A novel districting design approach for on-time last-mile delivery: An application on an express postal company[1]

M. Gabriela Sandoval
Department of Computers, Electronics and Mechatronics
Universidad de las Américas Puebla (UDLAP), San Andrés Cholula, Mexico
E-mail: *maria.sandovalel@udlap.mx*

Eduardo Álvarez-Miranda
School of Economics and Business, Universidad de Talca, Talca, Chile
Instituto Sistemas Complejos de Ingeniería, Chile
E-mail: *ealvarez@utalca.cl*

Jordi Pereira
Faculty of Engineering and Sciences, Universidad Adolfo Ibáñez, Viña del Mar, Chile
E-mail: *jorge.pereira@uai.cl*

Roger Z. Ríos-Mercado
Universidad Autónoma de Nuevo León (UANL)
Graduate Program in Systems Engineering, San Nicolás de los Garza, Mexico
E-mail: *roger.rios@uanl.edu.mx*

Juan A. Díaz
Department of Actuary, Physics and Mathematics
Universidad de las Américas Puebla (UDLAP) San Andrés Cholula, Mexico
E-mail: *juana.diaz@udlap.mx*

April 2021

[1] Preprint PISIS-RR21-02, Graduate Program in Systems Engineering, UANL, San Nicolás de los Garza, Mexico, April 2021

Abstract

Last-mile logistics correspond to the last leg of the supply chain, i.e., the delivery of goods to final customers and they comprise the core activities of postal and courier companies. Due to their role in the supply chain, last-mile operations are critical for customers' perception concerning the performance of the whole logistic process. In this sense, the sustained growth of e-commerce, which the irruption of the COVID-19 pandemic has abruptly catalyzed, has modified customers' habits and overtaxed the operational side of delivery companies. Many of these habits will remain after overcoming the sanitary crisis, which will permanently reshape the structure and emphasis of postal supply chains, demanding companies to implement organizational and operational changes to address these challenges.

In this work, we address a last-mile logistic design problem faced by a courier and delivery company in Chile, although the same problem is likely to arise in the last-mile delivery operation of any postal company, in particular in the operation of express delivery services. The company's operational structure is based on the division of an urban area into smaller territories (districts) and outsourcing the delivery operation of each territory to a last-mile contractor. Due to the increasing volume of postal traffic and decreasing service performance, mainly for express deliveries, the company is forced to redesign its current territorial arrangement. Such a redesign results in a novel optimization problem that resembles a classical districting problem with additional service requirements. This novel problem is first formulated as a mathematical programming model, and then a specially tailored heuristic is designed for solving it. The proposed approach is tested on instances from the real-life case study. The obtained results show significant improvements in terms of the percentage of on-time deliveries achieved by the proposed solution compared to the current districting design of the company. Furthermore, by performing a sensitivity analysis for different levels of demand, we show that the proposed approach is effective in providing districting designs capable of enduring significant increases in the demand for express postal services.

Keywords: Districting; Last-mile delivery; Postal delivery; Supply chain management; Heuristics.

1 Introduction and Motivation

In the management science community, districts correspond to territorial or administrative entities whose composition tries to balance the demand and the supply of a given *service* or *good*. Such services or goods might be associated with decision-making contexts such as electoral systems (Ricca et al., 2013), postal logistics (Jarrah and Bard, 2012), schooling policy (Caro et al., 2004), commercial territory planning (Ríos-Mercado and Fernández, 2009), health care management (Benzarti et al., 2013), emergency response planning (Mayorga et al., 2013), municipal solid waste collection (Lin and Kao, 2008), among others. In all these contexts, districts comprise basic functional units (e.g., customers, streets, zip codes, cities, and so on) of a larger geographical area (e.g., a city or a province). Defining these districts is known as *districting* (or *territorial design*). The goal is to find an optimal solution according to one or more planning criteria (e.g., the balance of the number of voters, equitable assignment of customers, and so on) that verifies other geometric and functional requirements (e.g., connectivity, compactness, or service performance).

Among the different applications, the design of delivery districts, especially in postal and last-mile logistics, is one of the most relevant ones (Kalcsics and Ríos-Mercado, 2019). It is common among logistic operators to divide urban areas into delivery zones (i.e., districts). Households and commercial and institutional customers form these districts. Then, the operators must assign these zones to one or more delivery workforce units (e.g., postmen), who are then responsible for delivering parcels, letters, and other postal products to the corresponding customers. The performance of these operations typically depends on the territorial structure of the districts and the postal traffic workload distribution among them: districts are expected to be *compact* (i.e., undistorted, without holes, and featuring a smooth boundary) and *balanced* (i.e., mail carriers are expected to have a similar workload on equivalent workday lengths). We refer the reader to Jarrah and Bard (2012) and Parriani et al. (2019) for two examples of applying these two criteria when designing delivery districts.

Over the last two decades, the postal service industry has undergone significant global changes. On the one hand, the traffic of traditional letter mails has dramatically declined and, on the other hand, the volume of parcels has significantly increased. According to Dragendorf et al. (2019), the mail-to-parcel volume ratio worldwide changed from 13:1 in 2000 to 2:1 nowadays, and it is expected to be 1:1 by 2025. Moreover, other sources have reported similar evidence (International Post Corporation, 2019; Universal Postal Union, 2019). The rise and consolidation of e-commerce plays a crucial role (see, e.g., Hong and Wolak, 2008; Nikali, 2008; Winkenbach and Janjevic, 2018). These changes have had a tremendous impact from a strategical to an operational level, especially on last-mile delivery operations where the demand for express courier services has substantially increased (see Dragendorf et al. (2019) for a brief report on how current market trends are transformed by last-mile delivery). Furthermore, the COVID-19 pandemic has imposed enormous stress on the whole postal delivery logistics, especially on express services, catalyzing several technological and operational changes that will permanently reshape the last-mile delivery operations (see OECD, 2020; Banker, 2020; Nyrop et al., 2020; Universal Postal Union, 2020, for recent analyses on the issue).

As a result, courier companies face enormous challenges in coordinating their operations and accomplishing their compromised service standards while preserving their economic viability (see Bates et al. (2018); Crew and Brennan (2015); Lee et al. (2016) for overviews on the challenges and opportunities for the next generation of last-mile distribution). Moreover, as these changes have occurred rather abruptly, postal service companies endure difficulties adapting to these challenges, particularly in developing countries (we refer the reader to Laseinde and Mpofu (2017), for a comprehensive study on challenges in postal operations in developing nations). One of these difficulties corresponds to the redesign of the districts associated with the delivery of mail and small/medium parcels. The design of these delivery areas usually results from a long negotiation among mail carriers, unions, delivery contractors, and postal company managers. Once these districts are defined, they are expected to remain unaltered for several years (we refer the reader to Demazière and Mercier, 2006, for a study on the singularities of postmen activities and the engagement with their delivery areas). Therefore, *redrawing* these districts (by adding or removing delivery points), or modifying the number of districts within a city, is not a simple task, even if the current design leads to declining performance metrics caused by changes in volume or structure of the required services. Attempts to carry out such changes are expected to be based on objective, fair, transparent, and consensual criteria, and procedures; furthermore, they shall be embedded in impartial decision-aid tools (see Verhoest and Sys, 2006; Uni Global Union, 2019, where different experiences on the use of decision-aid tools for delivery district modifications are exposed for national and private postal companies).

This work stems from a courier and express delivery company in Chile that is currently dealing with the challenges mentioned above. The company offers a national express delivery service for mail letters and small parcels, ensuring that they reach their destination the next working day before 11 am. The demand for this service has overwhelmingly increased over the last years, broadening its share of its revenues and straining the entire supply chain, especially the last-mile delivery operations. To fully understand the issues and challenges faced by the company, we briefly describe the country-level and their local-level operations. Letters and parcels are first sent from service points to the closest sorting hub located near the major airports. Express deliveries reach the closest sorting hub the same day they were handed to the company, and during the night shift, they are sorted according to their destination. Afterward, they are sent to their destination distribution depots using either ground or air transportation depending on the distance between the sorting center and the destination center. Once the letters and the parcels arrive at their corresponding destination depots, they are sorted according to their priority (i.e., express and non-express) and according to their address (i.e., by identifying their corresponding district). Finally, they are handed over to the personnel performing last-mile delivery that organizes deliveries within their route according to their experience.

The courier company has outsourced the last-mile delivery operations to independent contractors. The corresponding outsourcing contract stipulates a multi-year working relationship that defines the district assigned in exclusivity to each delivery contractor, an agreed-upon minimum payment from the courier to the contractor, and an additional payment per delivered parcel. These districts were settled by rules-of-thumb criteria and using outdated postal traffic records. Hence, as the company does not perform last-mile delivery by itself, it heavily relies on the contractors' ability to organize their workload and compliance to service levels, particularly for express delivery. Moreover, both the courier and the contractors tend to overestimate the contractors' capability for any parcel. Couriers tend to minimize the number of contractors to reduce minimum payments, and contractors tend to overestimate their delivery capacity to maximize their income. As a result, satisfactory on-time delivery of express services has been a significant issue. The company has recently set regional KPIs (key performance indicators) to measure the percentage of letters and parcels delivered on time and has established bonuses based on the performance in these KPIs to middle management and subcontractors. These policies have encouraged local managers and last-mile contractors to improve their delivery operations, opening the opportunity to negotiate and modify current delivery districts, improve contractors' workload, and increase the percentage of on-time deliveries.

This work presents a novel optimization methodology for addressing districting design decisions such as the one faced by the company in consideration. As we will describe, the proposed approach not only allows to design districts according to the company's performance measures, but it also allows to evaluate their expected performance under different scenarios.

Our contribution and paper outline . The main contribution of this paper is the development of a methodological framework for modeling and solving a novel districting design problem that aims at defining last-mile delivery areas according to a quality of service objective rather than to workload balancing considerations. The quality of service, in this case, is measured as the percentage of on-time deliveries. The resulting problem is coined as the on-time last-mile delivery districting problem (OTLMDP). Its two-stage nature characterizes the OTLMDP: mid-to-long term districting planning decisions (first stage) are made along with operational last-mile routing delivery decisions (second stage). In this second stage, we incorporate the variable nature of the express service demand (traffic volume and recipients' territorial deployment) to ensure that last-mile operators can accomplish improved levels of on-time deliveries. To solve the OTLMDP, we designed an ad-hoc heuristic featuring an initialization phase, a construction phase, and an exploration phase. These phases encompass specially designed grouping strategies, routing algorithms, and local-search operators.

The proposed approach is applied to a set of instances derived from the considered case study. In particular, we consider the case of Antofagasta, the fifth largest city of Chile, capital of the mining region of the country, and one of the areas where the problem mentioned above is severe. The attained results show that considering the number of districts currently deployed in Antofagasta, it is possible to design last-mile delivery districts featuring, on average, over 98% of on-time deliveries. Furthermore, considering the current market trends and the impact of the COVID-19 pandemics on the whole supply chain industry, particularly on postal and courier operators, we use the designed framework to carry out a sensitivity analysis regarding a different number of districts and also different demand level scenarios.

The paper is organized as follows. In Section 2, we present a review of related literature, and we highlight the methodological gap addressed by this work. In Section 3, we present a mathematical programming formulation for the OTLMDP. The details of the specially devised heuristic are outlined in Section 4. Further characteristics of the case study and a discussion of the obtained results are presented in Section 5. Finally, conclusions and avenues for future work are drawn in Section 6.

2 Literature review

Research on districting or territory design goes back to the early 1960. The main application areas were political districting (Ricca et al., 2013) and sales territory design and alignment (Ronen, 1983; Zoltners and Sinha, 2005). In the past twenty years, there has been an active research area for districting models, methods, and applications such as school districting (Caro et al., 2004), commercial territory design (Ríos-Mercado and Fernández, 2009; Salazar-Aguilar et al., 2011a,b; Ríos-Mercado and López-Pérez, 2013; López-Pérez and Ríos-Mercado, 2013; Ríos-Mercado and Escalante, 2016; Sandoval et al., 2020; Ríos-Mercado et al., 2021), public service districting (Bender et al., 2018), collection of waste electric and electronic equipment (Fernández et al., 2010; Ríos-Mercado and Bard, 2019; Ríos-Mercado et al., 2020), health care management (Mayorga et al., 2013; Gentry et al., 2015; Steiner et al., 2015; Yanık et al., 2019; Restrepo et al., 2020), market design and segmentation (Mansfield et al., 2003; Huerta-Muñoz et al., 2017), sales force sizing (Moya-García and Salazar-Aguilar, 2020), and finance (de la Poix de Fréminville et al., 2015), to name a few. There are some excellent surveys where the reader may find a detailed discussion of the models and solution algorithms (Duque et al., 2007; Kalcsics and Ríos-Mercado, 2019). There is a recent book due to Ríos-Mercado (2020) covering recent advances on districting and territory design applications, models, and algorithms, including surveys on police districting (Liberatore et al., 2020), health care (Yanık and Bozkaya, 2020), and computational geometry methods (Behroozi and Carlsson, 2020). This section focuses our discussion on districting applications related to postal, delivery operations, and routing, which is the subject of our work.

It is worth mentioning that a particular class of districting problems deals with edge-partition decisions (see, e.g., Muyldermans et al., 2002, 2003; Butsch et al., 2014; García-Ayala et al., 2016). In these types of problems, attributes are associated with the edges, not the nodes. These are still tactical models, in the sense that routing is not decided; however, these are the first attempts to consider arc services for deciding the territory design.

2.1 Districting application with embedded routing factors

One of the first papers to address districting and routing corresponds to Chapleau et al. (1985), who studied a school districting and routing problem. The first phase of their two-stage approach determines appropriate districts, including an adequate number of students to be served. Then, for each district, a route and the stops along this route are determined. They present numerical results with reasonably good results.

Haugland et al. (2007) study the problem of designing districts for vehicle routing problems with stochastic demands. In particular, demands are assumed to be uncertain when the districts are made, and these are revealed after the districting decisions are determined. They propose a two-stage stochastic model minimizing the expected routing cost. In the first stage, demand is considered a stochastic variable, and the goal is to define the districts so that the expected total travel cost is minimized. In order to balance the districts more evenly, they impose a constraint limiting the actual travel cost within each district to a given upper bound. Connectivity constraints are considered. They propose tabu search and multi-start heuristics, observing that tabu search outperforms multi-start in terms of solution quality.

Ríos-Mercado and Salazar-Acosta (2011) study a commercial districting problem in the bottled beverage distribution industry and introduce a combinatorial optimization model. The problem consists of grouping a set of city blocks into territories to maximize territory compactness. As planning requirements, the grouping seeks to balance both the number of customers and product demand across territories, maintain connectivity of territories, and limit the total routing costs. This work addresses both design and routing decisions simultaneously by considering a budget constraint on the total routing cost. The authors propose a greedy randomized adaptive search procedure that incorporates advanced features such as adaptive memory and strategic oscillation. They test their algorithm on a broad set of randomly generated instances based on real-world data showing a positive impact of the advanced components. Solution quality is significantly improved as well.

Lei et al. (2012) address a vehicle routing and districting problem with stochastic customers. Planning requirements include generating contiguous districts, servicing each customer (including regular and stochastic customers) within the same district by the same vehicle, visiting each customer vertex once by one vehicle while minimizing an objective function combining vehicle cost, routing cost, and a district compactness measure. They introduce a two-stage stochastic programming model in which the districting decisions are made in the first stage. The Beardwood–Halton–Hammersley formula is used to approximate the expected routing cost of each district in the second stage. District compactness is also considered as part of the objective function. They developed a large neighbourhood search heuristic. They tested the heuristic on modified Solomon instances and on modified Gehring and Homberger instances. Extensive computational results confirm the effectiveness of the proposed heuristic.

Schneider et al. (2014) develop a two-phase territory-based routing approach motivated by an application in the package shipping industry. They use this approach to investigate the design requirements for successfully handling time windows and study the influence of time window constraints on the performance of such an approach. They found that considering geographical aspects in the districting turns out to be critical for generating high-quality territories. Incorporating time window characteristics and historical demand data does not lead to a perceptible improvement of the solution quality.

Lei et al. (2015) introduce the multiple traveling salesperson and a districting problem with multi-periods and multi-depots. In this problem, the compactness of the districts, the dissimilarity measure of districts, and an equity measure of salesperson profit are part of the objective function, and the salesperson travel cost on each district is approximated by the Beardwood–Halton–Hammersley formula. They developed an adaptive large neighbourhood search metaheuristic for the problem. They tested their algorithm on modified Solomon and Gehring and Homberger instances. Computational results confirm the effectiveness of the proposed metaheuristic.

Lei et al. (2016) study a multi-objective dynamic stochastic districting and routing problem in which the customers of a territory stochastically evolve over several periods of a planning horizon. The number of service vehicles, the compactness of the districts, the dissimilarity measure of the districts, and an equity measure of vehicle profit are used as objectives. The problem is modeled and solved as a two-stage stochastic program. In each period, districting decisions are made in the first stage. The Beardwood–Halton–Hammersley formula is used to approximate the expected routing cost of each district in the second stage. They developed an enhanced multi-objective evolutionary algorithm (MOEA). They tested their algorithm on randomly generated instances and compared it with two state-of-the-art MOEAs. Computational results confirm the superiority and effectiveness of the proposed algorithm.

Konur and Geunes (2019) introduce an integrated districting, fleet composition, and inventory planning problem for a multi-retailer distribution system. In particular, they analyze the districting decisions for a set of retailers such that the retailers within the same district share truck capacity for their shipment requirements. They consider that the number of trucks of each type dedicated to a retailer district and retailer inventory planning decisions is jointly determined in a district formation problem. They present a mixed-integer-nonlinear programming formulation for this problem and develop a column generation-based heuristic approach for its set partitioning formulation. To do so, they first characterize essential properties of the optimal fleet composition and inventory planning decisions for a given retailer district. Then, they use these properties within a branch-and-price method to solve the integrated districting, fleet composition, and inventory planning problem. Their numerical experiments show the efficiency of the proposed approach.

Moreno et al. (2020) study a similar problem found in meat distribution. Candidate districts are generated using a modified K-means heuristic that evaluates the time required to deliver goods within the district according to an approximation formula, following Lei et al. (2015). The districting plan is then obtained by solving an integer programming formulation to select a subset of the candidate districts generated by the modified K-means heuristic.

Bender et al. (2020) propose a two-stage solution approach for the assignment of drivers and vehicles to customers that relies on districting instead of vehicle routing techniques and allows for daily adaptations. For the tactical planning level in the first stage, they developed a districting approach that involves determining the number of districts and the assignment of heterogeneous resources. They present three integer programming (IP) models for the tactical planning problem, which differ in the level of detail of their input data and in their expected compliance with the drivers' contractual working times. They also present a heuristic solution procedure for the tactical problem. The authors propose a mixed-integer programming model that adapts the tactical districting solution to the concrete demand realization of a day for the operational level on the second stage. In their numerical experiments, they analyzed the feasibility of using districting approaches for the problem at hand and the suitability of the three tactical planning models. They investigated the trade-off between compliance with the drivers' contractual working times and service consistency.

Zhou et al. (2021) address a districting and routing application in a dairy firm responsible for producing and distributing perishable products. The problem consists of grouping customers into geographic districts to minimize the total operational cost, computed as a function of the fixed costs of the districts and the routing costs. The authors present and assess a genetic algorithm enhanced with several search techniques. The proposed design is extensively tested on instances derived from the literature and real-world instances involving more than 1000 customers. The results show the effectiveness of their proposed heuristic in producing good-quality operational solutions.

2.2 Postal and Delivery Applications

To the best of our knowledge, the only work involving districting in postal or pick-up and delivery applications are due to Bard and Jarrah (2009) and Jarrah and Bard (2012). Bard and Jarrah (2009) address a clustering problem in regional pick-up and delivery operations by a shipping firm. Their goal is to build a set of compact districts that satisfy volume, time, and contiguity constraints while minimizing the number of districts (homogeneous vehicles). They introduce a mixed-integer linear programming (MILP) model and present a three-phase procedure for clustering the data points to tackle instances in the order of 6,000 to 50,000 basic units. The solution procedure makes use of metaheuristic and mathematical programming techniques. In the preprocessing phase, a fraction of the data points is aggregated to reduce the problem size. The preprocessing substantially decreases the computational burden without compromising solution quality. In the main step, an efficient adaptive search procedure is used to form the clusters. Randomness is introduced at each inner iteration to ensure a full exploration of the feasible region, and an incremental slicing scheme is used to overcome local optimality. In metaheuristic terms, these two refinements are equivalent to diversification and intensification search strategies. To improve the results, a set covering problem is solved in the final phase. They tested their heuristic on some real-world data sets provided by the company.

In follow-up work, Jarrah and Bard (2012) present an alternative approach to rationalize the design of work areas for drivers who pick up and deliver hundreds of packages a day. Considering the random nature of demand, visit frequency, and service time, their objective is to partition the customers into the minimum number of convex, continuous districts such that a single vehicle can service each district within the time available in a day. An additional requirement is that the aspect ratio of a work area must satisfy certain geometric conditions. They formulate the problem as a generic capacitated clustering problem with side constraints. They propose a solution algorithm that integrates several ideas, a combination of aggregation methods to achieve analytic tractability, column generation to determine good clusters, regeneration to diversify the exploration of the feasible region, and heuristic variable fixing to find good feasible solutions. They tested their methodology with real-world data provided by a leading carrier company ranging in size from around 6,000 to 45,000 basic units. The results showed that existing designs were significantly improved and that the number of drivers could be reduced.

2.3 Conclusions from the revised literature

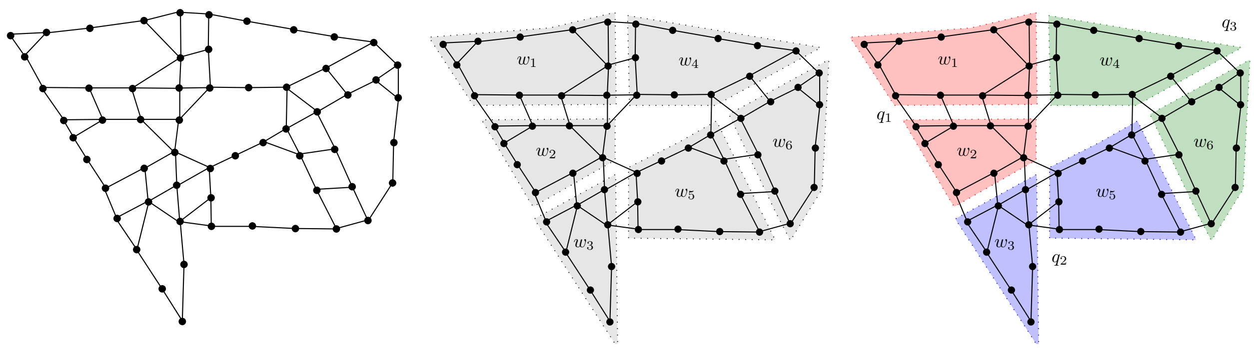
From the literature revision presented before, we can conclude that all districting design approaches that feature (last-mile) routing decisions use classical route-distance minimization criteria. Namely, the (district-wise) routing subproblem solutions are assessed in terms of a total distance measure, which is expected to be minimal. In contrast, our work is the first to consider a customer satisfaction-related measure: the proportion of customers served on time. This metric entirely changes the performance criterion's nature and makes all previous methodologies for similar problems inapplicable. In this sense, we introduce an alternative model that addresses both districting and routing decisions from a unique customer satisfaction criterion.

3 Districting design for on-time last-mile delivery

In this section we provide a mathematical framework for modeling the on-time last-mile delivery districting problem (OTLMDP). We first provide the notation and then present a formulation for the OTLMDP and for its nested routing problem.

Notation and preliminaries Let us assume that the considered urban area (e.g., a city) has been divided into a set V of customers, and let us assume that we have demand records from a set Δ of working days. Sets V and Δ are such that they effectively characterize how demand is geographically deployed and how demand varies on different days. Hence, a customer $i \in V$, is characterized by its geographical coordinates (a_i, b_i) and a demand vector $\mathbf{c}_i \in \mathbb{N}^{|\Delta|}$, so that c_i^d corresponds to the number of express orders to be delivered to customer i on day $d \in \Delta$. Let $(\mathbf{a}, \mathbf{b}) = ((a_1, b_1), \ldots, (a_{|V|}, b_{|V|}))$ and $\mathbf{c} = (\mathbf{c}_1, \ldots, \mathbf{c}_{|V|})$, also let $V^d \subset V$ be the set of all customers with programmed deliveries on day $d \in \Delta$, i.e., $V^d = \{i \in V \mid c_i^d > 0\}$. The street network is mapped into a set of edges $E \subseteq V \times V$, such that $e : \{i, j\} \in E$ indicates that customers $i, j \in V$

are adjacent in the corresponding street network. In Figure 1a we show an example of a network associated to a set of customers V and a street network mapped into a set of edges E. For any subset $\nu \subset V$, let $E(\nu) \subset E$ be the set of edges such that $\forall \{i,j\} : E(\nu)$ it holds that $i, j \in \nu$, i.e., the set of edges induced by ν. Additionally, let $\mathbf{l} : E \to \mathbb{R}_{>0}$ be a vector such that $l_{i,j}$ corresponds the time between customers i and j, for $\{i,j\} : e \in E$. Hence, the corresponding travel time vector $\mathbf{t} : V \times V \to \mathbb{R}_{\geq 0}$ is given by $t_{ij} = SP(i, j, E, \mathbf{l})$, where $SP(i, j, E, \mathbf{l})$ corresponds to the total travel time of the shortest-path between nodes $i, j \in V \mid_{i \neq j}$ considering the edge set E and the time vector $\mathbf{l}$. Finally, and without loss of generality, there is one element in V that corresponds to the *depot*; such element will denoted by ρ.



(a) Customers (V) and street network (mapped into E)

(b) (Connected) Basic units $w_1, \ldots, w_6$

(c) (Connected) Districts q_1, q_2 and q_3

Figure 1: Representation of the set of customers, edges, basic units and districts.

Let us assume that we are given a partition of V into a set W of disjoint subsets (i.e., $V = \bigcup_{w \in W} w$ and $\bigcap_{w \in W} w = \emptyset$). Elements from W are regarded as *basic units* (BUs) and they correspond to street blocks or any other convenient aggregation of customers according to urban or commercial criteria. Defining these BUs in terms of their topology and their composition may play a crucial role in the quality of the obtained solutions and in the performance of tailored algorithms for solving districting problems (further insights on this topic are outlined in Section 4). These basic units must be connected, namely, i.e., the subnetwork induced by $E(w)$ for every $w \in W$ corresponds to a connected component. In Figure 1b we display an example of how the network in Figure 1a can be divided into six connected basic units. Considering these definitions, a district corresponds to a collection of basic units; hence, a districting design, say $Q_p = \{q_1, \ldots, q_p\}$, corresponds to a partition of W into p disjoint sets, i.e., p districts. Let $\mathcal{Q}_p$ the set of all possible p-partitions. Given this definition, the k-th district q_k is a set of p_k basic units, $q_k = \{w_1^k, \ldots, w_t^k, \ldots, w_{p_k}^k\}$, each of them corresponding to a set of (connected) customers. Figure 1c shows a solution with three districts for the example network of six basic units shown in Figure 1b.

Let σ be the service time per order, which we will consider identical regardless of the type of order. Finally, let τ be the total service time *before* the express delivery deadline, i.e., only the orders handed over before τ are on-time.

Considering the notation and definitions, the OTLMDP can be defined as the problem of finding a p-partition of the set of customers, i.e., a set of p districts, so that the number of customers that are served on-time is maximized. We now present a formal mathematical programming formulation.

3.1 A mathematical programming formulation for the OTLMDP

We will now present a mathematical programming formulation for the OTLMDP. First, a districting design is said to be feasible if the corresponding p-partition $Q_p \in \mathcal{Q}_p$ is such that every set $q_k \in Q_p$ is connected, i.e., the subnetwork induced by $E(q_k)$ corresponds to a connected component. This means that the basic units that covered by district q_k (which yield connected subnetwork themselves), must also induce, altogether, a connected component. For a district $q_k \in Q_p$, composed of p_k basic units, and a day $d \in \Delta$, let $V^d(q_k) = \bigcup_{t=1}^{p_k}(w_t^k \cap V^d)$ be the set of customers with orders to be handed over on day d. Following these definitions, let $\pi(V^d(q_k))$, or simply π_k^d, be a permutation (i.e., an ordered sequence) of the customers in the k-th district with orders to be handed over on day d, and let Π_k^d be the set of all such permutations.

Suppose that customers in $V^d(q_k)$ are served by a delivery person according to a permutation π_k^d; and let $h(\pi_k^d, \mathbf{t}, \mathbf{c}, \sigma, \tau)$, or simply $h(\pi_k^d)$, be the *service performance function*, so that $h(\pi_k^d)$ corresponds to the total number of orders that are delivered on-time if customers $V^d(q_k)$ are served according to the route defined by the permutation π_k^d. Evidently, function $h(\cdot)$ depends on the permutation $\pi_{(\cdot)}^{(\cdot)}$, the travel time vector $\mathbf{t}$, the volume of express orders $\mathbf{c}$, the service time time σ and the deadline τ. Given this definition, let

$$\dot{\pi}_k^d = \underset{\pi_k^d \in \Pi_k^d}{\arg\max}\, h(\pi_k^d) \tag{(OTRP)-(k,d)}$$

be the best permutation for serving customers $V^d(q_k)$, i.e., the permutation that induces the route where more orders can be delivered on-time. For a given set of customers, $\dot{\pi}_k^{(\cdot)}$ can be interpreted as a *tour* or *routing* plan. Hence, finding $\dot{\pi}_k^{(\cdot)}$ implies finding a route of an ad-hoc single-vehicle routing problem, where the objective function differs from the traditional distance-based measure. We have coined this problem as the On-time routing problem (OTRP). The objective function of the OTRP is not a standard total distance minimization function (see Toth and Vigo, 2002, for a fundamental reference on routing problems), but an orienteering-like objective function as the goal is to maximize the number of on-time deliveries (we refer the reader to Gunawan et al., 2016, for a recent review on orienteering problems). In §3.2 we provide further details and a mixed-integer programming (MIP) formulation for the OTRP.

Additionally, for a given district $q_k \in Q_p$, let $a(w_t^k) > 0$ be the area of the t-th basic unit that constitute district q_k, and let $\rho(q_k)$ be the area of the circle that has the same perimeter as the territory induced by q_k. In both cases, the area is measured in units such as square meters. Given these definitions, let $f : q_{\cdot} \to [0, 1]$

be the compactness function, so that $f(q_k)$, i.e., the compactness of the k-th district, is given by

$$f(q_k) = \frac{\sum_{w_t^k \in q_k} a(w_t^k)}{\rho(q_k)}, \tag{1}$$

i.e., the ratio between the actual area of the district and the area of its "ideal" compact shape, a circle, and the closer to 1 the value of $f(\cdot)$ s, the more compact the corresponding district is considered. This compactness measure was first proposed by Cox (1927) and it is a common way of expressing compactness, specially when basic units are uniform as it is our case, as we will present later in this section (we refer the reader to Butsch, 2016, for a thorough review on different compactness measures in the context of territorial and districting design). In our setting, we will seek for districts whose compactness is *at least* α, with $\alpha \in [0, 1]$, which basically means to seek districts that ensure $f(\cdot) \geq \alpha$.

Considering these definitions, the OTLMDP can be defined as the problem of finding a districing design, say Q_p^*, such that the total number of express orders that are delivered on-time, from all districts and at all days, is maximum:

$$\text{(OTLMDP)} \qquad z^* = \max \sum_{d \in \Delta} \sum_{q_k \in Q_p} h(\dot{\pi}_k^d) \tag{OTLMDP.1}$$

$$\text{s.t.} \qquad f(q_k) \geq \alpha, \ \forall q_k \in Q_p \tag{OTLMDP.2}$$

$$q_k \text{ induces a connected network, } \forall q_k \in Q_p \tag{OTLMDP.3}$$

$$Q_p \in \mathcal{Q}_p. \tag{OTLMDP.4}$$

Objective function (OTLMDP.1) seeks for a districting design that maximizes the value of $h(\dot{\pi}_k^d)$ across all districts and all days, i.e., a design that maximizes the number of express orders that are handed over on-time. Constraints (OTLMDP.2) ensure that the measure of compactness of the sought districts is at least α. Similarly, constraints (OTLMDP.3) ensure that each of the selected districts induces a connected network. Finally, constraint (OTLMDP.4) ensures that the districting plan is a p-partition of the set of all customer (i.e., every customer is assigned to exactly one district).

3.2 An MIP formulation for the OTRP

Let us assume that we have computed a p-partition $Q_p \in \mathcal{Q}_p$, such that every district $q_k \in Q_p$, for $k \in \{1, \ldots, p\}$, is connected. We will now present a mixed integer programming (MIP) formulation for solving the underlying ((OTRP)-(k, d)) problem for any given district $q_k \in Q_p$ and day $d \in \Delta$.

First, let $A(q_k)$ be the set of arcs obtained by computing the so-called extended network $A(q_k) = \{(i,j),(j,i) \mid \forall i,j \mid_{i \neq j} \in V^d(q_k)\}$. The corresponding travel time vector $\mathbf{t}^k : A(q_k) \to \mathbb{R}_{\geq 0}$ is given by $t_{ij}^k = SP(i,j,E,\mathbf{l},q_k)$, where $SP(i,j,E,\mathbf{l},q_k)$ corresponds to the travel time of the shortest-path between nodes i and j considering the edge set $E(q_k)$ and vector $\mathbf{l}$. Given these definitions, let $\delta^-(i) \subset A(q_k)$ and

$\delta^+(i) \subset A(q_k)$ be, respectively, the set of *incoming* and *outgoing* arcs to and from customer $i \in V^d(q_k)$. Additionally, let $r \in V^d(q_k)$ be the closest customer to the company's depot ρ which, regardless of the route, will be the first customer served. Figure 2 displays an example of the delivery route obtained for a given day d under the above-mentioned circumstances. We can observe that the first served customer is r followed by the subset of customers that are served on-time (i.e., before the deadline encoded by τ), marked by symbol $\bullet$, and then by the customers that are not served on time, marked by symbol $\otimes$. The figure also shows the customers with no demand for day d, marked by symbol $*$.

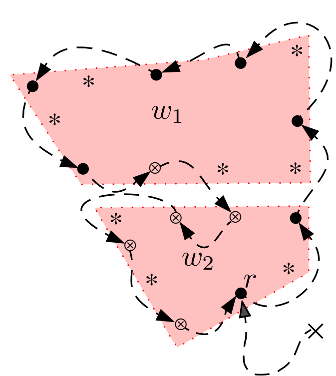


Figure 2: Example of an OTRP solution for a given day d on a district comprised by w_1 and w_2 ($\bullet$: customers served before τ (on time), $\otimes$: customers served after τ (late), $*$: customers with no demand on day d, r: customer closest to depot, $\times$: depot ρ).

Let $\mathbf{y} \in \{0,1\}^{|A(q_k)|}$ be a vector of binary variables such that $y_{i,j} = 1$ if arc $(i,j) \in A(q_k)$ is part of the last-mile route within district q_k on day $d \in \Delta$, and $y_{i,j} = 0$, otherwise. Likewise, let $\mathbf{z} \in \{0,1\}^{|V^d(q_k)|}$ be a vector of binary variables such that $z_i = 1$ if customer $i \in V^d(q_k)$ is served on-time by the delivery service, and $z_i = 0$, otherwise. Finally, let $\mathbf{u} \in \mathbb{R}_{\geq 0}^{|V^d(q_k)|}$ be a vector of non-negative continuous variables such that u_i corresponds to the arrival time at demand node $i \in V^d(q_k)$. Given these definitions, the OTRP embedded into the k-district and for day d, can be formulated as follows:

$$\text{(OTRP)-}(k,d) \quad \dot{\pi}_k^d = \arg\max \sum_{i \in V^d(q_k)} c_i z_i \tag{OTRP.1}$$

$$\text{s.t.} \quad \sum_{(i,j)\in\delta^-(j)} y_{ij} = 1, \qquad \forall j \in V^d(q_k) \tag{OTRP.2}$$

$$\sum_{(i,j)\in\delta^+(i)} y_{ij} = 1, \qquad \forall i \in V^d(q_k) \tag{OTRP.3}$$

$$u_i + \sigma c_i + t_{ij}^k \leq u_j + M_1(1 - y_{ij}), \qquad \forall (i,j) \in \delta^-(j),\ \forall j \in V^d(q_k) \setminus \{r\} \tag{OTRP.4}$$

$$u_r = 0 \tag{OTRP.5}$$

$$u_i - \tau \leq M_2(1 - z_i), \qquad \forall i \in V^d(q_k) \setminus \{r\} \tag{OTRP.6}$$

$$\mathbf{y} \in \{0,1\}^{|A(q_k)|},\ \mathbf{z} \in \{0,1\}^{|V^d(q_k)|} \text{ and } \mathbf{u} \in \mathbb{R}_{\geq 0}^{|V^d(q_k)|}. \tag{OTRP.7}$$

Objective function (OTRP.1) maximizes the number of orders that are handed over on-time. Hence, $\dot{\pi}_k^d$ encodes the best last-mile delivery route for serving customers $V^d(q_k)$. Constraints (OTRP.2) and (OTRP.3) ensure that every customer i in $V^d(q,k)$ is visited exactly once within the delivery route. Constraints (OTRP.4) account for the value of the arrival time u_j, for every customer $j \in V^d(q_k) \setminus \{r\}$ (where M_1 is an big-M auxiliary parameter): if $y_{ij} = 1$, the arrival time to customer j should be equal to the arrival time to the previously visited customer (u_i), plus the service time on customer i (σc_i), and the travel time from i to j ($t_{e:\{i,j\}}$). Without loss of generality, constraint (OTRP.5) sets to 0 the arrival time to customer r (i.e., $u_r = 0$). The activation of the $\mathbf{z}$ variables is defined by constraints (OTRP.6) (where M_2 is an additional big-M auxiliary parameter). These constraints state that if the arrival time at customer $i \in V^d(q_k) \setminus \{r\}$ is greater than the time limit τ, then variable z_i should have a positive value. Since it is a binary variable, it should have a value of 1. Hence, variable z_i indicates if customer $i \in V^d(q_k) \setminus \{r\}$ is served on time ($z_i = 1$) or if a delayed delivery occurs ($z_i = 0$). Finally, constraints (OTRP.7) set the nature of the decision variables. Note that formulation (OTRP.1)-(OTRP.7) shares similarities with the well-known Miller-Tucker-Zemlin formulations for routing problems (we refer the reader to Bektaş and Gouveia, 2014; Desrochers and Laporte, 1991; Kara et al., 2004, for revisions and enhancements of this class of formulations).

Evidently, for a day d, the performance of any district q_k is measured by comparing the corresponding value of $h(\dot{\pi}_k^d)$ (i.e., the maximum number of order delivered on-time) according to the total number of orders to be delivered on day d. As we will show in Section 5, we are particularly interested in measuring the performance of the districting design for the day with the largest number of orders, referred to as the *critical day* $\delta \in \Delta$. The idea behind this procedure is that if a districting plan performs reasonably well for such a day, then it is likely to perform well for any other day. In the next section we present the details of the devised ad-hoc scheme for solving the OTLMDP and the OTRP part of the problem.

4 A heuristic approach for solving the OTLMDP

Given an instance of the problem, the proposed heuristic finds a districting design (a p-partition) $Q_p = \{q_1, \ldots, q_k, \ldots, q_p\}$ and a routing plan for the critical day, $\delta \in \Delta$, $\boldsymbol{\pi}^\delta = \{\pi_1^\delta, \ldots, \pi_k^\delta, \ldots, \pi_q^\delta\}$, where π_k^δ is the permutation associated with the delivery route on the k-th district on day δ.

The algorithm is divided into three phases: a preprocessing phase on the input data that creates the basic units required by the solution procedure; a solution generation phase to build an initial feasible solution; and a local-search phase to improve the incumbent solution. The overall procedure is described in Algorithm 1.

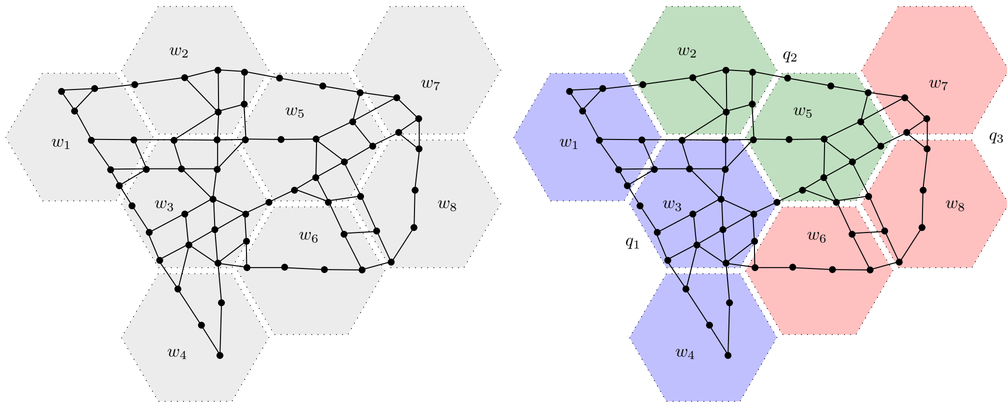
Algorithm 1: `main_algorithm()`

Input: An instance to the problem.
Output: $(Q_p^\star, \boldsymbol{\pi}^{\delta\star})$: A districting design $Q_p^\star$ and a routing plan $\boldsymbol{\pi}^{\delta\star}$ for each district on day δ.

$W \leftarrow$ `preprocessing`$(V, E, (\mathbf{a}, \mathbf{b}), \mathbf{c}, \lambda)$ //preprocessing phase
$(Q_p, \boldsymbol{\pi}^\delta) \leftarrow$ `solution_generation`$(p, W, V, E, \alpha, \mathbf{c}, \mathbf{t}, \kappa)$ //solution generation phase
$(Q_p^\star, \boldsymbol{\pi}^{\delta\star}) \leftarrow$`local_search`$(Q_p, \boldsymbol{\pi}^\delta)$ //local search phase
return $(Q_p^\star, \boldsymbol{\pi}^{\delta\star})$

As shown in Algorithm 1, the `preprocessing`() phase defines a partition of V into a set W of BUs. The partition is performed according to their coordinates $(\mathbf{a}, \mathbf{b})$ and their postal traffic encoded by vector $\mathbf{c}$. In particular, our algorithm defines BUs that result from *tessellating* the area of interest through regular hexagons; all hexagons that contain at least one demand node are candidates to be basic units. Hexagonal shapes were chosen because a hexagonal tessellation is the densest way to arrange circles in two dimensions. The *honeycomb conjecture* states that the hexagonal tessellation is the best way to divide a surface into the smallest total number of regions of equal area (Hales, 2001). Consequently, it should be easier to obtain compact shapes (compact districts) due to grouping hexagonal BUs. An additional motivation for grouping the demand nodes (customers) into polygons is to have *areas* instead of *single points*. They capture the topology of street blocks in urban areas and allow the method to easily incorporate new customers by allocating them to their corresponding BU. Figure 3 shows an example of such a tessellation (Figure 3a) and a possible district partition designed from such a geometrical composition (Figure 3b).

The tasks of the `preprocessing`() function are performed in three stages. In the first stage, a hexagonal tessellation or grid is defined, using a hexagon size λ, to ensure that each customer location falls within one hexagon and that the corresponding set of customers induces a connected component. If connectivity is not fulfilled, then λ is increased, and the process is repeated until meeting this requirement. In Figure 4a, we show the result obtained when applying the first stage to the demand points of our case study. In the second stage, all hexagons that do not contain any customers are removed from the grid. The resulting tessellation is shown in Figure 4b, where gray hexagons correspond to those removed from the tessellation defined in the first stage. As shown in Figure 4b, after removing empty BUs, the resulting tessellation (light



(a) 8 hexagonal BUs (b) 3-partition of the 8 BUs

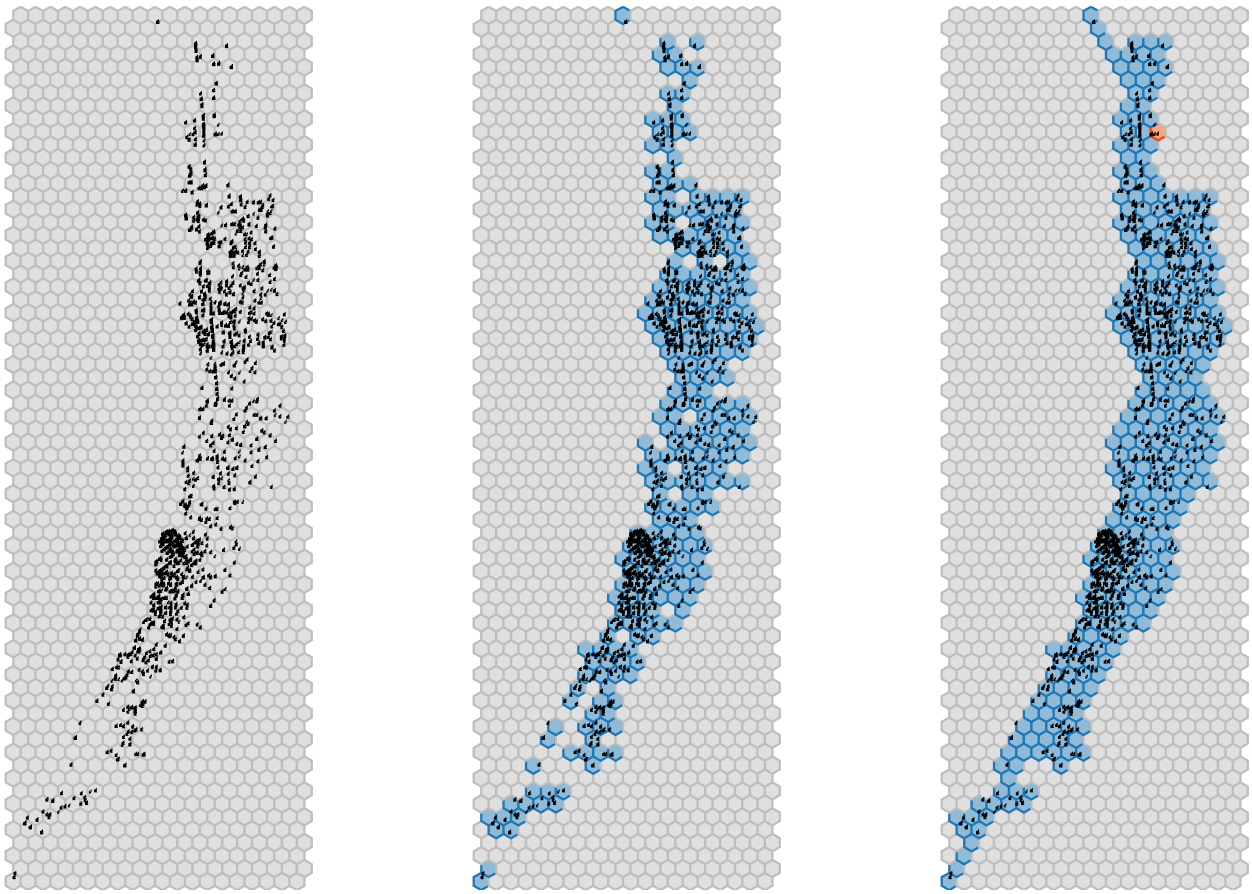
Figure 3: Example of a partition of V into 8 hexagonal BUs and a possible grouping of them in a 3-partition.

blue hexagons) does not necessarily define a single connected component. There can be *holes*, which must be avoided because potential future customers could be neglected, and compactness could be compromised. Both issues are addressed in a third stage. A specially tailored routine heuristically adds as few (empty) BUs as possible to *fill* the tessellation and amend the mentioned issues. In Figure 4c, we show the tessellation obtained after applying this third stage. This figure also highlights the BU containing the depot node ρ by coloring it red.

After defining the set of BUs W, a solution is built by calling the `solution_generation()` function, outlined in Algorithm 2. This function takes as input the desired number p of territories, the set of BUs W, the set of customers V, the set of edges E, the demand matrix $\mathbf{c}$, the distance vector $\mathbf{l}$, and the number of solutions to generate κ. The `solution_generation()` function is a randomized procedure that returns the best feasible solution, $(Q_p, \boldsymbol{\pi}^\delta)$, out of κ generated solutions. Hence, the value of κ can be tuned according to how much diversity is required and how much computational effort is allowed in this phase. Preliminary testing with the data showed that a value of $\kappa = 5$ was an adequate choice for the case under study.

The procedure to create each of the κ solutions is divided into three stages. The first stage creates p *partial* districts (lines 6 to 20 in Algorithm 2). The second stage *completes* these partial districts (lines 21 to 36), and the third stage builds the routes on the resulting districts considering only the critical day δ (lines 38 to 44).

The first stage begins by randomly selecting a BU, say s_1, from the set $W^\star = W$ (line 8). BU s_1 is regarded as a *seed*. BU s_1 is removed from set $W^\star$ (line 11; set $W^\star$ is used to store the set of unassigned BUs) and added both to the set of seeds S and to the partial district q_1 (lines 9 and 10). Afterwards, see line 13, function `fill_district()` is called, whose arguments are q_1, $W^\star$ and *fill_threshold* (in the case of the latter, its value is initially set to 0.2). This function moves elements from $W^\star$ to q_1 until the *size* of



(a) First stage　　(b) Second stage　　(c) Third stage

Figure 4: Hexagonal tessellations obtained, for our case study, in each stage of the `preprocessing()` phase.

the district reaches the $fill\_threshold$. The size of a district $q_{k'}$, $size(q_{k'})$, is defined as the ratio between the *weighted length* of the district, given by $\sum_{i \in q_{k'}} c_i t_{ri}^{k'}$ (i.e., the sum of distances of the customers to the depot ρ weighted by the corresponding customer's number of orders), and a proxy of the average length of all districts, i.e.,

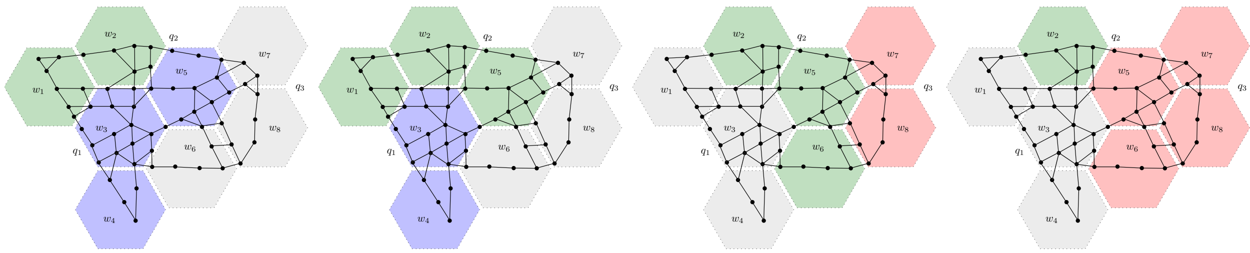
$$size(q_{k'}) = \frac{\sum_{i \in q_k} c_i t_{\rho i}^{k'}}{\frac{1}{p} \sum_{j \in V \setminus \rho} c_j t_{\rho j}}.$$

On the one hand, this measure captures the fact that the length of the route is determined by both, the number of orders and the distance to the depot. On the other hand, the value $size(q_{k'})$ approximates how balanced is the length of district $q_{k'}$ with respect to the *ideal* length $\frac{1}{p} \sum_{j \in V \setminus \rho} c_j t_{\rho j}$. Note that the output of the function `fill_district()` is a boolean value equal to `TRUE` if the district has any neighboring BU left in $W^\star$ and `FALSE` otherwise. For the remaining districts, indexed by $k = \{2, \dots, p\}$, an equivalent process is applied (lines 14 to 20). At each iteration, the corresponding seed is selected with the function `maxmin_dispersion`(), that returns the BU s_k in $W^\star$ containing the node that has the maximum-minimum distance to the elements in S (line 15). The BU s_k is removed from $W^\star$ (line 18) and added to S and to q_k (lines 16 and 17). Then, the function `fill_district`() is called as described before (line 19). At this point, the first stage is complete, and the algorithm has constructed p partial districts and a set $W^\star$ containing BUs that are still unassigned.

In the second stage (lines 21 to 36), the districts are gradually filled by increasing the value of $fill\_threshold$ and considering a random permutation at every iteration (line 24). Note that the gradual increase of the $fill\_threshold$ value avoids districts from *blocking* each other too soon. The filling process also ensures that no district should have a compactness measure lower than the compactness threshold α ($\texttt{compactness}(q_k) \leq \alpha$; line 29); if so, the current solution is discarded and the process restarts. The process finishes when no more BUs need to be assigned, i.e., all k districts are generated.

After generating a districting plan that meets the compactness constraint, the third stage (lines 38 to 44) builds the delivery routes from depot ρ to the customers of each territory considering the customers and the demand of the critical day δ (i.e., the day with the most demand). At each iteration, the ordered set π_k^δ stores the demand nodes of the k-th district, $V^\delta(q_k)$, in the order in which they are visited. An initial route is obtained by calling function `weighted_nearest_neighbor`() (line 41), which implements the well-known nearest-neighbor heuristic using as distance vector the travel time between nodes weighted by the number of orders of the destination node. This initial route is then improved by applying function `two_opt`() (line 42), which features the well-known `2-OPT` heuristic for the TSP (see Croes, 1958, for a seminal work on this matter), and evaluates each move by determining if the number of on-time served customers improves. Afterwards, the number of on-time deliveries, $h(q_k)$, is computed with function `served_orders`() (line 43). Finally, in lines 45 to 47 the obtained solution is conveniently organized, and in lines 48 to 50 we attempt to update the best solution computed so far. In consequence, the output of the `solution_generation`() algorithm, the pair $(Q_p^{\text{best}}, \boldsymbol{\pi}^{\delta,\text{best}})$, is the best solution out of κ (randomly) generated solutions.

The `local_search`() algorithm receives the solution computed by the `solution_generation`() algorithm and refines it by iteratively exploring the space of feasible solutions through a (randomized and greedy) local search scheme. The local search is based on a best-first strategy with some flexibility to explore worse solutions; therefore, there is a distinction between the incumbent solution ($Q_p^\star$) and the pivot solution (Q_p). The pivot solution is the one from which local moves are evaluated by exploring its neighbor solutions. The neighbors of the pivot solution are generated by selecting two adjacent districts, say t_0 and t_1, and creating two new (and better) districts from them by reorganizing their BUs. The designed exploration strategy first selects t_0, which corresponds to the district (from the input solution) with the least percentage of served customers to generate a neighbor solution (line 8 from Algorithm 3); t_0 is removed from the auxiliary set P_0 (line 9). Afterwards, in line 10, we define set P_1, which corresponds to the set of districts adjacent to t_0.



(a) From q_1 and q_2 (b) From q_1 and q_2 (c) From q_2 and q_3 (d) From q_2 and q_3

Figure 5: Example of solutions obtained from exploring the neighborhood of the solution shown in Figure 3b.

The iterative process encoded by the while-loop in lines 11 to 48 corresponds to the core of the exploration process. For a given district t_0, we find the adjacent district t_1 (from P_1) with the shortest route (line 12); both t_0 and t_1 are *merged* into the auxiliary set $W^\star$ (line 14). The while-loop in lines 18 to 46 computes at most n_s neighbor 2-partitions of set $W^\star$, i.e., neighboring feasible districts q'_0 and q'_1, by applying the `solution_generation`() algorithm (line 20), setting as input $p = 2$ (as only two districts must be generated), $W^\star$ (the BUs of t_0 and t_1), $V(W^\star)$ (the set of customers contained in $W^\star$), $E(V(W^\star))$ (the set of edges induced by $V(W^\star)$), α, $\mathbf{c}$, $\mathbf{t}$, and $\kappa = 1$ (as a single neighbor solution is sought at each iteration). Note that n_s can be tuned to control the size of the neighborhood to be explored at each phase (according to preliminary testing, setting $n_s = 5$ provided an adequate balance between running time and exploration breadth). In Figure 5 we show examples of how neighbor solutions, from the solution shown in Figure 3b, are computed: the solutions in Figures 5a and 5b are obtained by rearranging the BUs of the original districts q_1 and q_2, while the solutions in Figures 5c and 5d are obtained by rearranging the BUs of the original districts q_2 and q_3.

Along with the neighboring feasible solution induced by q'_0 and q'_1, we also compute $h(q'_0, q'_1)$, the number of on-time deliveries, of the new districts (considering day δ). Every time a neighbor solution is produced, three criteria are considered to decide if we *move* to this neighbor solution. The first criterion is to accept the move if the new number of on-time deliveries improves compared with the pivot solution ($h(t_0, t_1)$; lines 22 to 24). The second criterion considers that if the number of on-time deliveries is the same in the neighbor solution ($h(q'_0, q'_1) = h(t_0, t_1)$), but the balance between on-time deliveries between the two new districts improves, then the move is accepted (lines 25 to 27). The third criterion accepts a deterioration in the objective value if such move broadens the search space; i.e, we move to a worse solution if such move is allowed ($allow\_wm =$`TRUE`) and the number of on-time deliveries plus a threshold value ($h(q'_0, q'_1) + wm\_thresh$) is not less than $h(t_0, t_1)$ (lines 28 to 31).

Once a move is accepted ($move =$`TRUE`), the pivot solution Q_p is updated with the new arrangements of territories t_0 and t_1 (lines 34 to 38), and the sets P_0 and P_1 are emptied to repeat the process (lines 39 to 40). If no improvement is obtained after generating $n_s = 5$ neighbor solutions with the selected districts

t_0 and t_1 (i.e., ($move$ =`FALSE`), the while-loop in lines 11 to 47 is performed again; hence, a new adjacent district t_1 is retrieved from P_1, and the whole local-search process is carried out once more. Once all the elements in P_1 are tested without improvement, a new iteration of the while-loop in lines 7 to 48 is carried out and the next district in P_0 with the least percentage of on-time deliveries is selected as the new t_0; and set P_1 is updated accordingly.

After all the possible combinations of t_0 and t_1 are tested, the algorithm checks if the pivot solution Q_p is better than the incumbent $Q_p^\star$. Q_p is better than $Q_p^\star$, i.e., if the total number of on-time deliveries improves ($h(Q_p^\star) < h(Q_p)$), or if both values are the same ($h(Q_p^\star) = h(Q_p)$) but the district with worst performance in Q_p is better than the worst one in $Q_p^\star$ ($\min_{q \in Q_p^\star} h(q) \leq \min_{q' \in Q_p} h(q')$). If one of these conditions is fulfilled, the incumbent solution is updated (lines 47 to 50). The local-search process is repeated until the goal of serving 95% of the orders according to schedule is reached or wm_thresh reaches a value of 80 (lines 53 to 55). The value of this later parameter, and of the rest of the parameters used throughout the algorithm, were determined by a set of sensibility analysis tests with the data obtained for the case study.

All of the functions that are used in `solution_generation`() and `local_search`() (Algorithms 2 and 3, respectively), are described in detail in Table 1.

Algorithm 2: `solution_generation`$(p, W, V, E, \alpha, \mathbf{c}, \mathbf{t}, \kappa)$

Input: $p, W, V, E, \alpha, \mathbf{c}, \mathbf{t}, \kappa$

Output: $(Q_p^{\text{best}}, \boldsymbol{\pi}^{\delta,\text{best}})$: the best solution found out of κ randomly generated solutions

$K \leftarrow \{1, 2, \dots, p\}$
$(Q_p^{\text{best}}, \boldsymbol{\pi}^{\delta,\text{best}}) \leftarrow (\emptyset, \emptyset)$
for $i \in \{1, 2, \dots, \kappa\}$ **do**
    $continue \leftarrow$ `TRUE`
    while $continue =$ `TRUE` **do**
        $W^\star \leftarrow W$
        $continue \leftarrow$ `FALSE`
        $s_1 \leftarrow$ `random_element`$(W^\star)$ // select seeds and fill territories to initial threshold
        $S \leftarrow \{s_1\}$
        $q_1 \leftarrow \{s_1\}$
        $W^\star \leftarrow W^\star \setminus \{s_1\}$
        $fill\_threshold \leftarrow 0.2$
        $status_1 \leftarrow$ `fill_district`$(q_1, W^\star, fill\_threshold)$
        for $k \in \{2, \dots, p\}$ **do**
            $s_k \leftarrow$ `maxmin_dispersion`$(S, W^\star)$
            $S \leftarrow S \cup \{s_k\}$
            $q_k \leftarrow \{s_k\}$
            $W^\star \leftarrow W^\star \setminus \{s_k\}$
            $status_k \leftarrow$ `fill_district`$(q_k, W^\star, fill\_threshold)$
        end
        while $W^\star \neq \emptyset$ **do**
            //assign remaining basic units to territories
            $fill\_threshold \leftarrow fill\_threshold + 0.2$
            $R \leftarrow$ `random_permutation`(K)
            for $k \in R$ **do**
                if $status_k$ **then**
                    $status_k \leftarrow$ `fill_district`$(q_k, W, fill\_threshold)$
                end
                else if `compactness`$(q_k) \leq \alpha$ **then**
                    //restart solution
                    $continue \leftarrow$ `TRUE`
                    $W^\star \leftarrow \emptyset$
                    `EXIT FOR LOOP`
                end
            end
        end
    end
    for $k \in K$ **do**
        //build routes
        $\pi_k^\delta \leftarrow$ ordered set of the nodes in $V^\delta(q_k)$.
        $\pi_k^\delta \leftarrow$ `weighted_nearest_neighbor`$(\pi_k^\delta \cup r)$
        $\pi_k^\delta \leftarrow$ `two_opt`(π_k^δ)
        $h(q_k) \leftarrow$ `served_orders`(π_k^δ)
    end
    $Q_p^i \leftarrow \{q_1, \dots, q_p\}$
    $h^i \leftarrow \sum_{q \in Q_p} h(q)$
    $\boldsymbol{\pi}^{\delta,i} \leftarrow \{\pi_1^\delta, \dots, \pi_p^\delta\}$
    if $(Q_p^i, \boldsymbol{\pi}^{\delta,i})$ is better than $(Q_p^{best}, \boldsymbol{\pi}^{\delta,best})$ **then**
        $(Q_p^{\text{best}}, \boldsymbol{\pi}^{\delta,\text{best}}) \leftarrow (Q_p^i, \boldsymbol{\pi}^{\delta,i})$
    end
end
return $(Q_p^{\text{best}}, \boldsymbol{\pi}^{\delta,\text{best}})$

Algorithm 3: `local_search`$(Q_p, \boldsymbol{\pi}^{\delta})$

```
Input: (Q_p, π^δ)
Output: (Q_p^⋆, π^{δ,⋆}): the final OTLMDP solution computed by our heuristic
Q_p^⋆ ← Q_p and h(Q_p^⋆) ← Σ_{q∈Q_p} h(q) and
wm_threshold ← 0
allow_wm ← FALSE
continue_ls ← TRUE
while continue_ls = TRUE do
    P_0 ← Q_p
    while P_0 ≠ ∅ do
        t_0 ← district with the least percentage of on-time served orders in P_0
        P_0 ← P_0 \ t_0
        P_1 ← neighboring districts of t_0
        while P_1 ≠ ∅ do
            t_1 ← district with the shortest route length in P_1
            P_1 ← P_1 \ t_1
            W^⋆ ← t_0 ∪ t_1
            h(t_0, t_1) ← h_{t_0} + h_{t_1}
            continue_ls ← TRUE
            ns_count ← 0
            while continue_ls = TRUE do
                //creates two districts from the elements of W^⋆
                ({q'_0, q'_1}, h(q'_0, q'_1)) ← solution_generation(2, W^⋆, V(W^⋆), E(V(W^⋆)), α, c, t, κ)
                move ← FALSE
                if h(t_0, t_1) < h(q'_0, q'_1) then
                    move ← TRUE
                end
                if h(t_0, t_1) = h(q'_0, q'_1) and min(h_{t_0}, h_{t_1}) < min(h_{q'_0}, h_{q'_1}) then
                    move ← TRUE
                end
                if allow_wm = TRUE and h(t_0, t_1) ≤ h(q'_0, q'_1) + wm_thresh then
                    move ← TRUE
                    allow_wm ← FALSE
                end
                if move = TRUE then
                    //new pivot solution
                    continue_ls ← FALSE
                    Q_p ← Q_p ∪ {q'_0, q'_1} \ {t_0, t_1}
                    h(Q_p) ← Σ_{q_k∈Q_p} h_k
                    P_1 ← ∅
                    P_0 ← ∅
                end
                ns_count ← ns_count + 1
                if ns_count > 5 then
                    continue_ls ← FALSE
                end
            end
        end
    end
    if h(Q_p^⋆) < h(Q_p) or ((h(Q_p^⋆) = h(Q_p)) and (min_{q∈Q_p^⋆} h(q) ≤ min_{q'∈Q_p} h(q'))) then
        Q_p^⋆ ← Q_p and h(Q_p^⋆) ← h(Q_p) //new incumbent solution
        wm_threshold ← 0
    end
    allow_wm ← TRUE
    wm_threshold ← wm_threshold + 10
    if wm_threshold ≥ 80  or  h(Q_p^⋆) ≥ 0.95 · total_orders then
        continue_ls ← FALSE
    end
end
return (Q_p^⋆, π^{δ,⋆})
```

Table 1: Functions used in `solution_generation`() and `local_search`().

Function	Input	Output	Process
`random_element()`	A: a set of elements.	a: an element of the input set.	Selects an element from the input set according to a uniform distribution.
`fill_district()`	A: a set of BUs. B: a set of BUs. $threshold$: a float value.	$status$: a boolean value.	In an iterative process, elements from B that are neighbors of elements in A are added to A. The criteria considered to select an element is to check first if there are any elements with more than two neighbors in A, if not, the element with the highest number of orders is chosen. Every time an element is selected, it is added to set A and removed from set B. The iterative process is repeated until the size of the district reaches the limit set by the variable $threshold$. The $status$ variable is set to `TRUE` if there are still neighbors of A in set B at the end of the process, and `FALSE` otherwise.
`maxmin_dispersion()`	A: a set of BUs. B: a set of BUs.	b: an element from set B.	For every element in B it evaluates the minimum Euclidean distance from any element in A. It returns the element in B that has the maximum value for this minimum distance.
`random_permutation()`	A: a set of territory indices.	$A^\star$: an ordered set of elements in A.	It generates a random permutation of the elements in set A and stores it in the ordered set $A^\star$.
`compactness()`	A: a set of BUs*.	c: a float value	Computes the compactness of A according to (1).
`weighted_nearest_neighbor()`	A: an ordered set of demand nodes.	$A^\star$: an ordered set of demand nodes.	It builds a route with the elements of A from the depot node ρ and stored in $A^\star$. It uses a nearest neighbor heuristic in which the distance between two nodes is defined as the travel-time weighted by the number of orders in the destination node.
`two_opt()`	A: an ordered set of demand nodes.	$A^\star$: an ordered set of demand nodes.	It implements the `2-OPT` heuristic to improve the route defined by A and the result is stored in $A^\star$ The criterion to accept a move is determined by being able to serve more orders.
`served_orders()`	A: an ordered set of demand nodes.	s: an integer value.	It evaluates the number of served orders in the route defined by the ordered set A.
`min()`	A: a set of float values.	a: an element of set A.	Selects the minimum value of set A and returns it in a.
`max_index()`	A: an ordered set of float values.	i: the index of an element of set A.	Selects the index of the maximum value of set A and returns it in a.

5 District (re)design for an express postal company: Results and discussion

5.1 Description of the case study

To validate the applicability of the proposed method and the quality of the solution provided by the method, we conducted experiments using instances derived from the real-life case study that originated this research. We implemented the procedure described in Section 4 in C++ and run all of the tests on a Intel Core i7-5500U CPU @ 2.40 computer with 8GB of RAM running Windows 10 operating system.

The area of the case study is the city of Antofagasta. Antofagasta is the fifth most populated city in Chile, with over 350,000 inhabitants and the capital of the mining region of Chile. The city is located in the northern part of the country, isolated from other inhabited areas and other major cities, and follows the Pacific coast, creating a strange city shape that spans over forty kilometers long but less than two kilometers wide in its widest point. Given its population and its significant contribution to the country's economy and numerous emigrants, both from other areas of Chile and other countries, the city has an above-average demand for express courier and parcel services. Moreover, its shape hinders express delivery operations as the farthest points from the distribution center are located almost 30 kilometers apart and are linked by roads prone to traffic jams. Consequently, the city is considered as a good test-bed to check the ability of the method to find robust solutions to the problem at hand.

Table 2: Daily postal traffic information. For each day considered within the case study, number of customers and orders for express deliveries. The average value among the 19 days is also reported. *Day 8 corresponds to the critical day δ.

Day	Customers	Orders	Day	Customers	Orders
1	163	313	11	181	272
2	166	235	12	235	383
3	198	306	13	198	385
4	184	312	14	188	392
5	158	219	15	191	313
6	109	163	16	132	178
7	107	159	17	195	309
8*	340	722	18	131	251
9	198	385	19	215	355
10	173	258	Average	182.2	311.1

The company deploys its last-mile delivery efforts by dividing the city's urban area into ten districts ($p = 10$). This strategy respects long-term agreements made several years ago with the city council and the state-owned postal service company. The current on-time level of service is 88.8%, but this level of service falls below 75% on days with great demand. Hence, the current performance is significantly below the 95%

goal of on-time express deliveries. For the computational experiences, the company provided the postal traffic of 19 working days with the location and number of express deliveries associated with each day and customers. Table 2 provides the number of (nodes) and the number of express deliveries to perform within each of these 19 days. Note that the variation between the number of orders and the number of customers comes from many express deliveries requested by specific customers such as mining and retail companies. An analysis of Table 2 shows that the day identified as #8 is the busiest day, i.e., day δ. Consequently, during the heuristic procedure, routes are evaluated according to the demand of day δ.

During the first phase of the solution procedure described in Section 4, the city was divided into hexagons according to the hexagonal grid shown in Figure 4a. Each hexagon size was chosen such that at most fifteen customers (nodes) with demand in any of the considered days were contained within one hexagon. This level of granularity allowed us to run the heuristic within reasonable times. Additionally, the *fill_threshold* and α were set to 0.2 and 0.1, respectively. We observed that such values were adequate to avoid blocking during the constructive part of the heuristic procedure for the instance under study. Finally, the algorithm was run ten separate times to validate its performance. The average running time of each execution of the algorithm was below one minute of CPU time (48.8 seconds). Section 5.2 provides an analysis of the quality of the results and sensitivity analysis concerning different demand levels.

5.2 Results and discussion

Table 3 reports the results of ten independent runs of the proposed procedure considering different numbers of districts p; in particular, we consider from $p = 8$ to 12 (recall that the company is, preliminary, constrained to set $p = 10$).

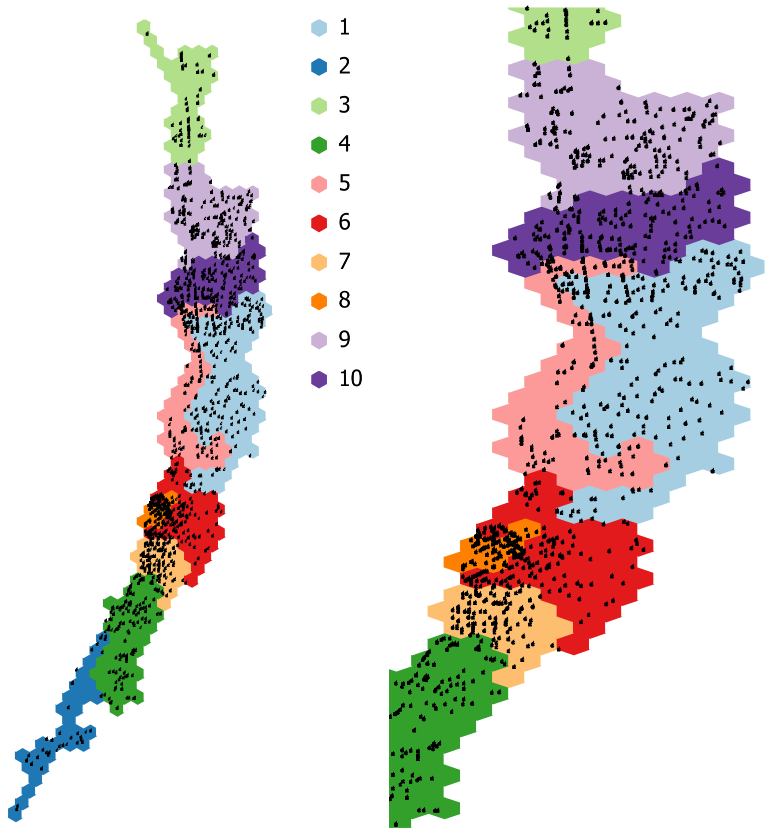
Table 3: Delivery operation performance, expressed by the average percentage of on-time express deliveries for the 19 days and for the critical day δ, attained by ten independent tests (rows 1 to 10) and different number of districts (from $p = 8$ to $p = 12$).

	Performance for all days					Performance for day δ				
	8	9	10	11	12	8	9	10	11	12
Test No.	% of on-time express deliveries					% of on-time express deliveries				
1	93.0	97.9	99.0	97.3	99.9	75.6	82.5	89.0	86.8	97.7
2	94.2	96.7	98.9	99.2	99.8	76.6	82.8	89.1	94.1	96.9
3	93.7	95.1	98.6	99.2	99.9	75.0	82.8	88.7	94.1	97.4
4	92.1	98.0	98.8	97.0	99.8	74.1	83.5	88.8	84.5	98.6
5	93.6	97.3	99.0	98.0	99.4	75.2	84.3	88.0	90.5	95.5
6	94.5	95.0	97.5	97.2	99.7	74.7	82.9	87.3	87.7	95.8
7	94.3	97.1	97.6	97.8	99.4	76.6	82.1	87.0	87.4	92.7
8	93.5	98.1	98.4	98.6	99.9	74.5	82.1	86.8	88.2	98.8
9	95.2	98.0	96.3	96.7	99.3	75.0	82.6	87.6	86.2	93.8
10	95.5	96.3	98.5	99.6	99.8	72.8	83.1	88.0	93.6	98.0
Average (%)	94.0	97.0	98.3	98.1	99.7	75.0	82.9	88.0	89.3	96.5

The columns “Performance for all days” correspond to the average percentage of on-time express deliveries within the 19 days, while the columns “Performance for day δ” correspond to the percentage of on-time express deliveries for the critical day δ. Let us consider the average performance on the 19 days. We observe that the results reported in the column corresponding to $p = 10$ show that our strategy defines a territorial design that improves the current one by 10% when considering the average percentage of on-time express deliveries (from 88.8% to 98.8%). Furthermore, it also ensures that, on average, the level of service is achieved (for all tests, we observe that the percentage of on-time express deliveries is above 95%). Additionally, if we consider $p = 9$, the reported results show that it is possible to achieve, on average, the desired level of service while simultaneously reducing the number of territories. Complementary, if we consider the performance for the critical day δ, we observe that for $p = 10$, the performance of the proposed design is below the expected service level (88.0%, in average); however, it is considerably better than the performance of the current districting design (below 75% on days with great demand). For the case of $p = 9$, we can observe that the performance associated with the critical day δ can be as low as 82.1% (tests No. 7 and 8), showing the risks of reducing the number of territories. Furthermore, when analyzing the results obtained for $p = 11$ and $p = 12$, we observe that although the average percentage of on-time express deliveries is nearly 100% (see values on columns “Performance for all days”), the performances for the critical day might still be below the desired service level. As a matter of fact, for $p = 12$, we can observe that the districting designs produced by tests No. 7 and 9 associate a percentage of on-time express deliveries below 95%.

For $p = 10$, among the ten independent runs, the best solution corresponds to the one associated with test No. 5 in Table 3; this solution is graphically depicted in Figure 6a. The customer locations are depicted as black dots, with the hexagons from each district being depicted in different colors. The visual representation shows that districts with a higher density of customers tend to be smaller and districts that are located further south; that is, further from the depot, located in the northern part of the city. Additionally, districts tend to divide the map either into territories that follow the city's shape (such as district #5) or slice the city horizontally (such as district #10). These patterns are natural within the current districts as the shape encourages deliveries that follow *easy to perform* routes, either going north to south or east to west during delivery. Figure 6b offers a more detailed view of the city's central area, where most customers are located.

Table 4 provides a more detailed analysis of the best-found solution with ten districts. In this table, we report the percentage of on-time deliveries per district and day of service. Given that the proposed method only considers the critical day within the heuristic, these results allow us to evaluate if such procedure affects the quality in the remaining days and detect if some districts perform significantly worse than others. The table reports a good performance among all days and routes. While district #6, one of the central districts, has worse performance metrics than the remaining districts, it is not consistently the worst district every day of service. Also, note that the designed districts enable a last-mile operation with an outstanding performance for most days. There are only a few cases where the performance is significantly low. If we



(a) Best solution, $p = 10$. (b) Best solution zoom.

Figure 6: Best found districting design for $p = 10$.

exclude critical day #8, there are only 4 cases in which the corresponding performance does not meet the expected level of service.

Furthermore, if we consider the total number of orders, it is only for the critical day that our design does not achieve the required service level. These results are quite promising as they improve the performance of the current design of the courier. Moreover, due to their consistency, they allow the last-mile manager to focus on more specific aspects of the decision-making setting to improve their operation. They look at, on a particular district (for instance, district #6) or a particular type of day (for instance, those days that might be circumstantially similar to the critical day #8 and, therefore, might also associate higher levels of demand).

Finally, to measure the *robustness* of the proposed districting design (and likely) change in the number of

Table 4: Delivery operation performance induced by the best found solution for $p = 10$. For each day within the case study (rows) the average percentage of on-time deliveries among all districts (column "Total (%)") and each territory (columns "1" to "10") are reported. The average among the 19 days is also reported (row "Average (%)").

		District									
		1	2	3	4	5	6	7	8	9	10
Day	Total (%)	% of on-time express deliveries									
1	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
3	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
4	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
5	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
6	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
7	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
8	84.9	68.0	100.0	97.0	89.0	79.0	73.0	93.0	92.0	81.0	98.0
9	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
10	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
11	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
12	98.2	100.0	100.0	100.0	97.0	100.0	89.0	100.0	100.0	100.0	100.0
13	95.6	100.0	100.0	100.0	77.0	100.0	100.0	100.0	100.0	100.0	100.0
14	96.7	100.0	100.0	100.0	100.0	100.0	84.0	100.0	100.0	96.0	100.0
15	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
16	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
17	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
18	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
19	97.5	100.0	100.0	100.0	86.0	100.0	95.0	100.0	100.0	100.0	100.0
Average (%)	98.6	98.3	100.0	99.8	97.3	98.9	96.9	99.6	99.6	98.8	99.9

express deliveries, we carried out a sensitivity analysis considering different levels of demand. We evaluated the performance of the districting design on synthetically generated instances, which were created by considering different levels of demand (measured by the No. of express orders). In other words, the proposed algorithm was used for solving only the underlying instances of the last-mile delivery problem (i.e., the OTRP), as the districting design is fixed. These additional instances were created according to the following two-step procedure:

- First, we set the number of customers of the day. Since the average number of orders is close to 300 orders per day, we consider an increasing number of orders from 300 to 500 in increments of 25. We stopped at 500 orders as it was clear that serving at least 95% of the orders on time could not be met with ten districts. A larger number of districts would be required to deal with such a demand, resulting in the results being reported in Table 3 for the day with maximum demand δ.

- Second, we *pool* the customers of the 19 days in the case study with their corresponding number of orders. Then, we randomly select customers and the number of orders, following a non-replacement strategy, until the number of orders is reached.

Note that demand points that appear on different days are added to the pool as separate entities. However, once a customer is chosen, it is removed from the pool. All other demand nodes with the same coordinates (with different service days) are also removed. For each number of orders, and to ensure consistency of the obtained results, we build ten synthetic days. In total, 90 additional instances resulted from this process.

Table 5: Sensitivity analysis of the delivery operation performance for different levels of demand ranging from 300 to 500 orders (column "No. orders"), we report the average percentage of on-time deliveries among all districts (column "Average (%)"), and the percentage of on-time deliveries among corresponding to each district (columns "1" to "10").

		District									
		1	2	3	4	5	6	7	8	9	10
No. orders	Average (%)	% of on-time express deliveries									
300	99.9	99.7	100.0	100.0	100.0	100.0	99.8	100.0	100.0	100.0	100.0
325	99.2	98.8	100.0	99.5	100.0	100.0	96.0	100.0	100.0	100.0	100.0
350	99.2	98.3	100.0	100.0	99.7	100.0	98.3	98.7	100.0	100.0	99.6
375	98.7	98.0	100.0	100.0	99.3	99.6	94.7	100.0	100.0	99.0	99.5
400	97.6	95.3	100.0	98.3	96.9	98.9	93.3	100.0	100.0	99.8	98.8
425	96.7	96.8	100.0	99.3	99.0	100.0	85.2	98.2	100.0	98.3	98.6
450	93.0	88.0	100.0	99.2	93.3	100.0	83.0	100.0	98.9	91.7	94.0
475	89.9	84.1	100.0	98.6	91.2	97.1	73.5	100.0	96.1	87.0	95.6
500	91.6	85.8	100.0	98.8	88.2	96.6	83.0	97.3	100.0	87.1	92.5

In Table 5, we report the (average) results obtained when solving the underlying last-mile delivery problem on each district and for different demand levels (recall that ten instances were generated for each value of the No. of orders). The obtained results show that the proposed districting design can perform according to the expected 95% on-time target even if the average number of orders is as high as 425 (which is 33% higher than its current level). Even if the demand increases to an average of 500 orders per day, our districting design ensures the expected service level in 5 out of 10 districts, with a minimum value of 83.0% among those below the on-time target. Furthermore, for such a level of demand, the average percentage of on-time express deliveries is 91.6%, which is better than the current performance of the courier (88.8% of on-time express deliveries). It is worth mentioning that the performance reported when considering an average of 325 orders (99.2% of on-time deliveries) is similar to the performance featured by the same districting design on the original input demand data (98.6% of on-time deliveries, see Table 4).

The results reported in Table 5 show the capacity of the proposed methodology for effectively providing districting designs capable of enduring future changes in the demand. Such a feature is attractive for the company, the customers, and mail carriers or contractors who seek certainty and stability for districts that they are committed to delivering.

6 Conclusions and future work

The irruption of the COVID-19 pandemic and the different sanitary measures adopted by several countries have increased the volume of postal traffic and modified customers' habits, stressing the supply chain and decreasing the service level of postal operators. Considering this context, in this work, we have addressed a new districting design problem arising in last-mile postal supply chain operations. Hence, instead of traditional dispersion or routing costs metrics, the addressed problem considers a customer-satisfaction criterion in the objective function. We have provided a mathematical formulation for this problem and designed a three-phase heuristic algorithm for finding feasible solutions. The construction phase of the heuristic is itself a three-phase procedure where p partial districts are first created, then the districts are completed by assigning the remaining basic units, and finally, routes are built within each district. The procedure is enhanced using a randomization scheme to provide further diversification. The local search stage attempts to improve the solution by exploring the space of feasible solutions through a randomized and greedy scheme. To this end, we propose a novel neighbor topology that intelligently exploits the particular structure of this problem.

The proposed framework was applied to a case study from a Chilean postal operator. The obtained results show that the proposed approach is capable of defining a districting design that, considering the whole urban area, improves the performance of the current design by allowing 98.8% of on-time express deliveries (compared to the 88.8% of the current design), achieving the 95.0% target for on-time express deliveries. Furthermore, it also ensures that this service level target is achieved in each district (as no district exhibits, on average, a performance lower than 96.6%).

Given the uncertainty and variability of the demand, a sensitivity analysis is conducted to assess the performance of the proposed solution for higher levels of demand. The obtained results show that the proposed districting design can endure a significant rise in the demand in the forthcoming future (up to a 40% increase) while still ensuring the 95.0% target for on-time express deliveries. This feature of the provided solutions is attractive for the company and the clients and for mail carriers or contractors who seek certainty and stability regarding the districts they are committed to delivering.

One direction for future work is a better characterization of how routing decisions are made and evaluated. Another line of work is the study on how future demand growths are integrated within the decision stage and how incentives are designed to align the goals of the outsourced companies responsible for delivery within each district to the company's overall strategy. Territories are held strict even if collaboration among outsourcing companies responsible for adjacent territories may lead to an overall increase of the performance indicators. The latter point involves changes on the current organization that go beyond the scope of this work. However, such changes may lead to significant improvements and new districting problems that consider the need to adapt territories when the workload is greater than the expected one. Finally, from a methodological perspective, it could be worth investigating how the designed algorithm's components can be complemented with other metaheuristic frameworks enhancing even further the solution quality.

Acknowledgments: Eduardo Álvarez-Miranda acknowledges the support of the National Agency of Research and Development (ANID), Chile, through the grant FONDECYT N.1180670 and the Complex Engineering Systems Institute ANID PIA/BASAL AFB180003. Roger Ríos-Mercado was supported by the Mexican Council for Science and Technology under its Frontier Science Program (grant FC2016-2/1948).

References

Banker, S. (2020). What will last-mile delivery look like post-coronavirus? FORBES, `https://www.forbes.com/sites/stevebanker/2020/07/24/what-will-last-mile-delivery-look-like-post-coronavirus/?sh=527662133b22`, 24 July.

Bard, J. F. and Jarrah, A. I. (2009). Large-scale constrained clustering for rationalizing pickup and delivery operations. *Transportation Research, Part B: Methodological*, 43(5):542–561.

Bates, O., Friday, A., Allen, J., Cherrett, T., McLeod, F., Bektas, T., Nguyen, T., Piecyk, M., Piotrowska, M., and Wise, S. (2018). Transforming last-mile logistics: Opportunities for more sustainable deliveries. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–14, Montreal, Canada.

Behroozi, M. and Carlsson, J. G. (2020). Computational geometric approaches to equitable districting: A survey. In Ríos-Mercado, R. Z., editor, *Optimal Districting and Territory Design*, volume 284 of *International Series in Operations Research and Management Science*, chapter 4, pages 57–74. Springer, Cham, Switzerland.

Bektaş, T. and Gouveia, L. (2014). *Requiem* for the Miller–Tucker–Zemlin subtour elimination constraints? *European Journal of Operational Research*, 236(3):820–832.

Bender, M., Kalcsics, J., and Meyer, A. (2020). Districting for parcel delivery services – A two-stage solution approach and a real-world case study. *Omega*, 96:3102283.

Bender, M., Kalcsics, J., Nickel, S., and Pouls, M. (2018). A branch-and-price algorithm for the scheduling of customer visits in the context of multi-period service territory design. *European Journal of Operational Research*, 269(1):382–396.

Benzarti, E., Sahin, E., and Dallery, Y. (2013). Operations management applied to home care services: Analysis of the districting problem. *Decision Support Systems*, 55(2):587–598.

Butsch, A. (2016). *Districting Problems: New Geometrically Motivated Approaches*. Phd thesis, Karlsruher Institut für Technologie, Karlsruhe, Germany.

Butsch, A., Kalcsics, J., and Laporte, G. (2014). Districting for arc routing. *INFORMS Journal on Computing*, 26(4):809–824.

Caro, F., Shirabe, T., Guignard, M., and Weintraub, A. (2004). School redistricting: Embedding GIS tools with integer programming. *Journal of the Operational Research Society*, 55(8):836–849.

Chapleau, L., Ferland, J. A., and Rousseau, J. M. (1985). Clustering for routing in densely populated areas. *European Journal of Operational Research*, 20(1):48–57.

Cox, E. (1927). A method of assigning numerical and percentage values to the degree of roundness of sand grains. *Journal of Paleontology*, 1(3):179–183.

Crew, M. and Brennan, T., editors (2015). *Postal and Delivery Innovation in the Digital Economy*, volume 50 of *Topics in Regulatory Economics and Policy*. Springer, Cham, Switzerland.

Croes, G. (1958). A method for solving traveling-salesman problems. *Operations Research*, 6(6):791–812.

de la Poix de Fréminville, P., Desaulniers, G., Rousseau, L.-M., and Perron, S. (2015). A column generation heuristic for districting the price of a financial product. *Journal of the Operational Research Society*, 66(6):965–978.

Demazière, D. and Mercier, D. (2006). Mail carriers' routes: Managerial standards, collective regulation and work activity strategies. *Sociologie du travail*, 48:e1–e21.

Desrochers, M. and Laporte, G. (1991). Improvements and extensions to the miller-tucker-zemlin subtour elimination constraints. *Operations Research Letters*, 10(1):27–36.

Dragendorf, J., Mohr, D., Ecker, T., Neuhaus, F., Briest, P., Beretzky, E., and Zimmermann, T. (2019). The endgame for postal networks: How to win in the age of e-commerce. Report, McKinsey & Company.

Duque, J. C., Ramos, R., and Suriñach, J. (2007). Supervised regionalization methods: A survey. *International Regional Science Review*, 30(3):195–220.

Fernández, E., Kalcsics, J., Nickel, S., and Ríos-Mercado, R. Z. (2010). A novel maximum dispersion territory design model arising in the implementation of the WEEE-directive. *Journal of the Operational Research Society*, 61(3):503–514.

García-Ayala, M. G., González-Velarde, J. L., Ríos-Mercado, R. Z., and Fernández, E. (2016). A novel model for arc territory design: Promoting Eulerian districts. *International Transactions of Operational Research*, 23(3):433–458.

Gentry, S., Chow, E., Massie, A., and Segev, D. (2015). Gerrymandering for justice: Redistricting U.S. liver allocation. *Interfaces*, 45(5):462–480.

Gunawan, A., Lau, H., and Vansteenwegen, P. (2016). Orienteering problem: A survey of recent variants, solution approaches and applications. *European Journal of Operational Research*, 255(2):315–332.

Hales, T. (2001). The honeycomb conjecture. *Discrete & Computational Geometry*, 25(1):1–22.

Haugland, D., Ho, S. C., and Laporte, G. (2007). Designing delivery district for the vehicle routing problem with stochastic demands. *European Journal of Operational Research*, 180(3):997–1010.

Hong, S. and Wolak, F. (2008). Relative prices and electronic substitution: Changes in household-level demand for postal delivery services from 1986 to 2004. *Journal of Econometrics*, 145(1-2):226–242.

Huerta-Muñoz, D. L., Ríos-Mercado, R. Z., and Ruiz, R. (2017). An iterated greedy heuristic for a market segmentation problem with multiple attributes. *European Journal of Operational Research*, 261(1):75–87.

International Post Corporation (2019). Global Postal Industry Report 2019: Key Findings. `https://www.ipc.be/sector-data/postal-sector/key-findings`.

Jarrah, A. I. and Bard, J. F. (2012). Large-scale pickup and delivery work area design. *Computers & Operations Research*, 39(12):3102–3118.

Kalcsics, J. and Ríos-Mercado, R. Z. (2019). Districting problems. In Laporte, G., Nickel, S., and Saldanha da Gama, F., editors, *Location Science*, chapter 25, pages 703–741. Springer, Cham, Switzerland, 2nd edition.

Kara, I., Laporte, G., and Bektaş, T. (2004). A note on the lifted miller–tucker–zemlin subtour elimination constraints for the capacitated vehicle routing problem. *European Journal of Operational Research*, 158(3):793–795.

Konur, D. and Geunes, J. (2019). Integrated districting, fleet composition, and inventory planning for a multi-retailer distribution system. *Annals of Operations Research*, 273(1):527–559.

Laseinde, O. and Mpofu, K. (2017). Providing solution to last mile challenges in postal operations. *International Journal of Logistics Research and Applications*, 20(5):475–490.

Lee, H., Chen, Y., Gillai, B., and Rammohan, S. (2016). Technological disruption and innovation in last-mile delivery. White paper, Stanford Graduate School of Business, Stanford.

Lei, H., Laporte, G., and Guo, B. (2012). Districting for routing with stochastic customers. *EURO Journal on Transportation and Logistics*, 1(1–2):67–85.

Lei, H., Laporte, G., Liu, Y., and Zhang, T. (2015). Dynamic design of sales territories. *Computers & Operations Research*, 56:84–92.

Lei, H., Wang, R., and Laporte, G. (2016). Solving a multi-objective dynamic stochastic districting and routing problem with a co-evolutionary algorithm. *Computers & Operations Research*, 67:12–24.

Liberatore, F., Camacho-Collados, M., and Vitoriano, B. (2020). Police districting problem: Literature review and annotated bibliography. In Ríos-Mercado, R. Z., editor, *Optimal Districting and Territory Design*, volume 284 of *International Series in Operations Research and Management Science*, chapter 2, pages 9–29. Springer, Cham, Switzerland.

Lin, H. and Kao, J. (2008). Subregion districting analysis for municipal solid waste collection privatization. *Journal of the Air & Waste Management Association*, 58(1):104–111.

López-Pérez, J. F. and Ríos-Mercado, R. Z. (2013). Embotelladoras ARCA uses operations research to improve territory design plans. *Interfaces*, 43(3):209–220.

Mansfield, R., Todd, D., and Wheeler, J. (2003). Company structure and market strategy. *Omega*, 6(2):133–138.

Mayorga, M. E., Bandara, D., and McLay, L. A. (2013). Districting and dispatching policies for emergency medical service systems to improve patient survival. *IIE Transactions in Healthcare Systems Engineering*, 3(1):39–56.

Moreno, S., Pereira, J., and Yushimito, W. (2020). A hybrid K-means and integer programming method for commercial territory design: a case study in meat distribution. *Annals of Operations Research*, 286:87–117.

Moya-García, J. G. and Salazar-Aguilar, M. A. (2020). Territory design for sales force sizing. In Ríos-Mercado, R. Z., editor, *Optimal Districting and Territory Design*, volume 284 of *International Series in Operations Research and Management Science*, chapter 10, pages 191–206. Springer, Cham, Switzerland.

Muyldermans, L., Cattryse, D., Van Oudheusden, D., and Lotan, T. (2002). Districting for salt spreading operations. *European Journal of Operational Research*, 139(3):521–532.

Muyldermans, L., Cattrysse, D., and Oudheusden, D. V. (2003). District design for arc-routing applications. *Journal of the Operational Research Society*, 54(11):1209–û1221.

Nikali, H. (2008). Substitution of letter mail for different sender-receiver segments. In Crew, M. and Kleindorfer, P., editors, *Competition and Regulation in the Postal and Delivery Sector*, chapter 7, pages 89–106. Edward Elgar Publishers, Cheltenham, UK.

Nyrop, N., Nathan, A., Lindquist, M., and Karlsen, J. (2020). COVID-19 will permanently change e-commerce in Denmark. Deloitte blog, posted 27 May 2020, URL: https://www2.deloitte.com/dk/da/blog/home-blog-collection/2020/blog-martin-nyrop-digital-customer.html.

OECD (2020). E-commerce in the time of COVID-19. http://www.oecd.org/coronavirus/policy-responses/e-commerce-in-the-time-of-covid-19-3a2b78e8/.

Parriani, T., Pozzi, M., Vigo, D., and Cruijssen, F. (2019). Creation of optimal service zones for the delivery of express packages. In Dell'Amico, M., Gaudioso, M., and Stecca, G., editors, *A View of Operations Research Applications in Italy, 2018*, volume 2 of *AIRO Springer Series*, pages 19–28. Springer, Cham, Switzerland.

Restrepo, M. I., Rousseau, L.-M., and Vallée, J. (2020). Home healthcare integrated staffing and scheduling. *Omega*, 95:102057.

Ricca, F., Scozzari, A., and Simeone, B. (2013). Political districting: From classical models to recent approaches. *Annals of Operations Research*, 204(1):271–299.

Ríos-Mercado, R. Z., editor (2020). *Optimal Districting and Territory Design*, volume 284 of *International Series in Operations Research and Management Science*. Springer, Cham, Switzerland.

Ríos-Mercado, R. Z., Álvarez-Socarrás, A. M., Castrillón, A., and López-Locés, M. C. (2021). A location-allocation-improvement heuristic for districting with multiple-activity balancing constraints and p-median-based dispersion minimization. *Computers & Operations Research*, 126:105106.

Ríos-Mercado, R. Z. and Bard, J. F. (2019). An exact algorithm for designing optimal districts in the collection of waste electric and electronic equipment through an improved reformulation. *European Journal of Operational Research*, 276(1):259–271.

Ríos-Mercado, R. Z. and Escalante, H. J. (2016). GRASP with path relinking for commercial districting. *Expert Systems with Applications*, 44:102–113.

Ríos-Mercado, R. Z. and Fernández, E. (2009). A reactive GRASP for a commercial territory design problem with multiple balancing requirements. *Computers & Operations Research*, 36(3):755–776.

Ríos-Mercado, R. Z., González-Velarde, J. L., and Maldonado-Flores, J. R. (2020). Tabu search with strategic oscillation for improving collection assignment plans of waste electric and electronic equipment. *International Transactions of Operational Research*. Forthcoming, DOI: 10.1111/itor.12917.

Ríos-Mercado, R. Z. and López-Pérez, J. F. (2013). Commercial territory design planning with realignment and disjoint assignment requirements. *Omega*, 41(3):525–535.

Ríos-Mercado, R. Z. and Salazar-Acosta, J. C. (2011). A GRASP with strategic oscillation for a commercial territory design problem with a routing budget constraint. In Batyrshin, I. and Sidorov, G., editors, *Advances in Soft Computing*, volume 7095 of *Lecture Notes in Artificial Intelligence*, pages 307–318. Springer, Heidelberg, Germany.

Ronen, D. (1983). Sales territory alignment for sparse accounts. *Omega*, 11(5):501–505.

Salazar-Aguilar, M. A., Ríos-Mercado, R. Z., and Cabrera-Ríos, M. (2011a). New models for commercial territory design. *Networks & Spatial Economics*, 11(3):487–507.

Salazar-Aguilar, M. A., Ríos-Mercado, R. Z., and González-Velarde, J. L. (2011b). A bi-objective programming model for designing compact and balanced territories in commercial districting. *Transportation Research Part C: Emerging Technologies*, 19(5):885–895.

Sandoval, M. G., Díaz, J. A., and Ríos-Mercado, R. Z. (2020). An improved exact algorithm for a territory design problem with p-center-based dispersion minimization. *Expert Systems with Applications*, 146:113150.

Schneider, M., Stenger, A., Schwahn, F., and Vigo, D. (2014). Territory-based vehicle routing in the presence of time-window constraints. *Transportation Science*, 49(4):732–751.

Steiner, M. T. A., Datta, D., Neto, P. J. S., Scarpin, C. T., and Figueira, J. R. (2015). Multi-objective optimization in partitioning the healthcare system of Parana state in Brazil. *Omega*, 52:53—-64.

Toth, P. and Vigo, D. (2002). *The Vehicle Routing Problem*. SIAM, Philadelphia.

Uni Global Union (2019). The economic and social consequences of postal services liberalization. Report, available at: `https://www.uni-europa.org/wp-content/uploads/2019/10/Report-postal\_EN.pdf`.

Universal Postal Union (2019). Postal Development Report 2019: Perspectives on the performance of postal operators worldwide. `https://www.upu.int/en/Publications/2IPD/Postal-Development-Report-2019`.

Universal Postal Union (2020). Postal Development Report 2020: Achieving higher performance amid a major crisis. `https://www.upu.int/en/Publications/2IPD/Postal-Development-Report-2020`.

Verhoest, K. and Sys, J. (2006). Liberalisation, privatisation and regulation in the Belgian postal services sector. Technical report, Katholike Universiteit Leuven, Leuven, Belgium.

Winkenbach, M. and Janjevic, M. (2018). Classification of last-mile delivery models for e-commerce distribution: A global perspective. In Taniguchi, E. and Thompson, R. G., editors, *City Logistics 1: New Opportunities and Challeenges*, chapter 11, pages 209–229. Wiley, New York.

Yanık, S. and Bozkaya, B. (2020). A review of districting problems in health care. In Ríos-Mercado, R. Z., editor, *Optimal Districting and Territory Design*, volume 284 of *International Series in Operations Research and Management Science*, chapter 3, pages 31–55. Springer, Cham, Switzerland.

Yanık, S., Kalcsics, J., Nickel, S., and Bozkaya, B. (2019). A multi-period multi-criteria districting problem applied to primary care scheme with gradual assignment. *International Transactions of Operational Research*, 26(5):1676–1697.

Zhou, L., Zhen, L., Baldacci, R., Boschetti, M., Daia, Y., and Lim, A. (2021). A heuristic algorithm for solving a large-scale real-world territory design problem. *Omega*. Forthcoming, DOI: 10.1016/j.omega.2021.102442.

Zoltners, A. A. and Sinha, P. (2005). Sales territory design: Thirty years of modeling and implementation. *Marketing Science*, 24(3):313–331.