thereby increasing the availability of products in the e-grocery system. As a result, the performance metrics of delivery cost per basket and carbon emissions per basket have a greater impact. Furthermore, the graph reveals that the fill rate performance is lower in experiments with a low XYZ/ABC overlap compared to those with a high XYZ/ABC overlap. This suggests that when the XYZ/ABC overlap is low, the weight ( $\alpha$ ) assigned to the product availability status in depots becomes more significant in determining the fill rate.

In Fig. 14, the relationship between  $\alpha$  and delivery cost per basket for the 2S policy is presented. It can be observed that the lowest cost per basket is predominantly achieved at  $\alpha$  values of 0.25 and 0.5. As  $\alpha$  increases beyond these values,

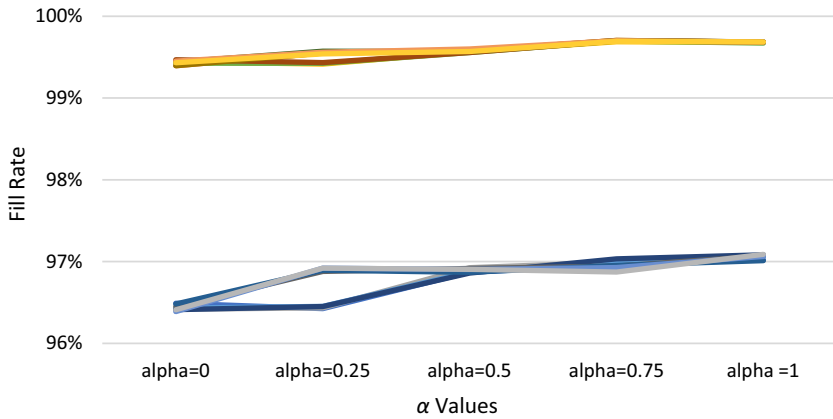


Fig. 13  $\alpha$  versus fill rate for 2S policy

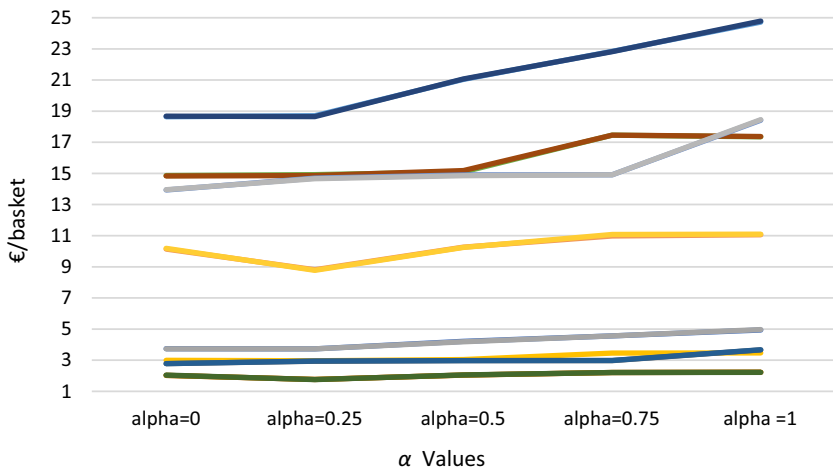


Fig. 14  $\alpha$  versus delivery cost per basket for 2S policy

the delivery cost per basket tends to rise gradually. This indicates that assigning a moderate weight to both product availability and distance metrics ( $\alpha = 0.25$  and  $0.5$ ) can lead to more favorable cost outcomes in the 2S policy.

Figure 15 visually presents the relation between varying  $\alpha$  values and the ensuing carbon emissions per basket, specifically within the context of the “Dual High Combined Score Depot Assignment” (2S) policy. Similar to the observations gleaned from Fig. 14, it becomes evident that the lowest levels of carbon emissions per basket are consistently associated with  $\alpha$  values of  $0.25$  and  $0.5$ . Notably, as the  $\alpha$  value exceeds these thresholds, a gradual increment in carbon emissions per basket becomes apparent.

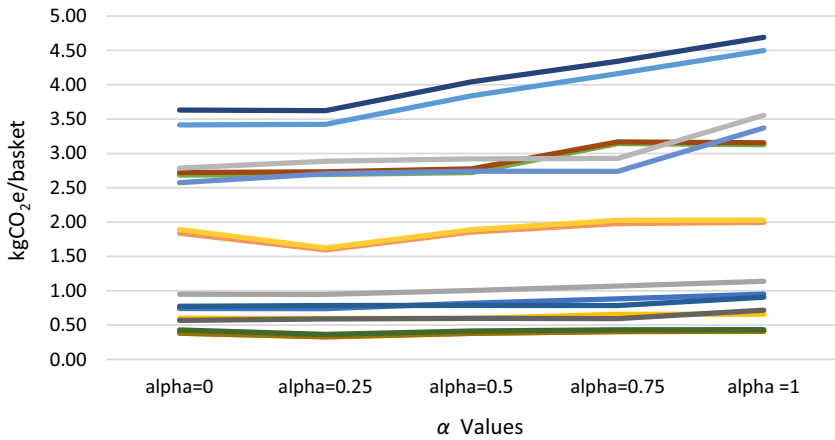


Fig. 15  $\alpha$  versus emissions per basket for 2S policy

This empirical trend accentuates that assigning a moderate weight to both product availability and distance metrics ( $\alpha=0.25$  and  $0.5$ ) can yield more favorable outcomes in terms of mitigating carbon emissions within the framework of the 2S policy.

After a thorough analysis of the experimental results, we further investigate the significant factors influencing each performance metric using ANOVA. The findings from this analysis will be discussed in the subsequent section.

### 5.3 Discussion

- *Proposition 1. The factors XYZ/ABC and the assignment policy affect the fill rate performance significantly.*
- The results from an ANOVA test suggest that the fill rate performance is heavily affected by the two factors: XYZ/ABC and the number of depot assignment policies for customers. Remember that customer preferred products are classified into XYZ classes based on a Pareto analysis of their ordering frequency. X, Y, and Z represent highly, moderately, and lowly ordered classes, respectively. E-grocery products are classified into ABC classes based on their inventory management. A, B, and C represent highly, moderately, and lowly selling classes, respectively. When those are at high levels, the fill rate increases significantly. In other words, when assessing the performance of the fill rate, these two elements should be given special attention.
- *Proposition 2. Delivery cost and emission per basket are affected by those four factors: XYZ/ABC, the assignment policy,  $k$ ,  $d\%$ , significantly.*
- The ANOVA results demonstrate that four factors have a significant effect on the cost and emission performance metric. Of these, the highest impact is caused by

the  $k$  factor, as an increase in the dispersion of households at a high level of  $k$  leads to a higher cost per basket. The second most significant factor is  $d\%$ , as a high level of  $d\%$  increases both the cost and emission outputs.

- *Proposition 3. The best policy for an e-grocery company looking to minimize environmental and economic impact without considering fill rate is policy C, as it usually results in decreased travel distance.*
- If the e-grocery company should consider a cost-efficient and sustainable strategy in its business, in this case, Policy C is typically the best choice since it tends to result in a shorter travel distance. However, in certain situations, the other policies may be more effective in terms of emissions per basket and fill rate performance. For example, if the distance factors are low ( $k=0$  and  $d\%=0$ ) and the probability of complementary shopping is high ( $g\%=1$ ) for a customer with  $XYZ/ABC=1$ , Policies A, S, 2C, 2A, 2S will not only result in an increased fill rate but a decrease in carbon emissions compared to the base policy C. This is likely because the improvement in product availability cuts down on the need for complementary shopping. However, delivery cost per basket is still the best in the C policy.
- *Proposition 4. By using customer historical purchasing data to create a pre-purchase depot allocation policy, it is possible to significantly improve the fill rate for the customer.*
- The results demonstrate that, on average, policy A increases the fill rate by 5.93% in comparison to policy C, which is based on customers' buying habits. Consequently, more sophisticated algorithms that use customer-based past information could potentially lead to even better product availability, resulting in higher customer satisfaction.
- *Proposition 5. Using customer-based historical data, a single-depot allocation policy is still a viable option for an e-grocery due to its simplicity.*
- Comparing C and S policies in Fig. 7, their respective delivery cost per basket performance metrics are similar, however, S policy's fill rate metric is much higher than that of policy C. This suggests that more score-based solutions should be explored to improve multi-objective delivery cost and fill rate-related issues.
- *Proposition 6. If preventing the environmental and economic effects of home delivery is a priority, and increasing the fill rate is desirable, then a data-driven policy which takes both availability and distance into account could be a good option. The focus on availability versus distance can be changed based on the e-grocery*
- *'s desired goals.* Policy S merges policy C and policy A by taking into account both distance and availability. The  $\alpha$  value can be adjusted depending on the goals of the e-grocery. For example, when the overlap between XYZ and ABC is high, policy S with  $\alpha=0.5$  works better than when the overlap is low.

## 6 Conclusion

This paper looks into e-grocery problems, which are mainly defined as when customers cannot purchase the products they intended to buy (pre-purchase out-of-stock) or when the desired products are not available for delivery (post-purchase out-of-stock). We analyze e-grocery allocation policies as a means of addressing these issues, while also considering the improvement of multiple objectives, such as maximizing product availability (i.e. fill rate) and making deliveries more sustainable and cost-efficient. We focus on two research questions: How e-grocery fulfillment policies affect this industry's performance from sustainable delivery perspectives (i.e., product availability, cost-efficiency and carbon emission)? And How customer and e-grocery relevant data can be utilized effectively for managing e-grocery fulfillment? The first research question involves the development of e-grocery fulfillment policies that allocate stores, depots, or fulfillment centers for processing online orders. The allocation policies are designed to allocate stores, depots, or fulfillment centers for processing online orders, addressing pre-purchase out-of-stock and post-purchase out-of-stock scenarios. The seven distinct policies, including Proximity-Based Single Depot Assignment (Policy C), Maximum Availability Depot Assignment (Policy A), Dual High Combined Score Depot Assignment (Policy 2S), and others, were formulated to provide a range of approaches that balance product availability, distance metrics, and customer preferences.

In RQ2, the effective utilization of customer and e-grocery-based data is explored to improve developed fulfillment policies, ensuring they are optimized for performance. This involves leveraging historical purchase data, product popularity trends, and delivery patterns to inform allocation strategies. We conduct a sensitivity analysis to ascertain how performance metrics are influenced by input data. We utilize an experimental design and apply ANOVA to identify factors significantly impacting performance metrics. Through the application of ANOVA, the intention is to mitigate potential biases stemming from the utilization of hypothetical data.

### 6.1 Research implications

This paper contributes to the existing academic literature by looking at product availability and home delivery issues in e-commerce, both of which have rarely been addressed together. We focus on the pre-purchase allocation of fulfillment centers to customers, which has not been thoroughly studied in the literature. This is especially pertinent due to the increasing environmental pressure on last-mile logistics and heightened customer satisfaction expectations in the e-grocery sector. Our research provides a significant contribution to the depot allocation problem in complex e-grocery management by considering customer-based solutions utilizing current data.

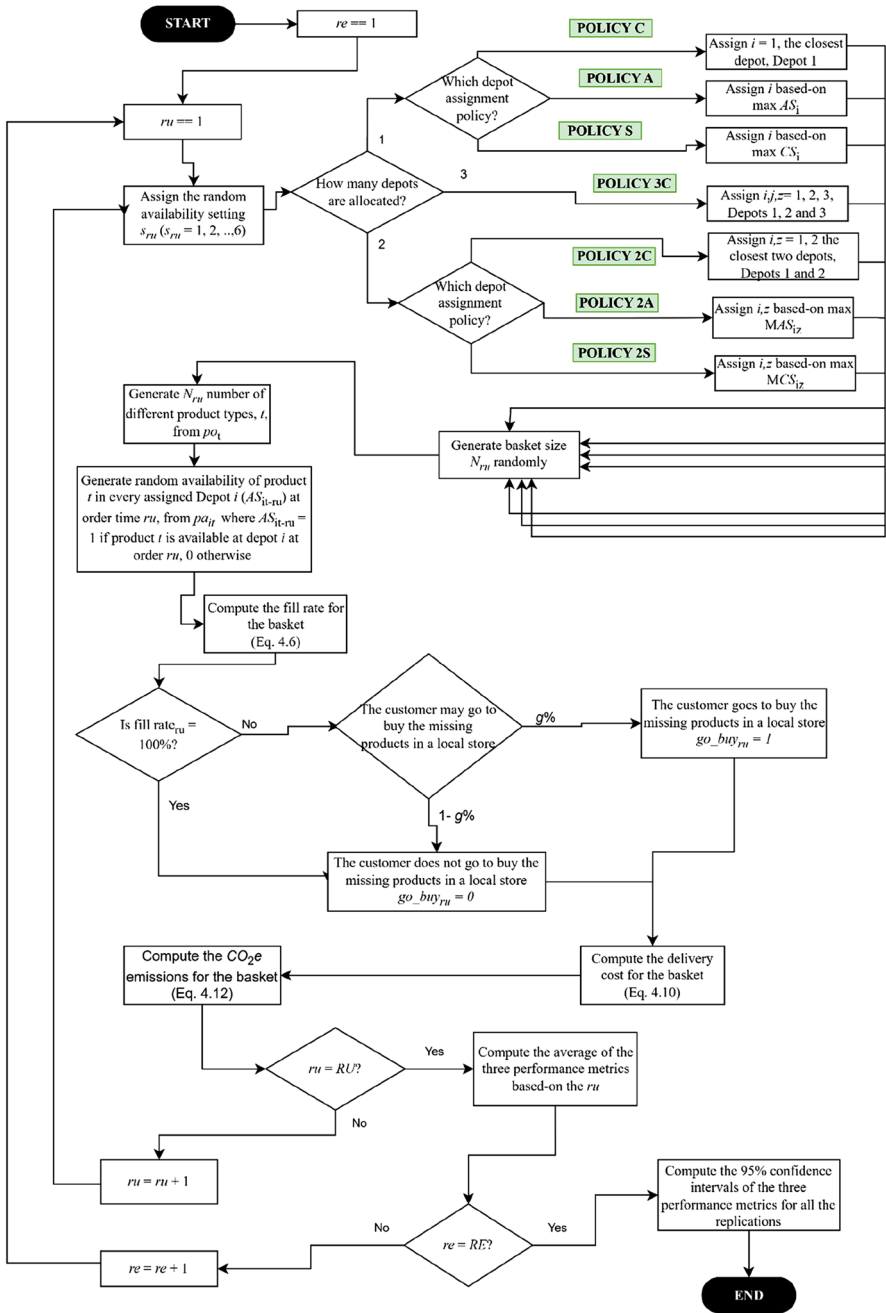
### 6.1.1 Managerial implications

This research explores how e-grocery companies can use the generated data and depot allocation to improve their competitiveness and cost efficiency. By taking advantage of customer-based historical data structure, e-groceries can increase product availability for online shoppers. Seven different depot assignment policies are studied to evaluate how e-grocery companies can use proactive allocation policies. Additionally, the research also considered two performance metrics beyond delivery cost: fill rate and emissions per basket, which can be adapted according to the e-grocery's needs. From these results, e-groceries may use the existing policies or create new ones to best suit their needs.

### 6.2 Limitations and future developments

This simulation study has some limitations, much like any others. For example, it was not possible to validate the simulation models directly on real data since there was no such real case and data. Future works could be done to apply these policies to real companies to compare the results from the computer-generated evidence. Additionally, our focus has primarily been on e-shoppers, neglecting consideration of physical store buyers in cases where a store operates in both realms. This aspect warrants attention for a more comprehensive understanding of the dynamics involved. The models fail to account for fixed costs when implementing various e-grocery methods in the distribution process. Furthermore, the depot allocation policies are only based on customer's purchasing preferences and depots' stock-out parameters, and could be improved by incorporating other advanced methodologies such as big data analytics and machine learning, artificial intelligence, etc. Furthermore, the study does not account for customer's purchasing probability for their next order, which could be addressed by integrating a Markov Process modelling approach. Lastly, the possibility of splitting deliveries from different depots to the customer point instead of transshipment between depots is something that could be further explored.

### Appendix A: Flow chart



**Data availability** Data available within the article.

## Declarations

**Conflict of interest** No potential competing interest was reported by the authors.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Acimovic, J., & Graves, S. C. (2015). Making better fulfillment decisions on the fly in an online retail environment. *Manufacturing & Service Operations Management*, 17(1), 34–51.
2. Ardjmand, E., Bajgirani, O. S., Rahman, S., Weckman, G. R., & Young, W. A., II. (2018). A multi-objective model for order cartonization and fulfillment center assignment in the e-tail/retail industry. *Transportation Research Part E: Logistics and Transportation Review*, 115, 16–34.
3. Azyuz, M. H., Muter, I., Erdogan, G., & Laporte, G. (2022). Minimum cost delivery of multi-item orders in e-commerce logistics. *Computers & Operations Research*, 138, 105613.
4. Bhargava, H. K., Sun, D., & Xu, S. H. (2006). Stockout compensation: Joint inventory and price optimization in electronic retailing. *Inform Journal on computing*, 18(2), 255–266.
5. Breugelmans, E., Campo, K., & Gijsbrechts, E. (2006). Opportunities for active stock out management in online stores: The impact of the stock-out policy on online stock out reactions. *Journal of Retailing*, 82(3), 215–228.
6. Das, J. N., Tiwari, M. K., Sinha, A. K., & Khanzode, V. (2023). Integrated warehouse assignment and carton configuration optimization using deep clustering-based evolutionary algorithm. *Expert Systems with Applications*, 212, 118680.
7. DEFRA. (2022). Greenhouse gas reporting: conversion factors. retrieved on 8 Aug 2023: <https://www.gov.uk/government/publications/greenhouse-gas-reporting-conversionfactors-2022>
8. DHL. (2023). Strategies for Managing Customer Expectations, Small Business Advice: Growing Your SME, retrieved on 20 Jul 2023: <https://www.dhl.com/discover/en-global/small-business-advice/growing-your-business/managing-customer-expectations>
9. Ermecke, K., Olbrich, R., & Brüggemann, P. (2023). They don't do what they say - The attitude-behavior gap in online and offline grocery shopping for organic products. In F. J. Martinez-Lopez (Ed.), *Advances in digital marketing and ecommerce - fourth international conference* (pp. 70–77). Cham: Springer International Publishing, Springer.
10. Forbes. (2023). Online Shopping and Its Advantages, retrieved on 8 Aug 2023: <https://www.forbes.com/advisor/in/credit-card/advantages-of-online-shopping/>
11. Govindarajan, A., Sinha, A., & Uichanco, J. (2021). Joint inventory and fulfillment decisions for omnichannel retail networks. *Naval Research Logistics (NRL)*, 68(6), 779–794.
12. Guidotti, R., Rossetti, G., Pappalardo, L., Giannotti, F., & Pedreschi, D. (2017). Next basket prediction using recurring sequential patterns. arXiv preprint [arXiv:1702.07158](https://arxiv.org/abs/1702.07158)
13. Hoad, K., Robinson, S., & Davies, R. (2010). Automating warm-up length estimation. *Journal of the Operational Research Society*, 61(9), 1389–1403.
14. Hoang, D., & Breugelmans, E. (2020). Sorry, your order has a substitutions: The effects of substitution policy in online grocery. In *International Conference on Advances in National Brand and Private Label Marketing* (pp. 145–155).

15. Hoang, D., & Breugelmans, E. (2023). "Sorry, the product you ordered is out of stock": Effects of substitution policy in online grocery retailing. *Journal of Retailing*, 99(1), 26–45. <https://doi.org/10.1016/j.jretai.2022.06.006>
16. Jasin, S., & Sinha, A. (2015). An lp-based correlated rounding scheme for multi-item e-commerce order fulfillment. *Operations Research*, 63(6), 1336–1351.
17. Jinzhong, S., & Jian, L. (2007). Inventory policy based on differential pricing with stockout compensation in b2c e-commerce. *IEEE International Conference on Grey Systems and Intelligent Services*, 2007, 1023–1027.
18. Jiu, S. (2022). Robust omnichannel retail operations with the implementation of ship-from-store. *Transportation Research Part E: Logistics and Transportation Review*, 157, 102550.
19. Kim, M., & Lennon, S. J. (2011). Consumer response to online apparel stockouts. *Psychology & Marketing*, 28(2), 115–144.
20. Kumar, M. E., Sharma, D. P., & Tapar, A. V. (2021). Out-of-stock justifications and consumers' behavioral outcomes-exploring the role of product type and sales level information in out-of-stock situations. *Journal of Retailing and Consumer Services*, 60, 102458.
21. Lamballais Tessensohn, T., Merschformann, M., Roy, D., de Koster, R. B., Azadeh, K., & Suhl, L. M. (2022). Dynamic policies for resource reallocation in a robotic mobile fulfillment system with time-varying demand. *European Journal of Operational Research*, 300(3), 937–952.
22. Lezcano, M. S., Harper, C. D., Nock, D., Lowry, G. V., & Michalek, J. J. (2023). Online grocery delivery: Sustainable practice, or congestion generator and environmental burden? *Transportation Research Part D: Transport and Environment*, 119(103722), 1361–9209. <https://doi.org/10.1016/j.trd.2023.103722>
23. Li, X., Li, J., Aneja, Y., Guo, Z., & Tian, P. (2019). Integrated order allocation and order routing problem for e-order fulfillment. *IIE Transactions*, 51(10), 1128–1150.
24. Liu, S. (2018). E-shop transshipment selection evaluation based on cloud model. *Journal of Shanghai Jiatong University (Science)*, 23(5), 643–649.
25. Mahar, S., & Wright, P. D. (2009). The value of postponing online fulfillment decisions in multi-channel retail/e-tail organizations. *Computers & operations research*, 36(11), 3061–3072.
26. Mahar, S., Bretthauer, K. M., & Venkataramanan, M. (2009). The value of virtual pooling in dual sales channel supply chains. *European Journal of Operational Research*, 192(2), 561–575.
27. Mahar, S., Modi, S., & Salzarulo, P. A. (2022). Appreciating how your bread is buttered: Improving online order allocation for cross-channel retailers. *International Journal of Production Research*, 60(15), 4845–4867.
28. Marques, P. A., Jorge, D., & Reis, J. (2022). Using lean to improve operational performance in a retail store and e-commerce service: A portuguese case study. *Sustainability*, 14(10), 5913.
29. Martin, J., Nencyz-Thiel, M., Dawes, J., Tanusondjaja, A., Cohen, J., McColl, B., & Trinh, G. (2020). Fundamental basket size patterns and their relation to retailer performance. *Journal of Retailing and Consumer Services*, 54, 102032.
30. McKinsey. (2021a). Making online grocery a winning proposition, retrieved on August, 8th, 2023: <https://www.mckinsey.com/industries/consumer-packaged-goods/our-insights/making-online-grocery-a-winning-proposition>
31. McKinsey. (2021b). Creating a competitive edge in omnichannel grocery fulfilment, retrieved on 8 Aug 2023: <https://www.mckinsey.com/industries/retail/our-insights/creating-a-competitive-edge-in-omnichannel-grocery-fulfillment>
32. McKinsey. (2022). Achieving profitable online grocery order fulfilment, retrieved on 8 Aug 2023: <https://www.mckinsey.com/industries/retail/our-insights/achieving-profitable-online-grocery-order-fulfillment>
33. Milella, A., Petitti, A., Marani, R., Cicirelli, G., & D'Orazio, T. (2020). Towards intelligent retail: Automated on-shelf availability estimation using a depth camera. *IEEE Access*, 8, 19353–19363.
34. Mkansi, M., Eresia-Eke, C., & Emmanuel-Ebikake, O. (2018). E-grocery challenges and remedies: Global market leaders perspective. *Cogent Business & Management*, 5, 1. <https://doi.org/10.1080/23311975.2018.1459338>
35. Mkansi, M., & Nsakanda, A. L. (2021). Leveraging the physical network of stores in e-grocery order fulfilment for sustainable competitive advantage. *Research in Transportation Economics*, 87, 100786. <https://doi.org/10.1016/j.retrec.2019.100786>

36. Nott, G., 2021. Is online or in-store food shopping better for the planet?, retrieved on 8 Aug 2023: <https://www.thegrocer.co.uk/sustainability-and-environment/is-online-or-in-store-food-shopping-better-for-the-planet/656712.article>
37. Peinkofer, S. T., Esper, T. L., & Howlett, E. (2016). Hurry! sale ends soon: The impact of limited inventory availability disclosure on consumer responses to online stockouts. *Journal of Business Logistics*, 37(3), 231–246.
38. Pizzi, G., & Scarpì, D. (2013). When out-of-stock products do backfire: Managing disclosure time and justification wording. *Journal of Retailing*, 89(3), 352–359.
39. PwC Survey. (2022). June 2022 Global Consumer Insights Pulse Survey, retrieved on 20 Jul 2023: <https://www.pwc.com/gx/en/news-room/press-releases/2022/global-consumer-insights-pulse-survey.html>
40. PwC Survey. (2023). June 2023 Global Consumer Insights Pulse Survey, retrieved on 20 Jul 2023: <https://www.pwc.com/gx/en/industries/consumer-markets/consumer-insights-survey.html>
41. Razmi, J., Hassani, A., & Babazadeh, R. (2015). A hybrid fuzzy bi-objective delivery planning model for legible e-tailing under uncertainty. *International Journal of Management Science and Engineering Management*, 10(1), 62–72.
42. Rivière, L., Artigues, C., Cheref, A., Jozefowicz, N., Huguet, M.-J., Ngueveu, S. U., & Charvillat, V. (2021). Multi-product, multi-supplier order assignment and routing for an e-commerce application in the retail sector. In *10th International Conference on Operations Research and Enterprise Systems* (pp. 438–445).
43. Robinson, S. (1997). Simulation model verification and validation: increasing the users' confidence. In *Proceedings of the 29th conference on Winter simulation* (pp. 53–59).
44. Rodríguez García, M., Gonzalez Romero, I., Ortiz Bas, A., & Prado-Prado, J. C. (2022). E-grocery retailing: From value proposition to logistics strategy. *International Journal of Logistics Research and Applications*, 25(10), 1381–1400.
45. Sandıkçı, B., & Sabuncuoğlu, İ. (2006). Analysis of the behavior of the transient period in non-terminating simulations. *European Journal of Operational Research*, 173(1), 252–267.
46. Scott, C. H., & Scott, J. E. (2006). Efficient allocation of online grocery order. *International Journal of Productivity and Quality Management*, 1(1–2), 88–102.
47. Seidel, S. (2021). One goal, one approach? A comparative analysis of online grocery strategies in France and Germany. *Case Studies on Transport Policy*, 9(4), 1922–1932.
48. Siawsohit, C., & Gaukler, G. M. (2021). Offsetting omnichannel grocery fulfilment cost through advance ordering of perishables. *International Journal of Production Economics*, 239, 108192.
49. Siragusa, C., Tumino, A., Mangiaracina, R., & Perego, A. (2022). Electric vehicles performing last-mile delivery in B2C e-commerce: An economic and environmental assessment. *International Journal of Sustainable Transportation*, 16(1), 22–33.
50. Torabi, S. A., Hassini, E., & Jaihoonian, M. (2015). Fulfillment source allocation, inventory transshipment, and customer order transfer in e-tailing. *Transportation Research Part E: Logistics and Transportation Review*, 79, 128–144.
51. Ulrich, M., Jahnke, H., Langrock, R., Pesch, R., & Senge, R. (2021). Distributional regression for demand forecasting in e-grocery. *European Journal of Operational Research*, 294(3), 831–842.
52. Vazquez-Noguerol, M., Comesana-Benavides, J., Poler, R., & Prado-Prado, J. C. (2020). An optimisation approach for the e-grocery order picking and delivery problem. *Central European Journal of Operations Research*, 30(3), 961–990.
53. Viu-Roig, M., & Alvarez-Palau, E. J. (2020). The impact of e-commerce-related last-mile logistics on cities: A systematic literature review. *Sustainability*, 12(16), 6492. <https://doi.org/10.3390/su12166492>
54. Wang, S., & Yang, Y. (2021). M-gan-xgboost model for sales prediction and precision marketing strategy of each product in online stores. *Data Technologies and Applications*, 55(5), 749–770.
55. WEF. (2021). The pandemic has changed consumer behaviour forever - and online shopping looks set to stay, retrieved on 8th Aug 2023: <https://www.weforum.org/agenda/2021/07/global-consumer-behaviour-trends-online-shopping/>
56. Xu, P. J., Allgor, R., & Graves, S. C. (2009). Benefits of reevaluating real-time order fulfillment decisions. *Manufacturing & Service Operations Management*, 11(2), 340–355.
57. Yang, S., Liao, Y., Shi, C., & Gao, C. (2014). Heuristics for solving an internet retailer's dynamic transshipment problem. *Expert Systems with Applications*, 41(11), 5382–5389.

58. Yuan, M., Pavlidis, Y., Jain, M., & Caster, K. (2016). Walmart online grocery personalization: Behavioral insights and basket recommendations. In S. Link & J. Trujillo (Eds.), *Advances in conceptual modeling*. Cham: Springer.

# Enhancing e-grocery order fulfillment: improving product availability, cost, and emissions in last-mile delivery

Ekren, Banu Y.

2025-06

Attribution 4.0 International

---

Ekren BY, Perotti S, Foresti L, Prativiera L. (2025) Enhancing e-grocery order fulfillment: improving product availability, cost, and emissions in last-mile delivery. *Electronic Commerce Research*, Volume 25, Issue 3, June 2025, pp. 2417-2455

<https://doi.org/10.1007/s10660-023-09799-x>

*Downloaded from CERES Research Repository, Cranfield University*