35th European Modeling & Simulation Symposium
20th International Multidisciplinary Modeling & Simulation Multiconference

2724-0029 © 2023 The Authors.
doi: 10.46354/i3m.2023.emss.030

Integrating Simulation and Optimization: A Case Study in Pamplona for Self-Collection Delivery Points Network Design

Irene Izco[1,*], Adrian Serrano-Hernandez[1], Bartosz Sawik[1,2] and Javier Faulin[1]

[1]GILT-OR Group, Institute of Smart Cities, Dept. Statitistics, Computer Science, and Mathematics, Public University of Navarra, Campus Arrosadia s/n, Pamplona, 31006, SPAIN
[2]Department of Business Informatics and Engineering Management, AGH University of Science and Technology, Mickiewicza 30, Kraków, 30059, POLAND

[*]Corresponding author. Email address: izco.121951@e.unavarra.es

Abstract

The disruptions experienced by the processes in the last mile delivery during the SARS-CoV-2 pandemic raised the dilemma of up-to-date last mile approaches for Urban Logistics (UL) issues. Self-Collection Delivery Systems (SCDS) have been proved to be an improvement for all the players of the SC, providing flexibility of time-windows and reducing overall mileage, delivery time and, consequently, gas emissions. Differing from previous works involving hybrid modeling for automated parcel lockers (APL) network design, this paper brings a System Dynamics Simulation Model (SDSM) to forecast online shopping demand in the Spanish city of Pamplona. A bi-criteria Facility Location Problem (FLP) is solved by means of an ε-constraint method, where ε is defined as the level of coverage of the total demand. The experiment run considers 90% of demand coverage, in order to obtain the most complex network possible. The simulation and demand forecast was carried out using Anylogic simulation software and the optimization procedure makes use of the Java-based CPLEX API solver.

Keywords: Last-mile Delivery; Automatic Parcel Lockers; System Dynamics Simulation; Facility Location Problem

1. Introduction

As a consequence of the regulations applied to overcome the pandemic in 2020, supply chains (SCs) experienced disruptions in their performance. Although e-commerce demand skyrocketed, the increase in online shares seems to be a transitory phenomenon. Anyhow, the pandemic has forced urban logistics to opt for alternative last-mile approaches. Hence, this situation has allowed the implementation of some well-known procedures (Faulin et al., 2019; Sawik et al., 2022a) into the urban logistics field. Self-Collection Delivery Systems (SCDS) have been on the radar of researchers and delivery companies, as an attractive alternative to traditional home delivery in last mile distribution. This approach provides flexibility to both couriers and customers, and reduces out-of-vehicle delivery times and vehicle dwell times as well as eliminates the no-shows issues. Therefore, the main advantage of SCDS is that road congestion is mitigated, leading to a reduction in gas emission. This paper focuses on automated parcel lockers (APLs) as particular SCDS, which are multi-compartment storage systems that allow couriers to dispose parcels. The main benefit of APLs is that they permit temporary and secure storage of online purchases until their pick-up by customers at a convenient time. Residential buildings, stores or petrol stations are potential sites for installing automated lockers, where high concentration of population and high e-shopping frequency live together.

Therefore, the main contribution lies in exploring the potential of integrating optimization and simulation techniques to address complex real-world problems, specifically in the context of creating an APL network for last-mile distribution. It demonstrates the advantages of incorporating system dynamics into the simulation of real-life features, particularly when dealing with interrelated demands from diverse customers. This investigation highlights the effectiveness of a hybrid model in effectively tackling intricate optimization problems within the realm of urban logistics.

2. State of the art

Simulation techniques have gained significant prominence in the field of urban logistics as effective tools for modeling and analyzing complex systems. On the one hand, Agent-based Modeling (ABM) and System Dynamics (SD) are widely employed simulation methods suitable for studying real-life systems. ABM, as highlighted in Mehdizadeh et al. (2022), offers a framework for simulating intricate decision systems, considering a diverse population. Unlike econometric approaches, ABM takes into account the dynamic nature of evolving features, such as mobility and behavioral changes in demand. Thus, ABM proves to be a flexible approach capable of capturing the inherent complexity of urban logistics, as demonstrated in this paper.

On the other hand, System Dynamics (SD) is proposed as a methodology for simulating dynamic models to address long-term policy issues both in public and industrial problems, as stated in Forrester (1968). Furthermore, Thaller et al. (2017) showcased a specific application of SD in urban logistics operations, while De La Torre et al. (2019) developed an SD model to examine customer behavior within the context of last-mile delivery. Building upon this research, Rabe et al. (2021) proposed a simulation-optimization approach that integrates a system dynamics simulation model with a multi-period capacitated facility location problem. In this context, the SDSM is designed with the aim to understand the behavior and the interdependences of the agents that form the APL system.

3. Materials and Methods

In this paper, we introduce a hybrid model that merges agent-based simulation modeling with a bi-criteria Facility Location Problem (FLP) Drezner and Hamacher (2004) to address the design of Automated Parcel Lockers (APL) networks. The objective is to simultaneously minimize the number of APLs required and maximize the coverage of demand. By integrating these two approaches, we aim to provide an effective solution for optimizing APL network design while considering multiple objectives. The model is defined over the set of nodes $i \in \mathcal{I}$ and $j \in \mathcal{J}$ representing the APL potential locations and potential customer demand points, respectively. As a novelty, a System Dynamics Simulation Model (SDSM) is built to estimate future three-year horizon demand of online parcel purchases, based on socio-economic factors in Pamplona. This model is applied to each of the demand points $i \in \mathcal{I}$ of the city to calculate the parcel demand allocated to APLs.

Table 1. Model variables.

Variable	Description
x_{ij}	1 if customer node $j \in \mathcal{J}$ is assigned to APL located at node $i \in \mathcal{I}$, 0 otherwise
y_i	Number of APLs located at node $i \in \mathcal{I}$

Once FLP is solved, the resulting APL network further impacts future demand in the SDSM. This impact is captured by the effect of APL proximity, which leads to an exponential increase in both the number of people willing to use APLs and the number of purchases per customer per week. This effect is formulated within the SDSM which describes the relationship between φ_{ijt}, the effect of APL proximity, and dl_{ij}, the distance from a customer node $j \in \mathcal{J}$ to an available (located) locker $i \in \mathcal{I}$ at any nonzero time t: $\varphi_{ijt} = 1 + e^{-dl_{ij}}, \quad \forall i \in \mathcal{I}, \forall j \in \mathcal{J} : t > 0$

The simulation begins with initial data on population, internet users, eShoppers, APL users, and parcel purchases, which are used to create the initial APL network in the city. These variables start evolving based on expected growth rates through the weekly dynamics of the System Dynamics model. On a monthly basis, the FLP is executed, considering the simulated data at that point, to determine the optimal number and location of APLs, which in turn feeds back into the simulation model. This new APL network subsequently affects the future demand, as the availability of APLs nearby positively influences the number of APL users and online purchases.

The FLP integrated within the simulation framework is a bi-criteria optimization model in which minimum number of lockers is desired (in or case that is the objective function 1) and the maximization of utilization of lockers is pursued , willing to cover the maximum demand of parcels of the system (constraint 2). The model is defined over the same set of nodes $i \in \mathcal{I}$ and $j \in \mathcal{J}$ for APL potential locations and demand points, respectively and it is defined as follows:

$$Min \sum_{i \in \mathcal{I}} y_i \tag{1}$$

subject to

$$\sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} d_j c_{i,j} x_{i,j} \geq \epsilon \sum_{j \in \mathcal{J}} d_j \tag{2}$$

$$M \sum_{j \in \mathcal{J}} c_{i,j} x_{i,j} \geq y_i, \quad \forall i \in \mathcal{I} \tag{3}$$

$$c_{i,j} x_{i,j} \leq y_i, \quad \forall i \in \mathcal{I}, \forall j \in \mathcal{J} \tag{4}$$

Table 2. Model parameters.

Parameter	Description
ε	Level of the coverage of total demand
maxDist	Maximum distance a customer is up to travel to pick up its parcel
maxNL	Maximum number of APLs that can be installed at the same location node $i \in \mathcal{I}$
dl_{ij}	Distance from customer node $j \in \mathcal{J}$ to an APL location $i \in \mathcal{I}$
c_{ij}	1 if distance dl_{ij} from customer node $j \in \mathcal{J}$ to an existing APL in location $i \in \mathcal{I}$ is smaller than *maxDist* and 0 otherwise
d_j	Number of parcel demand of customer node $j \in \mathcal{J}$
$\hat{y}_{i,(t-1)}$	Number of previously existing APL installed at location node $i \in \mathcal{I}$
a_i	APL capacity at location $i \in \mathcal{I}$

$$\sum_{j \in \mathcal{J}} d_j c_{i,j} x_{i,j} \leq a_i y_i, \quad \forall i \in \mathcal{I} \tag{5}$$

$$y_i \geq \hat{y}_{i,(t-1)}, \quad \forall i \in \mathcal{I} \tag{6}$$

$$\sum_{i \in \mathcal{I}} c_{ij} x_{ij} \leq 1, \quad \forall j \in \mathcal{J} \tag{7}$$

$$x_{ij} \in \{0, 1\}, \quad \forall i \in \mathcal{I}, \forall j \in \mathcal{J} \tag{8}$$

$$y_i \in [0, maxNL] \subset \mathbb{Z}^+, \forall i \in \mathcal{I} \tag{9}$$

As a result optimal APL location network and assignation of the customers to APLs is obtained. Minimum utilization of APLs for a given demand is accomplished by setting $\epsilon \in (0, 1)$ as the portion of total demand that is desired to be met as defined in constraint (2). Then, constraints (3) and (4) ensure that a locker installed whenever a demand point is assigned to a potential location (M stands for a sufficiently large number - highest estimation). Constraint (5) ensures capacity to be sufficient to meet the demand assigned to a given location. Installed lockers installed remain fixed from period to period, as stated in constraint (6). Constraint (7) ensures that a single locker is assigned to a customer, if assigned. Finally, expressions (8) and (9) define the variable (see Table 1) ranges, defining the maximum number of APLs that can be installed in a potential node. A binary parameter c_{ij} is used to force the assignation of an APL to client that is within a distance *maxDist*, maximum distance a customer is up to travel to pick up its parcel (see model parameters in Table 2).

The system dynamics model is applied to each of the demand points $j \in \mathcal{J}$ of the city in order to calculate the parcel demand allocated to APLs. The interdependences between the stocks and flows of the system are defined as population of each city node is positively affected by the initial population and its growth rate or the internet users depend on the internet users share, its growth rate and internet users themselves. The number of parcels using APLs is positively reinforced by the number of parcels purchased by online shoppers and APL users, which are previously positively affected by the number of e-shoppers, APL user growth rate and the APL users share.

The set of available lockers - first vertex presentation and the set of customer locations - second vertex presentation are connected by axes after the Facility Location Problem is optimized due to parcel demand. Axes represent the flow of customers picking up parcels from lockers. However, we are using Anylogic software for the simulation part, and within Anylogic we have an optimization model of the Facility Location Problem coded in Java. The optimization model is optimized with the use of CPLEX solver - the CPLEX OPL Studio, which is working together with Anylogic. We are doing 152 weeks demand simulations and within this time window we are optimizing the Facility Location Problem every four weeks, so demand is changing, and the network, too. Since we are using Anylogic for simulation, CPLEX solver for optimization, and we have created additional graphical tools to draw a network with location points (vertex) of customers and chosen (after optimization) location (vertex) of APL's (lockers). We know as the input date all potential customer locations, but we do not know which lockers locations are going to be opti-

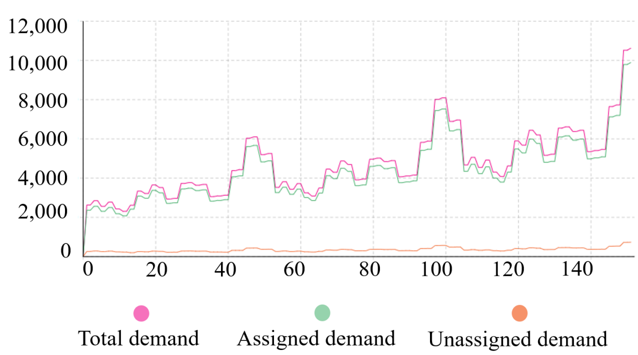

Figure 1. Parcel demand evolution during the simulation in a three-year time horizon for Pamplona APL network for 90% coverage

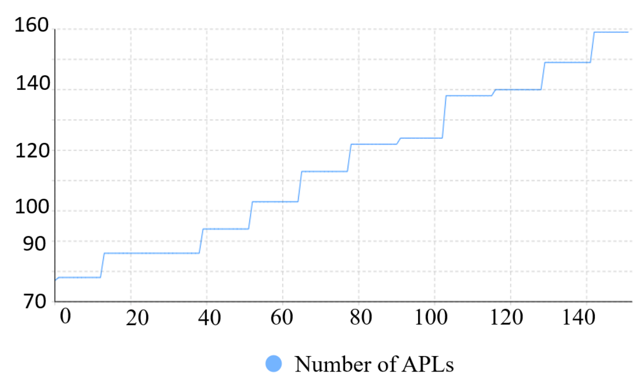

Figure 2. Number of installed APL evolution during the simulation in a three-year time horizon for Pamplona APL network for 90% coverage

mal. Set of all potential lockers locations in the city is also known, but as a solution, we always got part of the set of lockers vertex. To answer the question about what kind of graph represents the two networks, we believe that this is a directed graph, since our graph - or better to say new optimal network (graph) after every four-week optimization round within 152 simulation rounds - is a graph that is made up of a set of vertices connected by directed edges - arcs. Connection is between customer location (vertex) directed to nearest automated parcel locker location (vertex). Directed edge - arc represents number of packages collected weekly by customer from locker. Taking into account below information, we believe that the graph representing our network customer-locker is directed and acyclic.

4. Results and Discussion

The bi-criteria FLP is solved by means of the ε-constraint method, where the factor ε is defined as the level of coverage of total demand (i.e. $\varepsilon \in [0, 1]$). In this study the simulation-optimization model is run for $\varepsilon = 0.90$, and thus, considering that at least 90 % of the total demand must be covered with the obtained APL network. The motivation for this is to show the results of the most populated APL network that can be obtained with the model. Based on a real-world case from Pamplona with data retrieved from related literature (Rabe et al., 2021; Sawik et al., 2022b,a; Serrano-Hernandez et al., 2021) for simplicity reasons, a experiment considering a three years planning horizon has been tested.

As expected from input data, total demand shows an increasing tendency, maintaining the shape every year and having its peak before every Christmas (mid December) ($t = 45, 90, 135$). Coverage of demand fits strictly the utilization constraint (2). As shown in Figure 1 the assigned demand evolves parallel to the total parcel demand being always above its 90%. The coverage reached in the network at the end of the simulation is greater than the one required in the utilization constraint, a 93.1% ($\varepsilon = 0.9$). In this experiment, a total of 10, 627 parcels are purchased in the last week, being the highest demand coinciding with the last Christmas period ($t = 145$) considered in the simulation time.

The number of APLs installed in the network follows the same tendency of the demand (see Figure 2), increasing step-wise according to the number of parcels that must be attended in the system. This matches with the fact that fixed lockers are considered in this study, and thus the potential demand that can be attended is accumulated in the existing APLs. The final number of APLs in the network changes as expected with the desired 90% of demand coverage, with a total of 9,895 parcels assigned to 159 APLs. Figures 3 and 4 show the APL network obtained at the beginning and the end of the experiment, respectively.

The potential customer nodes are depicted as white dots and the installed APLs show up in blue once the FLP problem is solved and an APL is installed. At $t = 0$, after the FLP is solved for the first time, the lockers are dispersed homogeneously throughout the city covering all the districts (see Figure 3). At the end of the simulation, when $t = 151$, as highlighted in yellow in Figure 4 , the network has a new potential location opened and the number of open locker compartments in the previously installed locations increases (depicted in red in Figure 4), matching the increase of parcel demand.

In Figure 3 we presented a satellite view of the installed APLs in the network at the beginning period of simulation t = 0 in Pamplona. In Figure 4 we showed a satellite view of the installed APLs in the network at the end period of simulation t = 151 in Pamplona. We have decided not to add

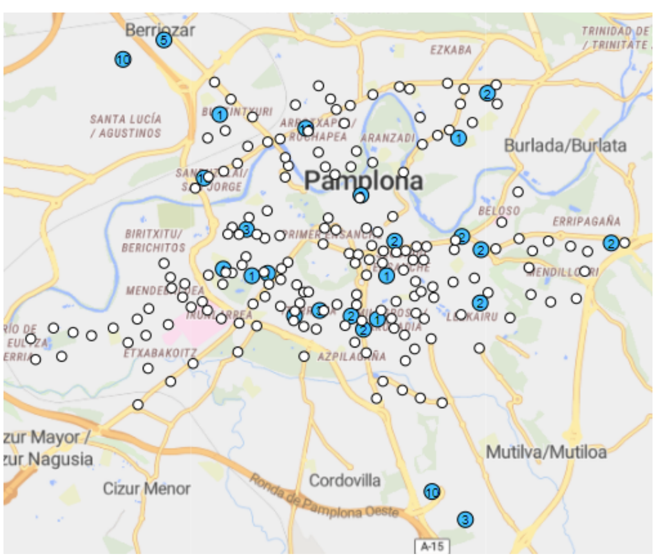

Figure 3. Satellite view of the installed APL in the network at the beginning period of simulation $t = 0$ in Pamplona.

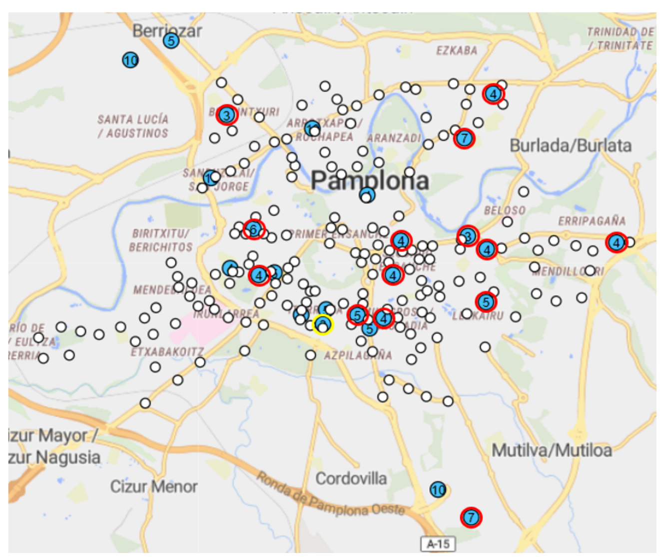

Figure 4. Satellite view of the installed APL in the network at the end period of simulation $t = 151$ in Pamplona.

in both figures directed edges - arcs between all customer locations (vertex) directed to the nearest automated parcel locker location (vertex). In both figures (3 and 4) the graph representing our network customer-locker is directed and acyclic, but adding directed edges - arcs between customer and locker will make both pictures illegible.

Besides, it is important to mention that alternative optimal solutions are obtained from the FLPs, which means that the number of APLs is kept the same, while their location in the network varies for each alternative solution. The potential of the model built for this study is that it has the capability of obtaining graphical results of the network, being able to locate the APLs on their exact coordinates in an interactive map of Pamplona (Figures 3 and 4).

5. Conclusions

This paper proposed the use of an integrated simulation-optimization approach merging system dynamics with exact optimization. The first conclusion that can drawn is the convenience of developing a System Dynamics Model to forecast the demand, as this allows to control this process in a more accurate way and gives more convenient results than other simulation approaches for estimating flows of people with individual behaviours. Secondly, relying on simulation-optimization methodology and considering APL user proximity effect has proven that the results obtained in the optimization model strictly influence the results of the simulation model. Lastly, increasing the level of detail in the demand data, and thus, considering individual customers' demand instead of aggregated demand in districts, has notably improved the quality of the obtained results. For instance, it is clearly observed that the rise of e-shoppers and purchase rate, increases the APL usage, as intuitively the number of lockers and the parcel demand are directly proportional.

Thus, we explore the potential of integrating optimization and simulation techniques to tackle challenging real-world problems. The paper demonstrates the advantages of utilizing system dynamics in simulating real-life features, especially when dealing with interrelated demands from various customers. Consequently, constructing a hybrid model emerges as a suitable approach to effectively handle intricate optimization problems in urban logistics.

However, it is important to note that this study has two notable limitations. Firstly, the available data used in this study may not always be accurate or up-to-date. This is due to reliance on older sources or data that may pertain to different regions, potentially introducing inconsistencies or inaccuracies in the analysis. Secondly, certain significant factors had to be omitted from the system due to the unavailability of relevant data. For instance, variables such as climate effects on demands and traffic conditions were not included due to the lack of data. These limitations highlight the need for future research to address these challenges. It is crucial to move forward by acquiring data from primary sources, which could involve conducting in-field surveys to gather more accurate and reliable data. Additionally, incorporating advanced statistical methods can help enhance the estimation of socioeconomic parameters. By addressing these limitations, future studies can provide a more comprehensive understanding of the behavior of the parcel demand system and improve the accuracy of their findings.

6. Acknowledgements

This work was partially funded by the Spanish Ministry of Science, Innovation, and Universities (RED2022-134703-T; PID2019-111100RB-C22/AEI/10.13039/501100011033) and the SEPIE Erasmus+ Program (2019-I-ES01-KA103-062602). We also appreciate the help of the Public University of Navarra for Young Researchers Projects Program (PJUPNA26-2022). Likewise, we appreciate the help of the AGH University of Science & Technology, Krakow, Poland (16.16.200.396) and the financial aid of the Polish Ministry of Science & Higher Education grants (N N519 405934; 6459/B/T02/2011/40), and the Polish National Science Centre research grant (DEC-2013/11/B/ST8/04458).

References

De La Torre, G., Gruchmann, T., Kamath, V., Melkonyan, A., and Krumme, K. (2019). *A System Dynamics-Based Simulation Model to Analyze Consumers' Behavior Based on Participatory Systems Mapping – A "Last Mile" Perspective*, pages 165–194. Springer, New York, NY.

Drezner, Z. and Hamacher, H. (2004). *Facility Location Applications and Theory*. Springer, New York, NY.

Faulin, J., Grasman, S., Juan, A., and Hirsch, P. (2019). *Sustainable Transportation and Smart Logistics: Decision-Making Models and Solutions*. Elsevier, Cambridge, MA.

Forrester, J. W. (1968). Industrial dynamics — after the first decade. *Management science*, 14(7):398–415.

Mehdizadeh, M., Nordfjaern, T., and Klöckner, C. A. (2022). A systematic review of the agent-based modelling/simulation paradigm in mobility transition. *Technological Forecasting and Social Change*, 184:122011.

Rabe, M., Gonzalez-Feliu, J., Chicaiza-Vaca, J., and Tordecilla, R. D. (2021). Simulation-optimization approach for multi-period facility location problems with forecasted and random demands in a last-mile logistics application. *Algorithms*, 14(2):41.

Sawik, B., Faulin, J., Serrano-Hernandez, A., and Ballano, A. (2022a). A simulation-optimization model for automated parcel lockers network design in urban scenarios in pamplona (spain), zakopane, and krakow (poland). In *2022 Winter Simulation Conference (WSC)*, pages 1648–1659. IEEE.

Sawik, B., Serrano-Hernandez, A., Muro, A., and Faulin, J. (2022b). Multi-criteria simulation-optimization analysis of usage of automated parcel lockers: A practical approach. *Mathematics*, 10(23):4423.

Serrano-Hernandez, A., Martinez-Abad, S., Ballano, A., Faulin, J., Rabe, M., and Chicaiza-Vaca, J. (2021). A hybrid modeling approach for automated parcel lockers as a last-mile delivery scheme: a case study in pamplona (spain). In *2021 Winter Simulation Conference (WSC)*, pages 1–12. IEEE.

Thaller, C., Niemann, F., Dahmen, B., Clausen, U., and Leerkamp, B. (2017). Describing and explaining urban freight transport by system dynamics. *Transportation research procedia*, 25:1075–1094.