The Capacitated Vehicle Routing Problem with Evidential Demands[☆]

Nathalie Helal[a,*], Frédéric Pichon[a], Daniel Porumbel[b], David Mercier[a], Éric Lefèvre[a]

[a]*Univ. Artois, EA 3926, Laboratoire de Génie Informatique et d'Automatique de l'Artois (LGI2A), F-62400 Béthune, France.*
[b]*Conservatoire National des Arts et Métiers, EA 4629, Cedric, 75003 Paris, France.*

Abstract

We propose to represent uncertainty on customer demands in the Capacitated Vehicle Routing Problem (CVRP) using the theory of evidence. To tackle this problem, we extend classical stochastic programming modelling approaches. Specifically, we propose two models for this problem. The first model is an extension of the *chance-constrained programming* approach, which imposes certain minimum bounds on the belief and plausibility that the sum of the demands on each route respects the vehicle capacity. The second model extends the *stochastic programming with recourse* approach: for each route, it represents by a belief function the uncertainty on its recourses, *i.e.*, corrective actions performed when the vehicle capacity is exceeded, and defines the cost of a route as its classical cost (without recourse) plus the worst expected cost of its recourses. We solve the proposed models using a metaheuristic algorithm and present experimental results on instances adapted from a well-known CVRP data set.

Keywords: Vehicle routing problem, Chance constrained programming, Stochastic programming with recourse, Belief function.

1. Introduction

In the Capacitated Vehicle Routing Problem (CVRP), we are given a fleet of vehicles with identical capacity located at a depot and a set of customers with known demands located on the vertices of a graph. The goal of this problem is to determine a route for each vehicle, such that the set of routes for all the vehicles has the least total cost, all customer demands are fully serviced, the capacity

---

of each vehicle is always respected and each customer is visited by exactly one route. The CVRP is NP-hard since it contains the traveling salesman problem as a particular case (one route and unbounded capacity). It can be written as an integer linear program. The CVRP has generated a large body of research, since it belongs to the class of local transportation or delivery problems affecting the most expensive component in the distribution network [8].

Yet, many industrial applications are confronted with uncertainty on customer demands in their distribution problems involving the CVRP, and the exact customer demands are mostly revealed when the servicing vehicles arrive at the customers. Accordingly, several authors (see, *e.g.*, [28, 29] and the references therein) tackled this issue by assuming that customer demands are random variables and the associated problem is the well-known Capacitated Vehicle Routing Problem with Stochastic Demands (CVRPSD). Two of the most widely-used frameworks for modelling stochastic problems, such as the CVRPSD, are the Chance-Constrained Programming (CCP) approach and the Stochastic Programming with Recourse (SPR) approach [7]. Modelling the CVRPSD via CCP amounts to using a probabilistic capacity constraint that requires the probability of respecting the capacity constraint to be above a certain threshold. The CCP modelling technique does not consider the additional cost of recourse (or corrective) actions necessary if capacity constraints fail to be satisfied. The SPR approach does consider situations needing recourses and it aims at minimizing the initially-planned travel cost plus the expected cost of the recourses executed along routes, *e.g.*, returning to the depot and unloading in order to bring to feasibility a violated capacity constraint.

The probabilistic approach to modelling uncertainty is not necessarily well-suited to all real-life situations. In particular, its ability to handle epistemic uncertainty (uncertainty arising from lack of knowledge) has been criticised [4, 1]. The typical approach to representing basic epistemic uncertainty is the set-valued approach [26]. It is sensible when, *e.g.*, all that is known about the customer demands is that they belong to some intervals. This kind of uncertainty in the CVRP is generally addressed using robust optimisation, where one optimises against the worst-case scenario, that is, one wants to obtain solutions that are robust to all realisations of customer demands that are deemed possible (see, *e.g.*, [49]). However, the set-valued approach to uncertainty representation may be too coarse and may thus lead to solutions that are too conservative, hence not useful.

In the last forty years, the necessity to account for all facets of uncertainty has been recognized and alternative uncertainty frameworks extending both the probabilistic and set-valued ones have appeared [4]. In particular, the theory of evidence introduced by Shafer [47], based on some previous work from Dempster [15], has emerged as a theory offering a compromise between expressivity and complexity, which seems interesting in practice as its successful application in several domains testifies (see [18] for a recent survey of evidence theory applications). This theory, also known as belief function theory, may be used to model various forms of information, such as expert judgements and statistical evidence, and it also offers tools to combine and propagate uncertainty [1].

In the context of the CVRP, the theory of evidence may be used to represent uncertainty on customer demands leading to an optimisation problem, which will be referred to as the CVRP with Evidential Demands (CVRPED). Using the theory of evidence in this problem seems particularly interesting as it allows one to account for imperfect knowledge about customer demands, such as knowing that each customer demand belongs to one or more sets with a given probability allocated to each set - an intermediary situation between probabilistic and set-valued knowledge. In this paper, we propose to address the CVRPED by extending the CCP and SPR modelling approaches into the formalism of evidence theory. Although the focus will be to extend stochastic programming approaches, we will also connect our formulations with robust optimisation.

To our knowledge, evidence theory has not yet been considered to model uncertainty in large-scale instances of an NP-hard optimisation problem like the CVRP. Indeed, it seems that so far, only other non classical uncertainty theories, and in particular fuzzy set theory [50, 9, 42, 11], have been used in such problems. Besides, modelling uncertainty in optimisation problems using evidence theory has concerned only continuous design optimisation problems [1] [41, 48] and continuous linear programs [40]. Specifically in [41], the reliability of the system is optimized, while uncertainty is handled by limiting the plausibility of constraints violation into a small degree; while in [48] the problem was handled differently, and the plausibility of a constraint failure was converted into a second objective to the problem that should be minimized. Of particular interest is the work of Masri and Ben Abdelaziz [40], who extended the CCP and SPR modelling approaches, in order to model continuous linear programs embedding belief functions, which they called the Belief Constrained Programming (BCP) and the recourse approaches, respectively. In comparison, in this work, we generalise CCP and SPR to an *integer* linear program involving uncertainty modelled by evidence theory. Borrowing from [40], we propose to model the CVRPED by methods that may be called the BCP modelling of the CVRPED and the recourse modelling of the CVRPED. For both models, the resolution algorithm is a simulated annealing algorithm; we use a metaheuristic, as the CVRPED derives from the CVRP, which is NP-hard.

The paper is structured as follows. Section 2 summarises the basic preliminaries on the CVRP and on the CVRPSD modelling via CCP and SPR, along with the necessary background on evidence theory. In section 3, the BCP model and the recourse model for the CVRPED are presented and some of their properties are studied. In Section 4 we solve the BCP model and the recourse model of the CVRPED using a simulated annealing algorithm and perform experiments on instances generated from CVRP benchmarks. In Section 5, we conclude and state the perspectives of the present work.

---

[1] Designing physical systems in the engineering field using optimisation techniques, so design costs are minimized, while the system performance is fulfilled [2].

2. Background

This section recalls necessary background on the CVRP, the CVRPSD and its stochastic programming formulations, as well as some concepts of belief function theory needed in this paper.

2.1. The CVRP

In the CVRP, a fleet of m identical vehicles with a given capacity limit Q, initially located at a depot, must collect[2] goods from n customers, with d_i such that $0 < d_i \leq Q$ the deterministic collect demand of client i, $i = 1, \ldots, n$. The objective in the CVRP is to find a set of m routes with minimum cost to serve all the customers such that total customer demands on any route must not exceed Q, each route starts and ends at the depot, and each customer is serviced only once.

Formally, it is convenient to represent the depot by an artificial client $i = 0$, whose demand always equals 0, i.e., $d_0 = 0$. The CVRP may be defined on a graph $G = (V, E)$ such that $V = \{0, \ldots, n\}$ is the vertex set and $E = \{(i, j) \, | i \neq j; \; i, j \in V\}$ is the arc set. V represents the customers and the depot that corresponds to vertex 0. A travel cost (or travel time or distance – these terms are interchangeable) $c_{i,j}$ is associated with every edge in E. Travel costs are such that $c_{i,j} = c_{j,i}$, $\forall \, (i, j) \in E$ and they satisfy the triangle inequality: $c_{i,j} \leq c_{i,l} + c_{l,j}$, $\forall i, l, j \in V$. Besides, $c_{i,i} = +\infty$, $\forall i \in V$ [51]. Let R_k be the route associated to vehicle k and $w_{i,j}^k$ a binary variable that equals 1 if vehicle k travels from i to j and serves j (except if j is the depot), and 0 if it does not. A proper formulation for the CVRP [8, 38] is:

$$\min \sum_{k=1}^{m} C(R_k), \tag{1}$$

where

$$C(R_k) = \sum_{i=0}^{n} \sum_{j=0}^{n} c_{i,j} w_{i,j}^k, \tag{2}$$

subject to

_____

[2]The problem can also be presented in terms of delivery, rather than collection, of goods.

$$\sum_{i=0}^{n}\sum_{k=1}^{m} w_{i,j}^k = 1, \qquad\qquad j = 1, \ldots, n, \qquad\qquad (3)$$

$$\sum_{i=0}^{n} w_{i,\ell}^k = \sum_{j=0}^{n} w_{\ell,j}^k, \qquad\quad k = 1, \ldots, m\,, \ell = 0, \ldots, n, \qquad (4)$$

$$\sum_{j=1}^{n} w_{0,j}^k \leq 1, \qquad\qquad\quad k = 1, \ldots, m, \qquad\qquad (5)$$

$$\sum_{\substack{i,j \in L \\ i \neq j}}\sum_{k=1}^{m} w_{i,j}^k \leq |L| - 1, \qquad\qquad L \subseteq V \setminus \{0\}, \qquad\quad (6)$$

$$\sum_{i=1}^{n} d_i \sum_{j=0}^{n} w_{i,j}^k \leq Q, \qquad\qquad k = 1, \ldots, m. \qquad\qquad (7)$$

Constraints (3) make sure that exactly one vehicle arrives at client j, $j = 1, \ldots, n$. Constraints (4) ensure the continuity of the routes (flow): if vehicle k leaves vertex ℓ, vehicle k must also enter vertex ℓ, ensuring that the route is a proper unbroken cycle in the graph. Constraints (5) oblige vehicle k, $k = 1, \ldots, m$, to leave at most one time the depot. The choices of the arcs that are represented by $w_{i,j}^k$ is also restricted by constraints (6), that forbids subtours solutions [8]. Without these latter constraints, we can have a vehicle performing the path $(i_1, i_2, \ldots, i_t)$ with $0 \notin \{i_1, i_2, \ldots, i_t\}$. Constraints (7) state that every vehicle cannot carry more than its capacity limit. We note that constraints (3) and (4) imply $\sum_{i=0}^{n}\sum_{k=1}^{m} w_{j,i}^k = 1$, $j = 1, \ldots, n$, i.e., exactly one vehicle leaves client j. Constraints (5) and (4) imply $\sum_{i=1}^{n} w_{i,0}^k \leq 1$, $k = 1, \ldots, m$, i.e., vehicle k is obliged to return at most one time to the depot. Finally, we remark that this model requires using *at most* m vehicles, since for some k, we might have $w_{i,j}^k = 0$, $i, j = 1, \ldots, n$.

Example 1. *Suppose $m = 2$ vehicles with capacity limit $Q = 10$, which must collect the demands of $n = 4$ customers with demands $d_1 = 3, d_2 = 4, d_3 = 5, d_4 = 6$. These customers are illustrated in Figure 1a (the depot is denoted by "0"), along with their associated travel cost matrix in Figure 1b. A candidate solution, i.e., a set of routes satisfying constraints (3)-(7), to this problem is shown in Figure 1c; the total travel cost (the value of the objective function in Equation (1)) of this solution is 24.1. An optimal solution, i.e., a set of routes satisfying constraints (3)-(7) and with minimum cost among the candidate solutions, is provided in Figure 1d; its total travel cost is 17.8.*

2.2. The CVRPSD

The CVRPSD is a variation of the CVRP, which introduces stochastic demands in the CVRP, i.e., $d_i, i = 1, \ldots, n$, are now random variables, such that

(a) The four clients with their demands

(b) The travel cost matrix

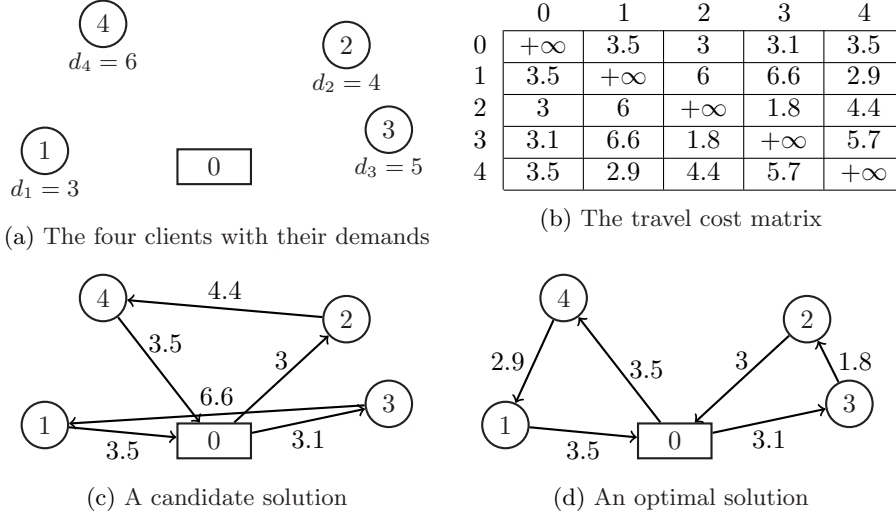(c) A candidate solution

(d) An optimal solution

Figure 1: A simple CVRP

$P(d_i \leq Q) = 1$ (these random variables are usually assumed to be independent). The CVRPSD is typically handled using the framework of stochastic programming, which models stochastic programs in two stages: an "a priori" solution is established in the first stage, and then in the second stage the realisations of the random variables – the actual demands in the case of the CVRPSD – are revealed and corrective actions are carried out if necessary on the first stage solution [28]. More precisely, the CVRPSD is either modelled as a so-called chance-constrained program [10] or as a stochastic program with recourse [7]; these two models are detailed in the next two sections.

2.2.1. The CVRPSD modelled by CCP

Chance constrained programming consists in finding a first stage solution for which the probability that the total demand on any route exceeds the capacity is constrained to be below a given threshold. Formally, a CCP formulation for the CVRPSD corresponds to the same optimisation problem described for the CVRP in Section 2.1 except that deterministic constraints represented by (7) are replaced by the following so called chance-constraints:

$$P \left(\sum_{i=0}^{n} d_i \sum_{j=0}^{n} w_{i,j}^k \leq Q \right) \geq 1 - \beta, \qquad k = 1, \ldots, m, \qquad (8)$$

where $1 - \beta$ is the minimum allowable probability that any route respects vehicle capacity and thus succeeds. Note that this model represents a so-called *individual* chance-constrained model, since the inequality must be satisfied for every k separately; see [45, 7] for more details.

This model does not consider the cost of corrective actions that may be necessary when the first stage solution is implemented. Indeed, when implementing this solution, it is unlikely yet possible that the vehicle capacity is exceeded, *i.e.*, route failures occur, when the actual demands are revealed and thus corrective actions may have to be carried out in the second stage.

2.2.2. The CVRPSD modelled by SPR

Stochastic programming with recourse deals explicitly with the possibility of a first stage solution failure, by incorporating into the objective of the problem the penalty cost of corrective, or *recourse*, actions such as allowing vehicles to return to the depot to unload. More specifically, in the SPR modelling of the CVRPSD, the *expected* penalty cost of the recourse actions happening in the second stage is considered and the problem is to find a set of routes which has the minimal expected cost defined as the cost of the first stage solution if no failures occur, plus the expected penalty cost of the recourse actions of the second stage. Formally, let $C_{\mathrm{E}}(R_k)$ denote the expected cost of a route R_k defined by

$$C_{\mathrm{E}}(R_k) = C(R_k) + C_{\mathrm{P}}(R_k), \tag{9}$$

with $C(R_k)$ the cost defined by (2) representing the cost of traveling along R_k if no recourse action is performed, and $C_{\mathrm{P}}(R_k)$ the expected penalty cost on R_k – $C_{\mathrm{P}}(R_k)$ may be defined in many different ways depending on the recourse policy used (see, *e.g.*, [12, 27, 39, 23]). Then, a SPR model for the CVRPSD consists in modifying the CVRP model presented in Section 2.1 as follows. The objective is to find a set of routes minimizing the sum of the expected costs of routes R_k, *i.e.*,

$$\min \sum_{k=1}^{m} C_{\mathrm{E}}(R_k), \tag{10}$$

subject to constraints (3)-(6) excluding constraints (5), which is replaced by

$$\sum_{j=1}^{n} w_{0,j}^k = 1, \qquad k = 1, \ldots, m, \tag{11}$$

that is *exactly m* vehicles must be used. Constraints (5) may be considered instead of (11), but then the problem becomes even more difficult to solve. In addition, note that the binary variables $w_{i,j}^k$ do not encode recourse actions: they represent only the initially planned solution routes, *i.e.*, the first stage solution.

2.3. Evidence theory

In this section, basic concepts as well as some more advanced notions of evidence theory [47], which are necessary in our study on the CVRPED, are recalled.

2.3.1. Basics of evidence theory

Let x be a variable taking its values in a finite domain $X = \{x_1, \ldots, x_K\}$. In this theory, uncertain knowledge about x may be represented by a *Mass Function* (MF) defined as a mapping $m^X : 2^X \to [0,1]$ such that $m^X(\emptyset) = 0$ and $\sum_{A \subseteq X} m^X(A) = 1$. The superscript ^{X} can be omitted when there is no risk of confusion. Each mass $m^X(A)$ represents the probability of knowing only that $x \in A$. Subsets $A \subseteq X$ such that $m^X(A) > 0$ are called the *focal sets* of m^X. To be consistent with the stochastic case terminology, a variable x whose true value is known in the form of a MF will be called an *evidential variable*.

Mass functions generalise both probabilistic and set valued representations of uncertainty since:

- a MF whose focal sets are singletons, *i.e.*, $m^X(A) > 0$ iff $|A| = 1$, corresponds to a probability mass function and is called a *Bayesian* MF;

- a MF having only one focal set, *i.e.*, $m^X(A) = 1$ for some $A \subseteq X$, corresponds to a set and is called a *categorical* MF.

Another special case of mass functions are those whose focal sets are nested, in which case they are called *consonant*.

Equivalent representations of a MF m^X are the *belief* and *plausibility* functions defined, respectively, as

$$Bel^X(x \in A) = \sum_{C \subseteq A} m^X(C), \quad \forall A \subseteq X,$$

$$Pl^X(x \in A) = \sum_{C \cap A \neq \emptyset} m^X(C), \quad \forall A \subseteq X.$$

The *degree of belief* $Bel^X(x \in A)$ can be interpreted as the probability that the evidence about x and represented by m^X, supports (implies) $x \in A$, whereas the *degree of plausibility* $Pl^X(x \in A)$ is the probability that the evidence is consistent with $x \in A$. For all $A \subseteq X$, we have $Bel^X(x \in A) \leq Pl^X(x \in A)$ and

$$Pl^X(x \in A) = 1 - Bel^X(x \in A^{\complement}), \tag{12}$$

where $A^{\complement}$ denotes the complement of A. Besides, if m^X is Bayesian, then $Bel^X(x \in A) = Pl^X(x \in A)$, for all $A \subseteq X$, and this function is a probability measure. If m^X is categorical, then $Bel^X(x \in A) \in \{0,1\}$ and $Pl^X(x \in A) \in \{0,1\}$, for all $A \subseteq X$, and the plausibility function restricted to the singletons corresponds to the indicator function of the set associated to m^X. If m^X is consonant, then its associated plausibility function is a possibility measure [54]: it verifies $Pl^X(x \in A \cup B) = \max(Pl^X(x \in A), Pl^X(x \in B))$, for all $A, B \subseteq X$.

Given a MF m^X and a function $h : X \to \mathbb{R}^+$, it is possible to compute the *lower expected value* $E_*(h, m^X)$ and *upper expected value* $E^*(h, m^X)$ of h

relative to m^X defined, respectively, as [16]

$$E_*(h, m^X) = \sum_{A \subseteq X} m^X(A) \min_{x \in A} h(x), \qquad (13)$$

$$E^*(h, m^X) = \sum_{A \subseteq X} m^X(A) \max_{x \in A} h(x). \qquad (14)$$

If m^X is Bayesian, then $E_*(h, m^X)$ and $E^*(h, m^X)$ reduce to the classical (probabilistic) expected value of h relative to the probability mass function m^X.

2.3.2. Comparisons of belief functions

The informative content of two set-valued pieces of information $x \in A$ and $x \in B$, $A, B \subseteq X$, about x is naturally compared by saying that $x \in A$ is more informative than $x \in B$ if $A \subset B$. An extension of this to compare the informative content of mass functions in terms of specificity is the notion of specialisation [24]: a MF m_1^X defined on X is said to be at least as informative (or *specific*) as another MF m_2^X defined on X, which is denoted by $m_1^X \sqsubseteq m_2^X$, if and only if there exists a non-negative square matrix $S = [S(A, B)]$, $A, B \in 2^X$, verifying

$$\sum_{A \subseteq X} S(A, B) = 1, \quad \forall B \subseteq X, \qquad (15)$$

$$S(A, B) > 0 \Rightarrow A \subseteq B, \quad A, B \subseteq X, \qquad (16)$$

$$m_1^X(A) = \sum_{B \subseteq X} S(A, B) m_2^X(B), \quad \forall A \subseteq X. \qquad (17)$$

The term $S(A, B)$ may be seen as the proportion of the mass $m_2^X(B)$ which "flows down" to A. Let us also recall that we have [24]

$$m_1^X \sqsubseteq m_2^X \Rightarrow [Bel_1^X(x \in A), Pl_1^X(x \in A)] \subseteq [Bel_2^X(x \in A), Pl_2^X(x \in A)], \forall A \subseteq X. \qquad (18)$$

Assume now that an ordering has been defined among the elements of X. By convention, assume that $x_1 < \ldots < x_K$. Let $A_{\underline{a}, \overline{a}}$ denote the subset $\{x_{\underline{a}}, \ldots, x_{\overline{a}}\}$, for $1 \leq \underline{a} \leq \overline{a} \leq K$ and let $\mathcal{I}$ denote the set of intervals of X: $\mathcal{I} = \{A_{\underline{a}, \overline{a}}, 1 \leq \underline{a} \leq \overline{a} \leq K\}$. Deciding whether an interval $A_{\underline{a}, \overline{a}}$, *i.e.* an interval-valued piece of information $x \in A_{\underline{a}, \overline{a}}$ about x, is smaller or equal to another interval $B_{\underline{b}, \overline{b}}$ can be done in several ways, and in particular the so-called lattice ordering denoted $\leq_{lo}$ is defined as [20]: $A_{\underline{a}, \overline{a}} \leq_{lo} B_{\underline{b}, \overline{b}}$ if $\underline{a} \leq \underline{b}$ and $\overline{a} \leq \overline{b}$. This ordering can be extended to arbitrary subsets A and B of X as follows: $A \leq_{lo} B$ if $\underline{a} \leq \underline{b}$ and $\overline{a} \leq \overline{b}$, where $\underline{a}$ and $\underline{b}$ (resp. $\overline{a}$ and $\overline{b}$) denote the indices of the lowest (resp. highest) values in A and B. More generally, following the extension above of inclusion between sets to mass functions, the ordering $\leq_{lo}$ of subsets can be extended to compare mass functions in terms of ranking as follows: a MF m_1^X is said to be at least as small as another MF m_2^X, which is

denoted by $m_1^X \preceq m_2^X$, if and only if there exists a non-negative square matrix $R = [R(A,B)]$, $A, B \in 2^X$, verifying

$$\sum_{A \subseteq X} R(A,B) = 1, \quad \forall B \subseteq X, \tag{19}$$

$$R(A,B) > 0 \Rightarrow A \leq_{lo} B, \quad A, B \subseteq X, \tag{20}$$

$$m_1^X(A) = \sum_{B \subseteq X} R(A,B) m_2^X(B), \quad \forall A \subseteq X. \tag{21}$$

In other words, the mass $m_2^X(B)$ can be shared among smaller (according to $\leq_{lo}$) subsets than B. We note that extensions of interval rankings to belief functions were already proposed in [17], but in the context of belief functions on the real line and the ranking $\leq_{lo}$ was not considered. To our knowledge, the definition of $\preceq$ appears thus to be new; it will be particularly useful in conjunction with Proposition 1, which is somewhat of a counterpart to Eq. (18) for $\preceq$, to exhibit a property of the BCP model for the CVPRED:

Proposition 1. *For any Q, $1 \leq Q \leq K$, we have*

$$m_1^X \preceq m_2^X \Rightarrow \begin{cases} Bel_1^X(x \in A_{1,Q}) \geq Bel_2^X(x \in A_{1,Q}), \\ Pl_1^X(x \in A_{1,Q}) \geq Pl_2^X(x \in A_{1,Q}). \end{cases} \tag{22}$$

The converse does not hold, i.e., *the implication in* (22) *is strict.*

Proof. See Appendix A. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Proposition 1 basically says that if a MF m_1^X is at least as small as a MF m_2^X, then the belief, according to the piece of evidence m_1^X, that the value of x is smaller or equal than a value x_Q is at least as great as the belief of the same event according to the piece of evidence m_2^X (and the same goes for the plausibility), as may be expected from the meaning of $\preceq$.

To sum up this section, $m_1^X \sqsubseteq m_2^X$ basically means that m_2^X represents a less precise piece of uncertain knowledge about x than m_1^X, whereas $m_1^X \preceq m_2^X$ means that m_2^X represents a piece of uncertain knowledge telling that x takes a higher value than what m_1^X tells.

2.3.3. Uncertainty propagation

Let $x^1, \dots, x^N$ be N variables defined on the finite domains $X_1, \dots, X_N$, respectively. A MF $m^{X_1 \times \dots \times X_N}$ defined on the Cartesian product $X_1 \times \dots \times X_N$ represent joint knowledge about the values of these variables.

Similarly as in probability theory, one can obtain joint knowledge about a subset of the evidential variables $x^1, \dots, x^N$ by *marginalising* MF $m^{X_1 \times \dots \times X_N}$ on the domains of these variables. For instance, and without lack of generality, the marginalisation of $m^{X_1 \times \dots \times X_N}$ on $X_1 \times X_2$ is the MF $m^{X_1 \times \dots \times X_N \downarrow X_1 \times X_2}$ on $X_1 \times X_2$ defined as, $\forall A \subseteq X_1 \times X_2$,

$$m^{X_1 \times \dots \times X_N \downarrow X_1 \times X_2}(A) = \sum_{\{B \subseteq X_1 \times \dots \times X_N, B^{\downarrow X_1 \times X_2} = A\}} m^{X_1 \times \dots \times X_N}(B), \tag{23}$$

10

where $B^{\downarrow X_1 \times X_2}$ denotes the projection of B onto $X_1 \times X_2$.

If MF $m^{X_1 \times \cdots \times X_N}$ satisfies, for all $A \subseteq X_1 \times \cdots \times X_N$,

$$m^{X_1 \times \cdots \times X_N}(A) = \begin{cases} \prod_{i=1}^{N} m^{X_1 \times \cdots \times X_N \downarrow X_i}(A^{\downarrow X_i}) & \text{if } A = \times_{i=1}^{N} A^{\downarrow X_i}, \\ 0 & \text{otherwise,} \end{cases} \quad (24)$$

then variables $x^1, \ldots, x^N$ are said to be *evidentially independent* (or independent for short) [47]. In practice, this happens when joint knowledge about variables $x^1, \ldots, x^N$, is built from marginal knowledge m^{X_i}, $i = 1, \ldots, N$, on each of these variables, supplied by sources assumed to be independent [43], as illustrated by Example 2 (other reasons for this to happen can also be found in [14, 13]).

Example 2. *Let* $X_1 = \{x_1^1, x_2^1, x_3^1\}$ *and* $X_2 = \{x_1^2, x_2^2\}$. *Furthermore, assume two sources providing the pieces of evidence* m^{X_1} *and* m^{X_2}, *respectively, about* x_1 *and* x_2, *defined as* $m^{X_1}(\{x_1^1, x_2^1\}) = 0.8$, $m^{X_1}(\{x_2^1, x_3^1\}) = 0.2$ *and* $m^{X_2}(\{x_1^2\}) = 0.7$, $m^{X_2}(X_2) = 0.3$. *Assuming that the sources are independent, we obtain*

$$\begin{aligned} m^{X_1 \times X_2}(\{x_1^1, x_2^1\} \times \{x_1^2\}) &:= m^{X_1}(\{x_1^1, x_2^1\}) \cdot m^{X_2}(\{x_1^2\}) = 0.56, \\ m^{X_1 \times X_2}(\{x_1^1, x_2^1\} \times X_2) &:= m^{X_1}(\{x_1^1, x_2^1\}) \cdot m^{X_2}(X_2) = 0.24, \\ m^{X_1 \times X_2}(\{x_2^1, x_3^1\} \times \{x_1^2\}) &:= m^{X_1}(\{x_2^1, x_3^1\}) \cdot m^{X_2}(\{x_1^2\}) = 0.14, \\ m^{X_1 \times X_2}(\{x_2^1, x_3^1\} \times X_2) &:= m^{X_1}(\{x_2^1, x_3^1\}) \cdot m^{X_2}(X_2) = 0.06. \end{aligned}$$

$m^{X_1 \times X_2}$ *clearly satisfies* (24).

Furthermore, let y be a variable with finite domain Y, such that $y = f(x^1, \ldots, x^N)$ for some mapping $f : X_1 \times \cdots \times X_N \to Y$. As shown in [25], uncertain knowledge $m^{X_1 \times \cdots \times X_N}$ about variables $x^1, \ldots, x^N$, induces MF m^Y about the value of y defined as

$$m^Y(B) = \sum_{f(A)=B} m^{X_1 \times \cdots \times X_N}(A), \quad \forall B \subseteq Y, \quad (25)$$

with $f(A) = \{f(x_{k^1}^1, \ldots, x_{k^N}^N) | (x_{k^1}^1, \ldots, x_{k^N}^N) \in A\}$ for all $A \subseteq X_1 \times \cdots \times X_N$.

3. Modelling the CVRPED

This section formalises and studies the CVRPED, which is an integer linear program involving uncertainty represented by belief functions. We obtain this problem when customer demands in the CVRP are no longer deterministic or random, but evidential, *i.e.*, the variables d_i, $i = 1, ..., n$, are evidential. Following what has been done for the case of linear programs involving evidential uncertainty [40], we may extend stochastic programming approaches to this integer linear program embedding belief functions: the CCP modelling of the CVRPSD is generalised into a BCP modelling of the CVRPED in Section 3.1, and the recourse modelling of the CVRPSD is generalised into a recourse modelling of the CVRPED in Section 3.2.

Note that, to simplify the exposition, we assume actual customer demands to be positive *integers*, hence the value of the demand of any customer belongs to the set $\Theta = \{1, 2, \ldots, Q\}$. In addition, since the CVPRED involves n evidential variables d_i, $i = 1, \ldots, n$, with respective domains $\Theta_i := \Theta$, $i = 1, \ldots, n$, then formally this means that knowledge about customer demands in this problem is represented by a MF m^{Θ^n} on $\Theta^n := \times_{i=1}^{n} \Theta_i$. In practical situations, it may be the case that only marginal knowledge in the form of a MF m^{Θ_i} may be available about the individual demand of each customer i, $i = 1, \ldots, n$. In such case, as explained in Section 2.3.3, m^{Θ^n} can be derived by assuming that these pieces of knowledge about individual customer demands have been provided by independent sources. In other words, if necessary and justified, evidential variables d_i, $i = 1, \ldots, n$, may be assumed to be independent, similarly as it may be done in the stochastic case. However, let us underline that the BCP and recourse modellings of the CVRPED proposed in the next two sections, are general in that they do not rely on such independence assumption, *i.e.*, they do not need m^{Θ^n} to satisfy a property of the form (24).

3.1. The CVRPED modelled by BCP

A generalisation of the CCP modelling of the CVRPSD to the case of evidential demands is proposed in this section. The model is provided in Section 3.1.1. Important particular cases of this model are discussed in Section 3.1.2. Influences of model parameters and of customer demand ranking on the optimal solution cost, are studied in Sections 3.1.3 and 3.1.4, respectively.

3.1.1. Formalisation

A BCP modelling of the CVRPED amounts to keeping the same optimisation problem described for the CVRP in Section 2.1 except that capacity constraints (7) are replaced by the following *belief*-constraints:

$$Bel\left(\sum_{i=1}^{n} d_i \sum_{j=0}^{n} w_{i,j}^k \leq Q\right) \geq 1 - \underline{\beta}, \qquad k = 1, \ldots, m, \qquad (26)$$

$$Pl\left(\sum_{i=1}^{n} d_i \sum_{j=0}^{n} w_{i,j}^k \leq Q\right) \geq 1 - \overline{\beta}, \qquad k = 1, \ldots, m, \qquad (27)$$

with $\underline{\beta} \geq \overline{\beta}$ and where $1 - \underline{\beta}$ (resp. $1 - \overline{\beta}$) is the minimum allowable degree of belief (resp. plausibility) that a vehicle capacity is respected on any route.

Remark 1. *From (12), constraints (27) are equivalent to*

$$Bel\left(\sum_{i=1}^{n} d_i \sum_{j=0}^{n} w_{i,j}^k > Q\right) \leq \overline{\beta}, \qquad k = 1, \ldots, m. \qquad (28)$$

Hence, constraints (26) and (27) amount to requiring that for any route there is a lot (at least $1 - \underline{\beta}$) of support (belief) of respecting vehicle capacity and not a lot (at most $\overline{\beta}$) of support of violating vehicle capacity.
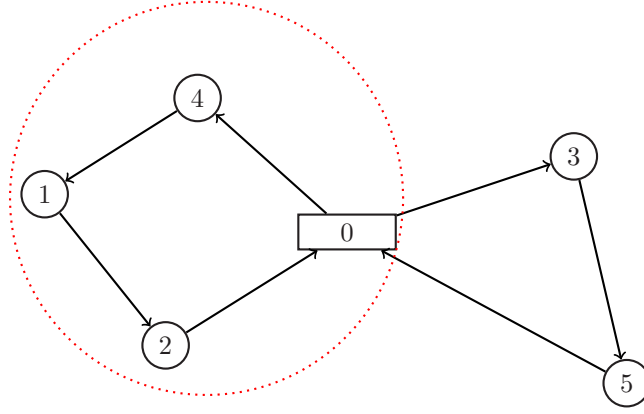
Figure 2: Evaluation of constraints (26) and (27) on the route $(0, 4, 1, 2, 0)$ (route circled in red).

Note that in order to evaluate the belief-constraints (26) and (27), the total demand on every route must be determined by summing all customer demands on that route. Suppose a route R having N clients, then the sum of customer demands on R is obtained using (25), where f is the addition of integers and where $m^{X_1 \times \cdots \times X_N}$ is the marginalisation of m^{Θ^n} on the domains of the evidential variables $dr_1, \ldots, dr_N$ associated with the N clients on the route, with X_i the domain of the evidential variable dr_i associated with the i-th client on R. The computation of the total demand on a route as well as the evaluation of constraints (26) and (27) for that route are illustrated by Example 3.

Example 3. *Suppose that $\underline{\beta} = 0.1$ and $\overline{\beta} = 0.05$ and that we have $n = 5$ customers and $m = 2$ vehicles with capacity limit $Q = 15$. Moreover, suppose knowledge about customer demands is represented by MF m^{Θ^n} defined on $\Theta^n = \Theta^5 = \Theta_1 \times \Theta_2 \times \Theta_3 \times \Theta_4 \times \Theta_5$ by:*

$$
\begin{aligned}
m^{\Theta^5}(\{(2,3,8,4,5), (3,5,6,7,4), (3,4,7,6,2)\}) &= 0.5, \\
m^{\Theta^5}(\{(5,5,6,4,7), (7,6,5,3,4)\}) &= 0.3, \\
m^{\Theta^5}(\{(4,6,7,4,6), (5,5,6,5,7)\}) &= 0.2. \quad (29)
\end{aligned}
$$

Consider the two routes represented in Figure 2. Let us compute the sum of the customer demands on the route $(0, 4, 1, 2, 0)$, i.e., the route that collects the demand of customer 4, then the demand of customer 1 and finally the demand of customer 2. Call this route R. On this route, there are $N = 3$ clients. The first client is client 4, hence according to the above notation, we have $dr_1 = d_4$ and $X_1 = \Theta_4$. Similarly, we have

$$
\begin{aligned}
dr_2 &= d_1, X_2 = \Theta_1, \\
dr_3 &= d_2, X_3 = \Theta_2.
\end{aligned}
$$

13

The marginalisation of m^{Θ^5} on $X_1 \times X_2 \times X_3$ is the MF $m^{\Theta^5 \downarrow X_1 \times X_2 \times X_3}$ defined as:

$$
\begin{aligned}
m^{\Theta^5 \downarrow X_1 \times X_2 \times X_3}(\{(4,2,3),(7,3,5),(6,3,4)\}) &= 0.5, \\
m^{\Theta^5 \downarrow X_1 \times X_2 \times X_3}(\{(4,5,5),(3,7,6)\}) &= 0.3, \\
m^{\Theta^5 \downarrow X_1 \times X_2 \times X_3}(\{(4,4,6),(5,5,5)\}) &= 0.2.
\end{aligned}
\tag{30}
$$

Now given $m^{\Theta^5 \downarrow X_1 \times X_2 \times X_3}$ and using Equation (25) such that f is the addition of integers, uncertainty on the sum of client demands on route R is represented by a MF denoted $m_{\sum}^{\Theta_R}$ and defined on the domain $\Theta_R := \{1,2,\ldots,N \cdot Q\} = \{1,\ldots,45\}$ by:

$$
\begin{aligned}
m_{\sum}^{\Theta_R}(\{9,15,13\}) &= 0.5, \\
m_{\sum}^{\Theta_R}(\{14,16\}) &= 0.3, \\
m_{\sum}^{\Theta_R}(\{14,15\}) &= 0.2.
\end{aligned}
\tag{31}
$$

The belief and plausibility that the sum of customer demands on R is smaller or equal than the vehicle capacity Q are then respectively:

$$
\begin{aligned}
Bel(d_4 + d_1 + d_2 \leq 15) &= m_{\sum}^{\Theta_R}(\{9,15,13\}) + m_{\sum}^{\Theta_R}(\{14,15\}) \\
&= 0.7, \\
Pl(d_4 + d_1 + d_2 \leq 15) &= 1.
\end{aligned}
$$

Hence, we have

$$
\begin{aligned}
Bel(dr_1 + dr_2 + dr_3 \leq Q) &< 1 - \underline{\beta} = 0.9, \\
Pl(dr_1 + dr_2 + dr_3 \leq Q) &> 1 - \overline{\beta} = 0.95.
\end{aligned}
$$

In other words, constraint (27) is satisfied on R but constraint (26) is not, and thus any set of routes containing this route, such as the one shown in Figure 2, is not a candidate solution.

Suppose further that the number of focal sets of MF $m^{X_1 \times \cdots \times X_N}$ is at most c, then the worst case complexity of evaluating each of the belief constraints (26) and (27) on this route is $\mathcal{O}(N \cdot Q^N \cdot c)$. This latter complexity emerges from the following: the Q^N factor is the maximal number of elements of a focal set of MF $m^{X_1 \times \cdots \times X_N}$. As we have N clients on a route, then for each element of a focal set of MF $m^{X_1 \times \cdots \times X_N}$, the addition of N integers must be performed, this explains $N \cdot Q^N$. The last factor in the complexity which is c, is related to performing the product $N \cdot Q^N$ for the c focal sets of MF $m^{X_1 \times \cdots \times X_N}$. Nonetheless, in a particular case, the worst case complexity drops down significantly:

Remark 2. *When the focal sets of $m^{X_1 \times \cdots \times X_N}$ are all Cartesian products of N intervals, i.e., for all $A \subseteq X_1 \times \cdots \times X_N$ such that $m^{X_1 \times \cdots \times X_N}(A) > 0$, we have $A = A^{\downarrow X_1} \times \cdots \times A^{\downarrow X_N}$ with, for $i = 1,\ldots,N$, $A^{\downarrow X_i} = [\![\underline{A_i}; \overline{A_i}]\!]$ for some*

integers $\underline{A_i}, \overline{A_i} \in X_i$ such that $\underline{A_i} \le \overline{A_i}$, the worst case complexity is $\mathcal{O}(N \cdot c)$. This complexity to evaluate constraint (26) for route R comes from the fact that (with dr_i the evidential variable associated with the i-th client on R):

$$
\begin{aligned}
Bel(\sum_{i=1}^{N} dr_i \le Q) &= \sum \{m^{X_1 \times \cdots \times X_N}(A)|A : \max_{a \in A} f(a) \le Q\} \\
&= \sum \{m^{X_1 \times \cdots \times X_N}(A)|A : \max_{(a_1, \cdots, a_N) \in A} \sum_{i=1}^{N} a_i \le Q\} \quad (32) \\
&= \sum \{m^{X_1 \times \cdots \times X_N}(A)|A : \sum_{i=1}^{N} \overline{A_i} \le Q\},
\end{aligned}
$$

that is, at worst for each of the c focal sets of $m^{X_1 \times \cdots \times X_N}$, the addition of N integers needs to be performed. The complexity to evaluate constraint (27) is the same since we have

$$
\begin{aligned}
Pl(\sum_{i=1}^{N} dr_i \le Q) &= \sum \{m^{X_1 \times \cdots \times X_N}(A)|A : \min_{(a_1, \cdots, a_N) \in A} \sum_{i=1}^{N} a_i \le Q\} \quad (33) \\
&= \sum \{m^{X_1 \times \cdots \times X_N}(A)|A : \sum_{i=1}^{N} \underline{A_i} \le Q\}.
\end{aligned}
$$

Remark 3. *From (32), it is clear that the complexity to evaluate constraint (26) for a given route R, depends on the complexity of finding for each focal set A of $m^{X_1 \times \cdots \times X_N}$, the element $(a_1, \cdots, a_N) \in A$ that maximises $\sum_{i=1}^{N} a_i$. Suppose this latter complexity is at worst $\mathcal{O}(M)$, $1 \le M \le Q^N$, for each focal set. Then, the worst case complexity to evaluate constraint (26) for a route is $\mathcal{O}(N \cdot M \cdot c)$. Remark 2 provides a case, i.e., a particular shape for the focal sets of $m^{X_1 \times \cdots \times X_N}$, such that $M = 1$. We note that other, more refined yet still leading to tractable values for M, shapes for these focal sets may be considered. For instance, borrowing from robust optimisation [6], suppose that each focal set A of $m^{X_1 \times \cdots \times X_N}$ can be written as*

$$
A = \{(a_1, \cdots, a_N)|a_i \ge \underline{A_i}, a_i \ge \overline{A_i}, \sum_{i=1}^{N} \frac{a_i - \underline{A_i}}{\underline{A_i}} \le \Gamma\}, \quad (34)
$$

for some lower bounds $\underline{A_i}$ and upper bounds $\overline{A_i}$ $i = 1, \ldots, N$, and some un-certainty budget Γ [6]; budget Γ in (34) limits the sum of the deviations from the minimum demands $\underline{A_i}, i = 1, \ldots, N$. Then, maximizing $\sum_{i=1}^{N} a_i$ for each focal set A is done over a more difficult, yet still manageable, shape than in Remark 2. Obviously, similar comments can be made about the complexity of constraint (27).

3.1.2. Particular cases of the BCP modelling of the CVRPED

It is interesting to remark that depending on the values chosen for β and $\overline{\beta}$ as well as the nature of the evidential demands d_i, $i = 1,...,n$, the BCP modelling of the CVRPED may degenerate into simpler or well-known optimisation problems.

In particular, if m^{Θ^n} is Bayesian, *i.e.*, we are dealing really with a CVRPSD, then we have, for $k = 1,\dots,m$,

$$Bel\left(\sum_{i=1}^{n} d_i \sum_{j=0}^{n} w_{i,j}^k \leq Q\right) = Pl\left(\sum_{i=1}^{n} d_i \sum_{j=0}^{n} w_{i,j}^k \leq Q\right), \qquad (35)$$

and the BCP modelling of the CVRPED can be converted into an equivalent optimisation problem, which is the CCP modelling of this CVRPSD, with β in (8) set to $\overline{\beta}$.

In contrast, if m^{Θ^n} is categorical and its only focal set is the Cartesian product of n intervals, *i.e.*, we are dealing with a CVRP where each customer demand d_i is only known in the form of an interval $[\![\underline{d_i}; \overline{d_i}]\!]$, then the total demand on any given route is also an interval (its endpoints are obtained by summing the endpoints of the interval demands of the customers on the route) and thus for any $k = 1,\dots,m$, $Bel^{\Theta}(\sum_{i=0}^{n} d_i \sum_{j=0}^{n} w_{i,j}^k \leq Q)$ either equals 1 or equals 0, with the former occurring iff $\sum_{i=0}^{n} \overline{d_i} \sum_{j=0}^{n} w_{i,j}^k \leq Q$, and $Pl^{\Theta}(\sum_{i=0}^{n} d_i \sum_{j=0}^{n} w_{i,j}^k \leq Q)$ either equals 1 or equals 0, with the former occurring iff $\sum_{i=0}^{n} \underline{d_i} \sum_{j=0}^{n} w_{i,j}^k \leq Q$. Then, since $\sum_{i=0}^{n} \overline{d_i} \sum_{j=0}^{n} w_{i,j}^k \leq Q \Rightarrow \sum_{i=0}^{n} \underline{d_i} \sum_{j=0}^{n} w_{i,j}^k \leq Q$, the belief-constraints (26) and (27) reduce when $\underline{\beta} < 1$ to the following constraints

$$\sum_{i=0}^{n} \overline{d_i} \sum_{j=0}^{n} w_{i,j}^k \leq Q, \qquad k = 1,\dots,m. \qquad (36)$$

In other words, in the case of interval demands, the BCP modelling amounts to searching the solution which minimises the overall cost of servicing the customers (1) under constraints (36), *i.e.*, assuming the maximum (worst) possible customer demands, and thus it corresponds to the minimax optimisation procedures encountered in robust optimisation [49].

If m^{Θ^n} is consonant, then we are dealing with a CVRP where uncertainty on customer demand is really of a possibilistic nature, and the BCP modelling may then be connected to fuzzy-based approaches, that is approaches where uncertainty on customer demands is represented by fuzzy sets such as in [50]. In addition, let us remark that if only marginal knowledge in the form of a consonant MF m^{Θ_i} having interval focal sets is available about the individual demand of each customer i, $i = 1,\dots,n$, then, as explained in Section 2.3.3, m^{Θ^n} can be obtained by assuming (if justified) independence of the demands, in which case it will yield a tractable situation thanks to Remark 2, whose conditions are then satisfied (the focal sets of m^{Θ^n} being in this case Cartesian products of intervals). However, m^{Θ^n} may also be derived from these pieces of marginal knowledge by making other assumptions about the demand dependence and in

particular by assuming that they are *non-interactive* [25, 5] - a more classical independence assumption in the fuzzy setting - in which case the focal sets of m^{Θ^n} will also be Cartesian products of intervals but they will also be nested (m^{Θ^n} will then be consonant).

If $\underline{\beta} = \overline{\beta}$, then constraints (27) can be dropped, that is, only constraints (26) need to be evaluated (if constraints (26) are satisfied then constraints (27) are necessarily satisfied due to the relation between the belief and plausibility functions). As a matter of fact, the BCP approach originally introduced in [40] is of this form (no constraint based on Pl is considered). Most importantly, when $\underline{\beta} = \overline{\beta}$ and the evidential variables d_i, $i = 1 \ldots, n$ are independent, the BCP modelling of the CVRPED can be converted into an equivalent optimisation problem, which is the CCP modelling (with β in (8) set to $\underline{\beta}$) of a CVRPSD where customer demands are represented by independent stochastic variables denoted $d_{\overline{i}}$, $i = 1, \ldots, n$, with associated probability mass function $p_{\overline{i}}$ obtained from $m^{\Theta_i} := m^{\Theta^n \downarrow \Theta_i}$ as follows: for each focal set $A \subseteq \Theta_i$ of m^{Θ_i}, the mass $m^{\Theta_i}(A)$ is transferred to the element $\theta = \max(A)$. Indeed, with such a definition of $p_{\overline{i}}$, it is easy to show that we have, for $k = 1, \ldots, m$,

$$Bel\left(\sum_{i=1}^{n} d_i \sum_{j=0}^{n} w_{i,j}^k \leq Q\right) = P\left(\sum_{i=1}^{n} d_{\overline{i}} \sum_{j=0}^{n} w_{i,j}^k \leq Q\right). \qquad (37)$$

Let us eventually remark that the case $\underline{\beta} = 1 > \overline{\beta}$ is the converse of the case $\underline{\beta} = \overline{\beta}$ in the sense that constraints (26) can be dropped (as they are necessarily satisfied) and only constraints (27) need then to be evaluated. Moreover, in this case, if the evidential variables d_i, $i = 1 \ldots, n$ are independent, the BCP modelling of the CVRPED can be converted into an equivalent optimisation problem, which is the CCP modelling (with β in (8) set to $\overline{\beta}$) of a CVRPSD where customer demands are represented by independent stochastic variables denoted $d_{\underline{i}}$, $i = 1, \ldots, n$, with associated probability mass function $p_{\underline{i}}$ obtained from m^{Θ_i} as follows: for each focal set $A \subseteq \Theta_i$ of m^{Θ_i}, the mass $m^{\Theta_i}(A)$ is transferred to the element $\theta = \min(A)$. Indeed, with such a definition of $p_{\underline{i}}$, it is easy to show that we have, for $k = 1, \ldots, m$,

$$Pl\left(\sum_{i=1}^{n} d_i \sum_{j=0}^{n} w_{i,j}^k \leq Q\right) = P\left(\sum_{i=1}^{n} d_{\underline{i}} \sum_{j=0}^{n} w_{i,j}^k \leq Q\right). \qquad (38)$$

3.1.3. Influence of $\underline{\beta}$, $\overline{\beta}$ and Q on the CVRPED-BCP optimal solution cost

In this section, we study the influence of the parameters $\underline{\beta}$, $\overline{\beta}$ and Q, on the optimal solution cost of the CVRPED modelled via BCP (in the remainder of this article, we will simply say CVRPED-BCP instead of CVRPED modelled via BCP).

The following propositions hold:

Proposition 2. *The optimal solution cost is non increasing in Q.*

Proof. See Appendix B. $\qquad\square$

Proposition 3. *The optimal solution cost is non increasing in $\underline{\beta}$.*

Proof. See Appendix C. $\qquad\square$

Proposition 4. *The optimal solution cost is non increasing in $\overline{\beta}$.*

Proof. The proof is similar to that of Proposition 3. $\qquad\square$

Informally, Propositions 2–4 state that if the decision maker is willing to buy vehicles with a higher capacity or to have vehicle capacity exceeded on any route more often, then he will obtain at least as good (at most as costly) solutions.

3.1.4. Influence of customer demand ranking on the CVRPED-BCP optimal solution cost

In this section, we study the influence on the CVRPED-BCP optimal solution cost, of considering uncertain knowledge representing a more pessimistic estimation of the demand of each customer than currently assumed, that is uncertain knowledge telling that the demand of each customer is higher than currently believed.

Specifically, let m^{Θ^n} and $m_+^{\Theta^n}$ be two MF representing uncertain knowledge about customer demands, such that evidential variables d_i, $i = 1, \ldots, n$, are independent according to both these mass functions. Furthermore, let $m^{\Theta_i} := m^{\Theta^n \downarrow \Theta_i}$ and $m_+^{\Theta_i} := m_+^{\Theta^n \downarrow \Theta_i}$, $i = 1, \ldots, n$. Denote by $\hat{C}_{Q,\underline{\beta},\overline{\beta}}$ and $\hat{C}_{Q,\underline{\beta},\overline{\beta}}^+$ the costs of optimal solutions to the CVRPED-BCP when customer demands are known in the form of m^{Θ^n} and $m_+^{\Theta^n}$, respectively, for some $\underline{\beta}$, $\overline{\beta}$ and Q.

The following proposition holds:

Proposition 5. $m^{\Theta_i} \preceq m_+^{\Theta_i}, i = 1, \ldots, n \Rightarrow \hat{C}_{Q,\underline{\beta},\overline{\beta}} \leq \hat{C}_{Q,\underline{\beta},\overline{\beta}}^+$.

Proof. See Appendix D. $\qquad\square$

Informally, Proposition 5 shows that the more pessimistic knowledge is about customer demands, the greater the cost of the optimal solution.

An immediate consequence of this result is:

Corollary 1. *Assume that the focal sets of MF m^{Θ_i} and $m_+^{\Theta_i}$, $i = 1, \ldots, n$, are all intervals and that $m_+^{\Theta_i}$ can be obtained from m^{Θ_i}, $i = 1, \ldots, n$, as follows: for each interval $A = [\![\underline{A}; \overline{A}]\!]$ such that $m^{\Theta_i}(A) > 0$, the mass $m^{\Theta_i}(A)$ is transferred to the interval $A^+ = [\![\underline{A}; \overline{A} + a^+]\!]$, with $a^+ \in [\![0; Q - \overline{A}]\!]$. Then, $\hat{C}_{Q,\underline{\beta},\overline{\beta}} \leq \hat{C}_{Q,\underline{\beta},\overline{\beta}}^+$.*

Another immediate consequence is:

Corollary 2. *Assume that the focal sets of MF m^{Θ_i} and $m_+^{\Theta_i}$, $i = 1, \ldots, n$, are all intervals and that m^{Θ_i} can be obtained from $m_+^{\Theta_i}$, $i = 1, \ldots, n$, as follows: for each interval $A^+ = [\![\underline{A^+}; \overline{A^+}]\!]$ such that $m_+^{\Theta_i}(A^+) > 0$, the mass $m_+^{\Theta_i}(A^+)$ is transferred to the interval $A = [\![\underline{A^+} - a; \overline{A^+}]\!]$, with $a \in [\![0; \underline{A^+} - 1]\!]$. Then, $\hat{C}_{Q,\underline{\beta},\overline{\beta}} \leq \hat{C}_{Q,\underline{\beta},\overline{\beta}}^+$.*

Remark 4. *In both Corollaries 1 and 2, it is easy to show that $m^{\Theta_i} \preceq m_+^{\Theta_i}, i = 1, \ldots, n$, which is the reason why these corollaries hold. Note that for Corollary 1, we can also easily show that $m^{\Theta_i} \sqsubseteq m_+^{\Theta_i}, i = 1, \ldots, n$, whereas for Corollary 2, we have $m_+^{\Theta_i} \sqsubseteq m^{\Theta_i}, i = 1, \ldots, n$. This shows that the CVRPED-BCP optimal solution cost will not necessarily be higher if knowledge about customer demand is less specific. As will be seen in the next section, a different conclusion is reached for the recourse model, and specifically a counterpart to Proposition 5, based on $\sqsubseteq$ rather than $\preceq$, holds.*

3.2. The CVRPED modelled by a recourse approach

A recourse approach for the CVRPED is proposed in this section. The general model, extending the one recalled for the CVRPSD in Section 2.2.2, is presented in Section 3.2.1. Then, in Section 3.2.2, we detail how uncertainty on recourse actions is obtained in this model and in Section 3.2.3 we provide a method to compute efficiently this latter uncertainty in an important particular case. Similarly to what has been done for the BCP model, we discuss particular cases of our general model in Section 3.2.4 and study the influence of customer demands specificity on the optimal solution cost in Section 3.2.5.

3.2.1. Formalisation

The CVRPED may be addressed using an extension of the other main approach to modelling stochastic programs, that is the recourse approach. We propose to extend the recourse approach to the CVRPED, for the following policy and assumptions studied for the stochastic case in [12, 27, 39, 23]. Each actual customer demand cannot exceed the vehicle capacity. In addition, when a vehicle arrives at a customer on its planned route, it is loaded with the actual customer demand up to its remaining capacity. If this remaining capacity is sufficient to pick-up the entire customer demand, then the vehicle continues its planned route. However, if it is not sufficient, *i.e.*, there is a failure, then the vehicle returns to the depot, is emptied, goes back to the client to pick-up the remaining customer demand and continues its originally planned route.

Consider a given route R containing N customers and, without lack of generality, that the i-th customer on R is customer i. According to the above setting, a failure cannot occur at the first customer on R. However, it can occur at any other customer on R, and there may even be failure at multiple customers on R (at worst, if the actual demand of each customer is equal to the capacity of the vehicle, failure occurs at each customer except the first one).

Formally, let us introduce a binary variable r_i that equals 1 if failure occurs at the i-th customer on R and 0 otherwise (by problem definition $r_1 = 0$). Then, the possible failure situations that may occur along R may be represented by the vectors $(r_2, r_3, \ldots, r_N) \in \{0,1\}^{N-1}$. To simplify the exposition, we define the set Ω as the space of binary vectors representing the possible failure situations along R: each failure situation $(r_2, r_3, \ldots, r_N)$ is then a binary vector belonging to $\Omega = \{0,1\}^{N-1}$. For instance, when R contains only $N = 3$ customers, we have $\Omega = \{(0,0), (1,0), (0,1), (1,1)\}$, where the binary vectors mean that the

vehicle needs to perform a round trip to the depot, respectively, "never", "when it reaches the second customer", "when it reaches the third customer", and "when it reaches both the second and third customers".

Furthermore, let $g : \Omega \to \mathbb{R}^+$ be a function representing the cost of each failure situation $\omega \in \Omega$, with ω being the binary vector $(r_2, r_3, \ldots, r_N)$ representing a failure situation. Since the penalty cost upon failure on customer i is $2c_{0,i}$ (a failure implies a return trip to the depot), the cost associated to failure situation ω is

$$g(\omega) = \sum_{i=2}^{N} r_i 2 c_{0,i}. \tag{39}$$

Let m^Ω be a MF representing uncertainty towards the actual failure situation occurring on R – as will be shown in the next section, evidential demands may induce such a MF.

Then, adopting a similar pessimistic attitude as in the recourse approach to belief linear programming [40], the upper expected penalty cost $C_P^*(R)$ of route R may be obtained using (14) as follows:

$$C_P^*(R) = E^*(g, m^\Omega). \tag{40}$$

Accordingly, the upper expected cost $C_E^*(R)$ of route R may be defined as

$$C_E^*(R) = C(R) + C_P^*(R), \tag{41}$$

with $C(R)$ the cost (2) of travelling along route R when no failure occurs.

The CVRPED under the above recourse policy, may then be modelled using a modified version of the CVRP model of Section 2.1. Specifically, our recourse modelling of the CVRPED aims at

$$\min \sum_{k=1}^{m} C_E^*(R_k), \tag{42}$$

subject to constraints (3) - (6), with constraints (5) replaced by constraints (11).

Evaluating the objective function (42) requires the computation for each route, of the MF m^Ω representing uncertainty on the actual failure situation occurring on the route. This is detailed in the next section.

3.2.2. Uncertainty on recourses

Consider again a route R containing N customers. In addition, let us first assume that client demands on N are known without any uncertainty, that is we know that the demand of client i, $i = 1, \ldots, N$, is some value $\theta_i \in \Theta$. Then, it is clear that the above recourse policy amounts to the following definition for the binary failure variables r_i:

$$r_i \;=\; \begin{cases} 1, & \text{if } q_{i-1} + \theta_i > Q, \\ 0, & \text{otherwise}, \end{cases} \qquad \forall i \in \{2, \ldots, N\} \tag{43}$$

where q_j, $j = 1, \ldots, N$, denotes the load in the vehicle after serving the j-th customer such that $q_j = \theta_1$ for $j = 1$ and, for $j = 2, \ldots, N$,

$$q_j = \begin{cases} q_{j-1} + \theta_j - Q, & \text{if } q_{j-1} + \theta_j > Q, \\ q_{j-1} + \theta_j, & \text{otherwise.} \end{cases} \tag{44}$$

In other words, when it is known that the demand of the i-th customer is θ_i, $i = 1, \ldots, N$, then we have a precise demand vector on R that induces a precise binary failure situation vector $(r_2, r_3, \ldots, r_N)$, with r_i defined by (43). This can be encoded by a function $f : \Theta^N \to \Omega$, s.t. $f(\theta_1, \ldots, \theta_N) = (r_2, r_3, \ldots, r_N)$. For example, suppose we have $N = 3$ customers on route R, with respective demands $\theta_1 = 3, \theta_2 = 3$ and $\theta_3 = 5$, and the vehicle capacity limit is $Q = 5$. In such case, $f(\theta_1, \theta_2, \theta_3)$ implies the failure situation vector $(r_2 = 1, r_3 = 1)$.

In the general case, client demands on R are known in the form of a MF $m^{X_1 \times \cdots \times X_N}$, which is the marginalisation of m^{Θ^n} on the domains of the evidential variables $dr_1, \ldots, dr_N$ associated with the N clients on the route, with X_i the domain of the evidential variable dr_i associated with the i-th client on R. In such case, using (25) with f defined in the preceding paragraph, uncertainty on the actual failure situation on R is represented by a MF m^Ω defined as

$$m^\Omega(B) = \sum_{f(A)=B} m^{X_1 \times \cdots \times X_N}(A), \quad \forall B \subseteq \Omega. \tag{45}$$

Computing m^Ω defined by (45) involves evaluating $f(A)$ for any focal set A of $m^{X_1 \times \cdots \times X_N}$. Evaluating $f(A)$ for some $A \subseteq X_1 \times \cdots \times X_N$, implies $|A|$ (and thus at worst Q^N) times the evaluation of function f at some point $(\theta_1, \ldots, \theta_N) \in \Theta^N$. Hence, computing Equation (45) is generally intractable. Nonetheless, in a particular case, it is possible to compute $f(A)$, and thus Equation (45), with a much more manageable complexity:

Remark 5. *When the focal sets of $m^{X_1 \times \cdots \times X_N}$ are all Cartesian products of N intervals, i.e., for all $A \subseteq X_1 \times \cdots \times X_N$ such that $m^{X_1 \times \cdots \times X_N}(A) > 0$, we have $A = A^{\downarrow X_1} \times \cdots \times A^{\downarrow X_N}$ with, for $i = 1, \ldots, N$, $A^{\downarrow X_i} = [\![\underline{A_i}; \overline{A_i}]\!]$, it becomes possible to compute $f(A)$ with a complexity of the order 2^N, as detailed in the next section, and thus in this case if $m^{X_1 \times \cdots \times X_N}$ has at most c focal sets, the worst-case complexity to evaluate Equation (45) is $\mathcal{O}(2^N \cdot c)$.*

3.2.3. Interval demands

Let us consider a route R with N customers, such that the demand of customer i, $i = 1, \ldots, N$, is known in the form of an interval of positive integers, which we denote by $[\![\underline{A_i}; \overline{A_i}]\!]$, where $\underline{A_i} \geq 1$ and $\overline{A_i} \leq Q$. In this case, the failure situation on R belongs surely to $f\left([\![\underline{A_1}; \overline{A_1}]\!] \times \cdots \times [\![\underline{A_N}; \overline{A_N}]\!]\right) \subseteq \Omega$. Hereafter, we provide a method to efficiently compute $f\left([\![\underline{A_1}; \overline{A_1}]\!] \times \cdots \times [\![\underline{A_N}; \overline{A_N}]\!]\right)$.

In a nutshell, this method consists in generating a rooted binary tree, which represents synthetically yet exhaustively what can possibly happen on R in terms of failure situations.

More precisely, this tree is based on the following remark. Suppose a vehicle travelling along R and all that is known about its load when it arrives at the i-th customer on R is that its load belongs to an interval $[\![\underline{q}; \overline{q}]\!]$. Let us denote by q_i its load after visiting the i-th customer. Then, there are three exclusive cases:

1. either $\overline{q} + \overline{A_i} \leq Q$, hence there will surely be no failure at that customer and all that is known is that $q_i \in [\![\underline{q}; \overline{q}]\!] + [\![\underline{A_i}; \overline{A_i}]\!]$;

2. or $\underline{q} + \underline{A_i} > Q$, hence there will surely be a failure at that customer and all that is known is that $q_i \in [\![\underline{q}; \overline{q}]\!] + [\![\underline{A_i}; \overline{A_i}]\!] - Q$;

3. or $\underline{q} + \underline{A_i} \leq Q < \overline{q} + \overline{A_i}$, hence it is not sure whether there will be or not a failure at that customer. However, we can be sure that if there is no failure at that customer, *i.e.*, the sum of the actual vehicle load and of the actual customer demand is lower or equal to Q, then it means that $q_i \in [\![\underline{q} + \underline{A_i}; Q]\!]$; and if there is a failure at that customer, then it means that $q_i \in [\![1; \overline{q} + \overline{A_i} - Q]\!]$.

By applying the above reasoning repeatedly, starting from the first customer and ending at the last customer, whilst accounting for and keeping track of all possibilities and their associated failures (or absence thereof) along the way, one obtains a binary tree. The tree levels are associated to the customers according to their order on R. The nodes at a level i represent the different possibilities in terms of imprecise knowledge about the vehicle load after the i-th customer, and they also store whether these imprecise pieces of knowledge about the load were obtained following a failure or an absence of failure at the i-th customer. The pseudo code of the complete tree induction procedure is provided in Algorithm 1 and illustrated afterwards by Example 4.
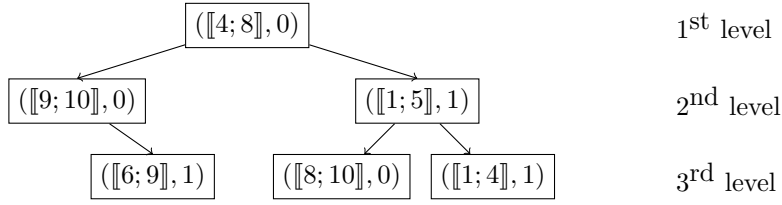
$$([\![4;8]\!],0) \qquad \text{1}^{\text{st}} \text{ level}$$

$$([\![9;10]\!],0) \qquad ([\![1;5]\!],1) \qquad \text{2}^{\text{nd}} \text{ level}$$

$$([\![6;9]\!],1) \qquad ([\![8;10]\!],0) \quad ([\![1;4]\!],1) \qquad \text{3}^{\text{rd}} \text{ level}$$

Figure 3: Recourse tree constructed for Example 4

Algorithm 1 Induction of Recourse Tree (RT)

Input: interval load $[\![\underline{q};\overline{q}]\!]$, Boolean failure variable r, next customer number i
Output: final tree $Tree$
1: create a root node containing interval load $[\![\underline{q};\overline{q}]\!]$ and Boolean failure r
2: **if** $i = N + 1$ **then**
3:      return $Tree = \{\text{root node}\}$
4: **else if** $\overline{q} + \overline{A_i} \leq Q$ **then**
5:      $[\![\underline{q_L};\overline{q_L}]\!] = [\![\underline{q};\overline{q}]\!] + [\![\underline{A_i};\overline{A_i}]\!]$
6:      $r_L = 0$
7:      $Tree_L = RT([\![\underline{q_L};\overline{q_L}]\!], r_L, i + 1)$
8:      attach $Tree_L$ as left branch of $Tree$
9: **else if** $\underline{q} + \underline{A_i} > Q$ **then**
10:      $[\![\underline{q_R};\overline{q_R}]\!] = [\![\underline{q};\overline{q}]\!] + [\![\underline{A_i};\overline{A_i}]\!] - Q$
11:      $r_R = 1$
12:      $Tree_R = RT([\![\underline{q_R};\overline{q_R}]\!], r_R, i + 1)$
13:      attach $Tree_R$ as right branch of $Tree$
14: **else**
15:      $[\![\underline{q_L};\overline{q_L}]\!] = [\![\underline{q} + \underline{A_i};Q]\!]$
16:      $r_L = 0$
17:      $Tree_L = RT([\![\underline{q_L};\overline{q_L}]\!], r_L, i + 1)$
18:      attach $Tree_L$ as left branch of $Tree$
19:      $[\![\underline{q_R};\overline{q_R}]\!] = [\![1;\overline{q} + \overline{A_i} - Q]\!]$
20:      $r_R = 1$
21:      $Tree_R = RT([\![\underline{q_R};\overline{q_R}]\!], r_R, i + 1)$
22:      attach $Tree_R$ as right branch of $Tree$
23: **end if**

Example 4. *Let us illustrate Algorithm 1 on a route R where $Q = 10$ and containing 3 customers, with $[\![4;8]\!]$, $[\![5;7]\!]$ and $[\![7;9]\!]$ the imprecise demands of the first, second and third customers, respectively. Since the demand of the first customer is $[\![4;8]\!]$, and there is no failure by definition at the first customer, and the customer following the first customer is the second customer, the tree is obtained with $RT([\![4;8]\!], 0, 2)$ and is shown in Fig. 3.*

For each leaf of the tree, by concatenating in a vector the Boolean failure

23

variable r_i at level $i = 2, \ldots, N$, written on the path from the root to the leaf, we obtain the binary failure situation vector $(r_2, r_3, \ldots, r_N)$. Hence, all the leaves of the tree, yield the subset $B \subseteq \Omega$. For instance, the rightmost leaf of the tree in Fig. 3 yields the failure situation vector $(r_2 = 1, r_3 = 1)$, the leftmost leaf yields $(r_2 = 0, r_3 = 1)$ and the remaining leaf yields $(r_2 = 1, r_3 = 0)$. The tree in this example yields thus the set $B = \{(1, 0), (0, 1)(1, 1)\}$.

Proposition 6. *The set B built using the tree generated by Algorithm 1 verifies* $B = f\left(\llbracket \underline{A_1}; \overline{A_1} \rrbracket \times \cdots \times \llbracket \underline{A_N}; \overline{A_N} \rrbracket\right)$.

Proof. See Appendix E. $\qquad\square$

The maximum number of leaf nodes in the tree is 2^{N-1}. Thus, the algorithmic complexity to obtain set $B \subseteq \Omega$ is of the order 2^N.

3.2.4. Particular cases of the recourse modelling of the CVRPED

In this section, some comments are provided on the behaviour of our recourse modelling, especially with respect to some particular evidential demands.

If m^{Θ^n} is Bayesian, *i.e.*, we are dealing really with a CVRPSD, then m^Ω is Bayesian on any given route R. Hence, the upper expected penalty cost $C_P^*(R)$ reduces to the classical (probabilistic) expected value of cost function g with respect to the probability mass function m^Ω, and thus our recourse modelling of the CVRPED clearly degenerates into the recourse modelling of the aforementioned CVRPSD.

We showed in Section 3.1.2 that the CVRPED-BCP can be converted, when $\underline{\beta} = \overline{\beta}$ and the evidential variables d_i, $i = 1 \ldots, n$, are independent, into an equivalent CVRPSD modelled via chance constrained programming, by transforming each evidential demand represented by MF m^{Θ_i} into a stochastic demand represented by probability mass function $p_{\overline{i}}$ obtained from m^{Θ_i} by transferring the mass $m^{\Theta_i}(A)$ to the element $\theta = \max(A)$. Example 5 shows that under the recourse approach, this latter transformation cannot be used in general to convert a CVRPED into an equivalent CVRPSD.

Example 5. *Suppose we have one available vehicle with capacity limit $Q = 14$, $n = 3$ clients with $\llbracket 2; 8 \rrbracket$, $\llbracket 3; 8 \rrbracket$ and $\llbracket 3; 8 \rrbracket$ the imprecise demands of clients 1, 2 and 3, respectively. The depot is denoted by 0 and the travel cost matrix $C = (c_{i,j})$ where $i, j \in \{0, 1, 2, 3\}$ is shown in Table 1. Under the recourse approach, the optimal solution to this CVRPED instance is the route defined by the path $(0, 3, 2, 1, 0)$ (its upper expected cost is 9.3). Using the above-mentioned transformation to transform the evidential demands into stochastic demands, we obtain $p_{\overline{1}}(8) = 1, p_{\overline{2}}(8) = 1$, and $p_{\overline{3}}(8) = 1$, and under the recourse approach the optimal solution to this CVRPSD instance is either the route defined by the path $(0, 2, 1, 3, 0)$ or $(0, 3, 1, 2, 0)$ (the expected cost of each one of these routes being 9.2), which are different from the optimum found for the CVRPED.*

Furthermore, let us remark that for a given route R containing N clients whose demands are known in the form of a MF $m^{X_1 \times \cdots \times X_N}$, its upper expected

Table 1: Travel cost matrix C

	0	1	2	3
0	$+\infty$	1	1.1	3
1	1	$+\infty$	1	2.1
2	1.1	1	$+\infty$	2.1
3	3	2.1	2.1	$+\infty$

cost $C_E^*(R)$ is necessarily reached for some probability measure $P^{X_1 \times \cdots \times X_N}$ belonging to the set $\mathcal{P}(m^{X_1 \times \cdots \times X_N})$ of probability measures compatible with $m^{X_1 \times \cdots \times X_N}$ and defined by

$$\mathcal{P}(m^{X_1 \times \cdots \times X_N}) = \{P | \forall A \subseteq X_1 \times \cdots \times X_N, Bel^{X_1 \times \cdots \times X_N}(A) < P(A)\}. \quad (46)$$

However, this measure cannot be determined easily in advance; it can be exhibited once possible recourses on the route and their associated costs are known. Moreover, we note that it is impossible to find a sensible transformation of evidential demands into stochastic demands such that each solution has the same costs in both the evidential and stochastic approaches, as shown by Example 6.

Example 6. *Suppose a problem where we have $m = 2$ available vehicles and $n = 5$ clients with demands represented by MF m^{Θ^n} defined as*

$$m^{\Theta^n}(\llbracket 7 \rrbracket \times \llbracket 6 \rrbracket \times \llbracket 1; 2 \rrbracket \times \llbracket 1 \rrbracket \times \llbracket 1 \rrbracket) = 1. \quad (47)$$

Consider the two following possible solutions, containing each two routes:

$$\begin{aligned} S_1 &= \{R_1^1, R_2^1\}, \\ S_2 &= \{R_1^2, R_2^2\}, \end{aligned}$$

with $R_1^1 = (0, 1, 2, 3, 0)$, $R_2^1 = (0, 4, 5, 0)$ and $R_1^2 = (0, 1, 2, 3, 5, 0)$, $R_2^2 = (0, 4, 0)$. The upper expected cost of S_i, $i = 1, 2$, is $C_E^*(S_i) := C_E^*(R_1^i) + C_E^*(R_2^i)$. Consider now the possible transformations of the evidential demands m^{Θ^n} into stochastic demands p^{Θ^n}: it seems sensible that p^{Θ^n} be chosen in the set of probability distributions compatible with m^{Θ^n}, that is in

$$\mathcal{P}(m^{\Theta^n}) = \{(p^{\Theta^n}(7, 6, 1, 1, 1) = \alpha, p^{\Theta^n}(7, 6, 2, 1, 1) = 1 - \alpha) | \alpha \in [0, 1]\}. \quad (48)$$

Let $C_E^\alpha(S_i)$ denote the expected cost of solution S_i, $i = 1, 2$, under stochastic demands represented by probability distribution $p_\alpha^{\Theta^n}$ defined by $p_\alpha^{\Theta^n}(7, 6, 1, 1, 1) = 1 - \alpha, p_\alpha^{\Theta^n}(7, 6, 2, 1, 1) = \alpha$, for some $\alpha \in [0, 1]$. Assume that customer 5 is further away from the depot than customer 3, that is $c_{0,5} > c_{0,3}$. Then, we obtain

$$C_E^*(S_1) - C_E^\alpha(S_1) = 2c_{0,3} \cdot (1 - \alpha) \quad (49)$$

and

$$C_E^*(S_2) - C_E^\alpha(S_2) = 2c_{0,5} \cdot \alpha + 2c_{0,3} \cdot \alpha. \quad (50)$$

The upper expected cost $C_E^*(S_1)$ of S_1 is thus reached for probability distribution $p_1^{\Theta^n}$, whereas the upper expected cost $C_E^*(S_2)$ of S_2 is reached for probability distribution $p_0^{\Theta^n}$. Hence, for S_1 to have the same costs in both the evidential and stochastic approaches and for S_2 to also have the same costs in both the evidential and stochastic approaches, different transformations of evidential demands into stochastic demands must be used.

The objective function (42) of our recourse model relies on optimising against $C_{\mathrm{E}}^*(R)$ which is the upper, *i.e.*, worst, *expected* cost of a route. In particular, if m^{Θ^n} is categorical, then $C_{\mathrm{E}}^*(R)$ is the worst *possible* cost of R. Hence, optimising (42) has some similarities with the protection against the worst case popular in robust optimisation [49].

Though, another approach could be followed [52], where one optimises against the lower, *i.e.*, best, expected cost $C_{*\mathrm{E}}(R) = C(R) + C_{*\mathrm{P}}(R)$, where $C_{*\mathrm{P}}(R)$ is evaluated using (13) such that $C_{*\mathrm{P}}(R) = E_*(g, m^{\Omega})$. This approach is appropriate when we are interested in the most optimistic solution. More complex decision schemes could also be considered, such as interval dominance [52], which would rely on $C_{\mathrm{E}}^*(R)$ and $C_{*\mathrm{E}}(R)$ and yield in general a set of optimal (non-dominated) solutions. Borrowing from what is done in label ranking [22], an interesting study would then be to identify from the set of non-dominated solutions, some parts of routes that would be more relevant (or preferred) to be included in a solution, over some irrelevant ones. This is left for future work.

Finally, another interesting particular case is when m^{Θ^n} is obtained from marginal knowledge about individual customer demands in the form of consonant MF m^{Θ_i}, $i = 1, \ldots, n$, having interval focal sets and the assumption that the demands are independent or non-interactive. As recalled in Section 3.1.2, the focal sets of m^{Θ^n} are then Cartesian products of intervals under both assumptions, in which case the recourse modelling of the CVRPED becomes tractable according to Remark 5.

3.2.5. Influence of customer demand specificity on the CVRPED-recourse optimal solution cost

In this section, we study the behaviour of the optimal solution cost of the CVPRED modelled via the recourse approach (in the remainder of this article, we will simply say CVRPED-recourse rather than the CVRPED modelled via the recourse approach), when knowledge specificity about customer demands decreases.

Specifically, let m^{Θ^n} and $m_\star^{\Theta^n}$ be two MF representing uncertain knowledge about customer demands, such that evidential variables d_i, $i = 1, \ldots, n$, are independent according to both these mass functions. Furthermore, let $m^{\Theta_i} := m^{\Theta^n \downarrow \Theta_i}$ and $m_\star^{\Theta_i} := m_\star^{\Theta^n \downarrow \Theta_i}$, $i = 1, \ldots, n$. Denote by $\hat{C}_{Rec}$ and $\hat{C}_{Rec}^\star$ the costs of optimal solutions to the CVRPED-recourse when customer demands are known in the form of m^{Θ^n} and $m_\star^{\Theta^n}$, respectively.

The following proposition holds:

Proposition 7. $m^{\Theta_i} \sqsubseteq m_\star^{\Theta_i}, i = 1, \ldots, n \Rightarrow \hat{C}_{Rec} \leq \hat{C}_{Rec}^\star$.

Proof. See Appendix F. □

Informally, Proposition 7 shows that the less specific knowledge is about customer demands, the greater the cost of the optimal solution.

An immediate consequence of this result is:

Corollary 3. *Assume that for $i = 1, \ldots, n$, $m_\star^{\Theta_i}$ is built from m^{Θ_i} as follows: for each $A \subseteq \Theta_i$ such that $m^{\Theta_i}(A) > 0$, the mass $m^{\Theta_i}(A)$ is transferred to a subset $A^\star$ such that $A \subseteq A^\star \subseteq \Theta_i$. Then, we have $\hat{C}_{Rec} \leq \hat{C}_{Rec}^\star$.*

Proposition 7 and Propositions 2–5 provide theoretical properties of the CVRPED solutions obtained under the recourse and the BCP approaches, respectively, when using exact optimisation methods. For now, such methods can not solve large instances of the CVRP, from which the CVRPED derives. As a matter of fact, Section 4 reports solution strategies to these two CVRPED models using a metaheuristic algorithm.

4. Solving the CVRPED

This section presents a metaheuristic algorithm to solve the two proposed CVRPED models and reports some experimental tests of this algorithm. More precisely, a simulated annealing algorithm for both the BCP and the recourse models is first described in Section 4.1. Next, benchmarks for the CVRPED are presented in Section 4.2. Finally, experimental tests using these benchmarks are provided in Sections 4.3 and 4.4 for the BCP and the recourse models, respectively.

4.1. A simulated annealing algorithm for the CVRPED

As for many NP-hard problems like the CVRP, exact solution methods might need a prohibitively large time to solve large instances. When uncertainty is introduced into the CVRP, the problem can easily become even more difficult. Metaheuristics are algorithms built on general (meta) concepts that search the solution space in a reasonable time and thus may be employed as a successful alternative to solve such a combinatorial optimisation problem. Various metaheuristic algorithms exist like: simulated annealing [37], genetic algorithms [36], tabu search [30, 31], etc. In this study, we will use simulated annealing, which is a well-known local search optimisation method. Generally speaking, a local search algorithm moves iteratively from solution to solution in the space of candidate solutions (the search space) by applying local changes, until a satisfying near-optimal solution is found.

Indeed, to try to find the global minimum of the cost function, the simulated annealing moves from solution to solution using either descent (improving) moves or deterioration moves, hoping that these non-improving moves will eventually help the process escape local optima [37]. One can say that it starts from a known initial configuration of a system (candidate solution) with a high temperature and then uses general neighbourhood search strategies to explore other

27

candidate solutions by following neighbourhood transitions (moves). For every temperature, the configuration of the system (candidate solution) is rearranged (transformed) by a series of neighbourhood moves. A rearranged configuration becomes the new candidate solution with a probability depending on the current temperature, *i.e.*, the lower the temperature, the lower the probability to accept non-improving moves. The temperature of the system is gradually lowered, and the process continues until reaching the freezing temperature of the system.

Our instantiation of simulated annealing for the CVRPED is provided by Appendix G (see in particular its pseudo-code given by Algorithm 2). The same pseudo-code is used for both the BCP and the recourse approaches. One can find in Algorithm 2 a modelling technique parameter MD that takes the value $MD =$"BCP" when the CVRPED-BCP is solved. In this case, the initial configuration is generated either randomly or using a first-fit greedy approach[3] while respecting the constraints of the BCP model from Section 3.1.1. The `neighbourhood_configuration(...)` routine generates a neighbourhood configuration that respects the CVRPED-BCP constraints, and is based on the following two operators that are applied consecutively on the current configuration $\mathcal{C}$ at each iteration. These operators are called `fix_minimum` and `replace_highest_average`. They are described below and illustrated using examples in Appendix H.

– `Fix_minimum`: This operator is applied on 80% of the iterations. It is based on selecting and freezing the positions in routes of the five customers, with the shortest distances to their right side customer[4]. This is done by computing distances between each pair of consecutive customers on all routes, including distances to the depot. Accordingly, `fix_minimum` selects the five smallest distances values and fixes their corresponding left side customers. Next, `fix_minimum` selects five random customers that exclude the depot and the customers fixed before, and removes them from their route. Subsequently, every customer removed will be inserted in a random route, while satisfying the problem constraints. The insert position of each customer on the selected route is determined based on the shortest distance separating it from its new left side customer.

– `Replace_highest_average`: This neighbourhood operator calculates the average distance separating every customer from its neighbours in a current route configuration. Computing the average distance for a client i reduces to calculating $\frac{c_{i-1,i}+c_{i,i+1}}{2}$, assuming clients $i-1$, i and $i+1$ are consecutive in the route[5]. Afterwards, `replace_highest_average` selects five customers having the five highest average distances and removes them from their routes. The removed clients are then randomly inserted in the available routes, as long

---

[3]A blind algorithm that inserts clients in routes in turn: each client is inserted in the first route that can serve it.

[4]Given the pair of customer $< i - 1, i >$, the right side customer is the i-th customer.

[5]The notation $c_{i-1,i}$ indicates the travel cost between $i-1$ and i, as defined in Section 2.1, which in our algorithm is assumed to be the euclidean distance separating customers.

as the problem constraints are respected. Furthermore, every removed customer i is inserted into the route position leading to the smallest average distance that will separate this customer from its new $(i-1)$ th and $(i+1)$ th neighbours.

If the above operators do not lead to a new neighbourhood configuration that satisfies the CVRPED-BCP constraints, we apply the operators a second time. If the second attempt fails as well, then the configuration $\mathcal{C}$ is not modified, $i.e.$, $C^* = \mathcal{C}$.

The complexity of an iteration for the BCP model emerges from evaluating the CVRPED-BCP constraints for the neighboring configuration, in particular the belief constraints (26) and (27), the complexity of which is provided in Section 3.1.1.

We solve the CVRPED-recourse with the same simulated annealing, by setting the modelling technique parameter to $MD =$ "recourse" in Algorithm 2. The `initial_config(...)` method proceeds by generating initial configurations that are subject to the CVRPED-recourse constraints mentioned in Section 3.2.1, using either a first fit greedy approach or randomly. The routine `neighbourhood_configuration(...)` applies three consecutive neighbourhood operators to each iteration: `fix_minimum`, `replace_highest_average` and `flip_route`.

– `Fix_minimum` and `replace_highest_average` operate similarly as in the BCP case, except that the problem constraints are now the CVRPED-recourse constraints. Recall that the recourse model lifts the capacity constraints, in the sense that any capacity excess (overflow) is addressed by the objective function using recourse decisions. This means that as `fix_minimum` and `replace_highest_average` are iteratively applied by the simulated annealing, we may end up with routes holding excessive total demands. Consequently, these routes will systematically fail, while other routes will hold limited total customer demands. Therefore, we incorporate to each of the `fix_minimum` and `replace_highest_average` operators, a method that maintains relatively balanced total demands on routes, during the process of moving customers from a route to another. More specifically, before a customer is inserted into a new route, total customer demands on each route yields a probability, indicating if a route is favourable for servicing an additional customer. The probability associated to each route is inversely proportional to the total customer demands on a route, $i.e.$, the smaller the total customer demands is on a route, the more it is probable to choose this route to include an additional customer.

– The operator `flip_route` is applied on 25% of the iterations. It reverses the order of a route if this improves its upper expected cost. Indeed, a route R having the path $(0, 1, \ldots, N, 0)$ and its reverse R^{-1} with the path $(0, N, \ldots, 1, 0)$ do not have necessarily the same upper expected penalty cost.

The complexity of an iteration for the recourse model corresponds to evaluating the uncertainty on the recourses of each one of the m routes of a configuration

$\mathcal{C}$, the complexity of which is given in Section 3.2.2.

Let us finally emphasize that the evidential approaches proposed in this paper are not limited to using simulated annealing or metaheuristics. Some of the exact methods used for the CVRPSD could be extended to these approaches. This is the case, for instance, of the Column Generation (CG) algorithm proposed in [12] for the recourse modelling of the CVRPSD. A CG algorithm relies on a linear program with prohibitively many variables (columns) associated to feasible routes. One could use our evidential approach to evaluate the total cost (base cost plus recourse cost) of each generated column (route), especially when the columns are constructed by Dynamic Programming (DP).

More exactly, each column (route) can be constructed by generating a cycle in a graph of DP states associated to sub-problems satisfying the Bellman principle of optimality. Assuming joint focal sets about customer demands are Cartesian products of intervals, one could define a DP state $\left(v, [\underline{q_v}, \overline{q_v}]\right)$ for each vertex $v \in V$ and for each interval of residual supply $[\underline{q_v}, \overline{q_v}]$ at v (quantity remaining in a vehicle after servicing clients up to v and performing recourses). The total cost of a route in a given state can be determined, for instance, as the sum of the traversed edges plus the upper expected cost of the recourse actions. In a loose sense, this is similar to the DP approach from [12] that constructs a state for each vertex v and for each possible value of the cumulative expected demand (of all the clients in the route up to v) and of the variation of this cumulative demand.

4.2. The CVRPED benchmarks

We generated two instance sets CVRPED and CVRPED$^+$, based on the set A of the Augerat test bed for the CVRP [46]. Each instance in these two sets corresponds to an instance in Augerat set A and has the same customer coordinates and capacity limit as this instance.

For each instance of the first CVRPED set, the knowledge on customer demands m^{Θ^n} is obtained by assuming that the evidential client demands $d_i, i = 1, \ldots, n$ of this instance are independent. Moreover, each d_i is associated to the mass function m^{Θ_i} defined by

$$m^{\Theta_i}(\{d_i^{det}\}) = 0.8, \qquad (51)$$
$$m^{\Theta_i}([\underline{z_i}, \overline{z_i}]) = 0.2, \qquad (52)$$

with d_i^{det} the original deterministic demand of client i in the corresponding instance of Augerat set A, and with $\underline{z_i}$ and $\overline{z_i}$ drawn at random in $(d_i^{det}, Q]$ and $[\underline{z_i}, Q]$, respectively.

For each instance of the second set CVRPED$^+$, the evidential client demands d_i are also assumed to be independent, and their associated mass function is denoted by $m_+^{\Theta_i}$ and defined from m^{Θ_i} as follows:

$$m_+^{\Theta_i}([d_i^{det}, d_i^{det} + a_i^+]) = 0.8, \qquad (53)$$
$$m_+^{\Theta_i}([\underline{z_i}, \overline{z_i}]) = 0.2, \qquad (54)$$

with a_i^+ drawn randomly in $[0, \underline{z_i} - d_i^{det} - 1]$. Note that $m^{\Theta_i} \sqsubseteq m_+^{\Theta_i}$ and $m^{\Theta_i} \preceq m_+^{\Theta_i}$, $i = 1, \ldots, n$.

In the next sections, an experimental study based on the CVRPED and CVRPED$^+$ instances[6] is presented for the BCP and recourse models solved using the algorithm described in Section 4.1. The programs were written in Java and the experiments were conducted on the 5 nodes of a cluster. The configuration of each node is as follows: 2 processors Intel R Xeon R E5-2630 v3 with 8 cores per processor having a 48GB memory shared between the 2×8 cores of the node. Each instance was executed on one core that has a memory of 2.8 GB.

4.3. Experimental study for the CVRPED-BCP

Our set of experiments on the CVRPED-BCP involve varying the values $\underline{\beta}$ and $\overline{\beta}$ involved in the constraints (26) and (27) for the CVRPED and the CVRPED$^+$ instances separately, where for each variation each instance was solved 30 times.

4.3.1. The CVRPED-BCP cost variation based on $\underline{\beta}$ and $\overline{\beta}$

In this first part, we show results of our experiments for the CVRPED and the CVRPED$^+$ instances in Tables 2 and 3, respectively. Indeed, we solved the CVRPED-BCP for two different values that we chose for the pair $(\underline{\beta}, \overline{\beta})$, such that $\underline{\beta} \geq \overline{\beta}$, and both values were employed for the CVRPED and the CVRPED$^+$ instances, separately. The columns figuring in each of these tables, are explained in the following. The first column is the name of each instance. The first field in this column (l in Table 2 and l^+ in Table 3), exposes the identification number of an instance and the second field "A" stands for aleatory indicating that the coordinates of the problem graph vertices were generated randomly in the original Augerat set A [3]. The third field designates the number n of vertices, while the last field provides the number m of vehicles. The "Best cost", "Std. dev." and "Avg. runtime" columns show respectively, the best solution, the standard deviation and the average running time that we obtained for each indicated value of the pair $(\underline{\beta}, \overline{\beta})$.

We notice the costs of the best solutions obtained with $\underline{\beta} = 0.4$, $\overline{\beta} = 0.25$ are lower than the costs of the best solutions obtained with $\underline{\beta} = 0.2$, $\overline{\beta} = 0.15$ in Table 2 as well as in Table 3, that is, the most constraining pair $(\underline{\beta}, \overline{\beta})$ induces the worst costs, as can be expected from Propositions 3 and 4. This shows that while our solving algorithm is not an exact optimisation method, it does exhibit experimentally a sound behaviour with respect to parameters $\underline{\beta}$ and $\overline{\beta}$.

4.3.2. The CVRPED-BCP cost variation based on client demand ranking

This section compares the BCP results on the CVRPED instances with the BCP results on the CVRPED$^+$ instances. Specifically, Table 4 compares the

---

[6]Our data sets can be found on the web site of the LGI2A laboratory [34].

Table 2: Results of the simulated annealing algorithm for the CVRPED-BCP using the CVRPED instances

Instance l-A-nn-mm :	$\underline{\beta} = 0.4, \overline{\beta} = 0.25$			$\underline{\beta} = 0.2, \overline{\beta} = 0.15$		
l instance id,	Best	Std.	Avg.	Best	Std.	Avg.
n clients, m vehicles	cost	dev.	runtime	cost	dev.	runtime
1-A-n32-m12	1418,3	3,7	3881s.	1850,9	5,3	3733s.
2-A-n33-m13	1055,3	0	4199s.	1491,6	17,5	4496s.
3-A-n33-m13	1073,1	6	4495s.	1480,2	0,6	4549s.
4-A-n34-m14	1320,6	0,1	3818s.	1749,1	0	3852s.
5-A-n36-m12	1318,9	2,7	5316s.	1718,6	0,2	4914s.
6-A-n37-m13	1110,6	5	4918s.	1358,8	34,5	6158s.
7-A-n37-m14	1597,9	0,5	4135s.	2113,9	2,8	3756s.
8-A-n38-m13	1154,5	0,9	5041s.	1571,1	5,1	5002s.
9-A-n39-m15	1485	8,1	4654s.	1944,8	0,9	4622s.
10-A-n39-m14	1403,9	8,6	4894s.	1906,9	0,2	5108s.
11-A-n44-m17	1693,4	10,1	4956s.	2158,2	1,2	4951s.
12-A-n45-m17	1660,3	0,1	5093s.	2184,8	5,1	5169s.
13-A-n45-m18	1890,1	5,7	4991s.	2573,2	0,7	5211s.
14-A-n46-m17	1552,2	5,4	5323s.	1980,3	38	5707s.
15-A-n48-m17	1872,4	10,8	5996s.	2397,5	3,6	6395s.
16-A-n53-m19	1806,1	11,9	7405s.	2358,2	10	6928s.
17-A-n54-m19	2052,6	13,8	7578s.	2636,8	60,8	6747s.
18-A-n55-m22	1755,3	9,5	6310s.	2352,6	9,8	6172s.
19-A-n60-m22	2263,9	17	8169s.	2969,1	34,8	7449s.
20-A-n61-m24	1793,8	9,9	6965s.	2345	42,3	7319s.
21-A-n62-m22	2532,3	19,1	8212s.	3207,2	32,2	7752s.
22-A-n63-m24	2946,9	14,6	7164s.	3918,9	22	7216s.
23-A-n63-m25	2179	8,9	6969s.	2881,1	6,9	7238s.
24-A-n64-m23	2629,1	16,8	7979s.	3261,5	13,6	7848s.
25-A-n65-m25	2214,7	16,4	7537s.	3070,9	6	7601s.
26-A-n69-m25	2056,5	11,1	8876s.	2668,1	52,9	8377s.
27-A-n80-m27	3507,2	21,5	11110s.	4524,9	21,5	9751s.

Table 3: Results of the simulated annealing algorithm for the CVRPED-BCP using the CVRPED$^+$ instances

Instance l^+-A-nn-mm :	$\underline{\beta} = 0.4, \overline{\beta} = 0.25$			$\underline{\beta} = 0.2, \overline{\beta} = 0.15$		
l^+ instance id,	Best	Std.	Avg.	Best	Std.	Avg.
n clients, m vehicles	cost	dev.	runtime	cost	dev.	runtime
1^+-A-n32-m16	1830,8	28,3	3087s.	2225,4	0	3565s.
2^+-A-n33-m16	1428,8	0,7	4048s.	1676,9	0,4	3790s.
3^+-A-n33-m14	1196	0	4330s.	1502	24,2	4604s.
4^+-A-n34-m16	1596	0,01	3389s.	1993	0	3719s.
5^+-A-n36-m16	1755,3	4	3928s.	2145,7	0	4178s.
6^+-A-n37-m18	1379,2	8,7	5179s.	1761,2	0	4347s.
7^+-A-n37-m19	1957	0	3344s.	2542,6	0	4112s.
8^+-A-n38-m16	1437,4	2,2	4017s.	1846,4	2,2	4479s.
9^+-A-n39-m19	1915,5	11,2	3915s.	2203,5	0	4089s.
10^+-A-n39-m18	1759	4,2	4255s.	2148,1	0	4140s.
11^+-A-n44-m26	2234,2	1,5	4218s.	2796,6	0,4	5698s.
12^+-A-n45-m22	2165,7	3.1	4360s.	2690,8	0	4601s.
13^+-A-n45-m22	2287,7	0,9	4339s.	3099,4	0	4783s.
14^+-A-n46-m22	1950,4	6,7	4603s.	2690,1	0	4795s.
15^+-A-n48-m22	2359,8	5,9	5742s.	2956,3	1,8	5500s.
16^+-A-n53-m25	2411,2	9,4	5747s.	3199,8	0	5883s.
17^+-A-n54-m24	2591	11,4	6025s.	3165,7	0	5746s.
18^+-A-n55-m27	2237,1	5,1	5716s.	2803,1	0	5493s.
19^+-A-n60-m27	2744,8	9,5	6224s.	3415,5	6,4	6548s.
20^+-A-n61-m30	2313,4	9,1	6125s.	3059,2	0	5877s.
21^+-A-n62-m31	3217,7	8,9	6814s.	4291	3,1	6276s.
22^+-A-n63-m30	3833	9,7	6057s.	4942,1	4,9	6257s.
23^+-A-n63-m31	2755	5,8	5757s.	3638,6	0,4	6127s.
24^+-A-n64-m29	3311,2	27,9	6656s.	4123,1	1,7	6848s.
25^+-A-n65-m34	2748,1	7,7	6762s.	3509,1	9	6842s.
26^+-A-n69-m33	2573,4	8,5	7135s.	3300,5	41,1	7125s.
27^+-A-n80-m38	4995,8	14,7	7444s.	5986,3	13,5	7612s.

Table 4: Results of the simulated annealing algorithm for the CVRPED-BCP for the CVRPED and CVRPED$^+$ instances

CVRPED instances			CVRPED$^+$ instances		
Instance	Best cost	Best cost	Instance	Best cost	Best cost
id l	$\underline{\beta} = 0.4, \overline{\beta} = 0.25$	$\underline{\beta} = 0.2, \overline{\beta} = 0.15$	id l^+	$\underline{\beta} = 0.4, \overline{\beta} = 0.25$	$\underline{\beta} = 0.2, \overline{\beta} = 0.15$
1	1418,3	1850,9	1^+	1830,8	2225,4
2	1055,3	1491,6	2^+	1428,8	1676,9
3	1073,1	1480,2	3^+	1196	1502
4	1320,6	1749,1	4^+	1596	1993
5	1318,9	1718,6	5^+	1755,3	2145,7
6	1110,6	1358,8	6^+	1379,2	1761,2
7	1597,9	2113,9	7^+	1957	2542,6
8	1154,5	1571,1	8^+	1437,4	1846,4
9	1485	1944,8	9^+	1915,5	2203,5
10	1403,9	1906,9	10^+	1759	2148,1
11	1693,4	2158,2	11^+	2234,2	2796,6
12	1660,3	2184,8	12^+	2165,7	2690,8
13	1890,1	2573,2	13^+	2287,7	3099,4
14	1552,2	1980,3	14^+	1950,4	2690,1
15	1872,4	2397,5	15^+	2359,8	2956,3
16	1806,1	2358,2	16^+	2411,2	3199,8
17	2052,6	2636,8	17^+	2591	3165,7
18	1755,3	2352,6	18^+	2237,1	2803,1
19	2263,9	2969,1	19^+	2744,8	3415,5
20	1793,8	2345	20^+	2313,4	3059,2
21	2532,3	3207,2	21^+	3217,7	4291
22	2946,9	3918,9	22^+	3833	4942,1
23	2179	2881,1	23^+	2755	3638,6
24	2629,1	3261,5	24^+	3311,2	4123,1
25	2214,7	3070,9	25^+	2748,1	3509,1
26	2056,5	2668,1	26^+	2573,4	3300,5
27	3507,2	4524,9	27^+	4995,8	5986,3

best costs from Table 2 (CVRPED instances) with the bests costs from Table 3 (CVRPED+ instances), for the same instance id.

Recall that for each client i in an l instance, its MF m^{Θ_i} is at least as small as the associated MF to client i in the l^+ instance, $i.e.$, $m^{\Theta_i} \preceq m_+^{\Theta_i}$, $i = 1, \ldots, n$. Proposition 5 (and more specifically Corollary 1) predicted an increase in the cost of an optimal solution to the CVRPED-BCP when knowledge about clients demands is more pessimistic. We can observe this behaviour in the results presented in Table 4: for each pair $(\underline{\beta}, \overline{\beta})$ the best cost obtained with the CVRPED$^+$ instances is higher than the one obtained with the CVRPED instances. This constitutes another experimental validation of the behaviour of our algorithm.

4.4. Experimental study for the CVRPED-recourse

In the experiments conducted for the CVRPED-recourse, we used the same generated CVRPED and CVRPED$^+$ instances used for the CVRPED-BCP experiments. Our results are reported in Table 5.

The columns "Instance id l" and "Instance id l^+" in this table represent the CVRPED and the CVRPED$^+$ instances id, respectively. Each one of these instances was solved 30 times and the best, average and standard deviation of the costs along with the average running times are reported in the respective columns "Best cost", "Avg cost", "Stand. dev." and "Avg. runtime" for the CVRPED and the CVRPED$^+$ instances, separately. In the "Penalty cost" column for the CVRPED instances (respectively the "Penalty cost" column for the CVRPED$^+$ instances), the contribution of the expected penalty costs to the overall costs of the best solutions to the CVRPED instances (respectively the CVRPED$^+$ instances) is provided as percentages. In the case of the CVRPED instances, it varies between 16% to 25%. As for the CVRPED$^+$ instances, it varies between 11% and 23%.

Recall that for each client i in an l instance, its MF m^{Θ_i} is at least as specific as the associated MF to client i in the l^+ instance, $i.e.$, $m^{\Theta_i} \sqsubseteq m_+^{\Theta_i}$, $i = 1, \ldots, n$. As expected from Proposition 7 (and more specifically from Corollary 3), best costs obtained with the CVRPED$^+$ instances are higher than those obtained with the CVRPED instances, which shows that our algorithm for the recourse model exhibits experimentally also a sound behaviour.

5. Conclusions

In this article, we proposed to represent uncertainty on customer demands in the capacitated vehicle routing problem via the theory of evidence. We tackled this problem by generalising the most popular approaches to stochastic programming: chance constrained programming and stochastic programming with recourse. We obtained belief constrained programming and evidential recourse approaches. We studied the optimal solution cost behaviour with respect to the model parameters and customer demand ranking in the case of the belief constrained programming model, and with respect to customer demand specificity in the case of the recourse model. In addition, by considering particular cases of evidential demands, we were able to connect our models not only to stochastic programming but also to robust optimisation. In the last part of this article, we solved both models by a simulated annealing metaheuristic algorithm that uses a combination of operators that aim at minimising the objective function of the problem. We reported the results of our experiments on instances of this difficult optimisation problem. Our experiments showed that our algorithm behave accordingly to the theoretical results studied in this paper.

Future work will include i) comparing our evidential models to the other models, and particularly the stochastic ones, using historical data on customer demands in order to show empirically the advantages of our evidential models; ii) extending to the evidential framework other stochastic variations of the CVRP, such as the CVRP with stochastic customers [28]; iii) extending our evidential models to the case of incomplete knowledge about the dependency between the evidential variables [21]; iv) performing a sensitivity analysis that would allow us to identify certain "key" customers, such that better knowledge about their demands leads to better solutions in each one of these models; and

Table 5: Results of the simulated annealing algorithm for the CVRPED-recourse for the CVRPED and CVRPED$^+$ instances

CVRPED instances					CVRPED$^+$ instances				
Instance id l	Best cost	Penalty cost	Avg cost	Stand. dev.	Avg. runtime	Instance id l^+	Best cost	Penalty cost	Avg cost	Stand. dev.	Avg. runtime
1	1750,3	16,8%	1783,9	16,1	1958s.	1$^+$	2252,6	18,3%	2283,5	13,2	1272s.
2	1327,5	16,2%	1353,2	13,6	1704s.	2$^+$	1650,6	17,7%	1676	9,6	1329s.
3	1296,1	18%	1338,8	16,4	1642s.	3$^+$	1490,3	16,6%	1510,6	11,3	1540s.
4	1661,9	19,8%	1698,7	24,6	1728s.	4$^+$	1999,6	19,7%	2044,9	18,7	1428s.
5	1670,1	24,2%	1741,9	29	2673s.	5$^+$	2205,3	16,9%	2247,3	18,6	1554s.
6	1391,2	20,1%	1425,8	12,5	3586s.	6$^+$	1697,5	14,9%	1737	13,2	1612s.
7	1895,6	24,4%	1947,3	21,2	2286s.	7$^+$	2561,5	17%	2593,8	17,5	1382s.
8	1493,8	16,1%	1525,6	15,9	2450s.	8$^+$	1769,7	16,6%	1802,9	18,9	1686s.
9	1851	21,1%	1897,6	27,4	2580s.	9$^+$	2319,9	19,9%	2355	20,5	1783s.
10	1715,2	22,2%	1755,6	22,9	3264s.	10$^+$	2099,3	20,7%	2146,3	20,8	1863s.
11	2127,8	20,7%	2216,7	25,3	2349s.	11$^+$	2858,5	11,8%	2889	15,5	1491s.
12	2147,1	17%	2193,7	21,1	2344s.	12$^+$	2667,9	18 %	2705,2	22	1808s.
13	2530,2	22,7%	2629,7	33,6	2427s.	13$^+$	3084,7	15,7%	3145,9	29,5	1755s.
14	1994,9	24,9%	2089,4	32,6	2948s.	14$^+$	2483,1	17,2%	2524,8	23,3	1950s.
15	2499,5	21,2%	2559,1	30,2	3348s.	15$^+$	3135,5	15,8%	3168,4	21,3	2293s.
16	2420,4	18,3%	2499,7	35	4715s.	16$^+$	3100	14,9%	3132,5	17,7	2294s.
17	2709,6	19,8%	2792,8	39,1	4129s.	17$^+$	3366,9	16,6%	3427	29,9	2572s.
18	2301,7	16,3%	2348,4	27,3	2844s.	18$^+$	2788,3	13,5%	2837,6	24,8	2110s.
19	3083	23,4%	3190,1	45,9	4193s.	19$^+$	3696,4	16%	3759,3	32,8	2706s.
20	2322,7	15,7%	2378	30,7	3398s.	20$^+$	2960,8	15,4%	3000,5	19,2	2484s.
21	3317,8	23,6%	3426,4	54,9	6604s.	21$^+$	4437,5	17,5%	4517,8	46,8	2382s.
22	4158,8	21%	4261,4	51,9	4148s.	22$^+$	5249,4	22,2%	5395,1	47,3	2692s.
23	2966,7	19,6%	3043,6	39,9	4217s.	23$^+$	3578,1	19,4%	3648,8	40,6	2727s.
24	3528,3	23,9%	3631,1	54,3	6365s.	24$^+$	4435,1	17,2%	4548,4	42,8	3162s.
25	2889,7	22,4%	3040	55,1	3857s.	25$^+$	3665,9	15,4%	3712,6	23,4	2573s.
26	2712,5	19,2%	2847,3	49,1	4461s.	26$^+$	3466,7	14,2%	3525,2	32,7	2748s.
27	5016,2	24,2%	5137,7	69,2	10401s.	27$^+$	6790,5	16,5%	6953,5	51,5	3357s.

v) considering more general uncertainty frameworks than evidence theory to model uncertainty on customer demands. Lower previsions [53], whose interest to model uncertainty about constraint parameters in optimisation problems has been investigated in [44], may be such a framework. In particular, in the case of independent demands, 2-monotone lower probabilities would offer more generality while being still tractable, thanks to the results of [19].

Acknowledgment

References

[1] N. Ben Abdallah, N. Mouhous-Voyneau, and T. Denoeux. Combining statistical and expert evidence using belief functions: Application to centennial sea level estimation taking into account climate change. *International Journal of Approximate Reasoning*, 55(1):341–354, 2014.

[2] H. Agarwal. *Reliability Based Design Optimization: Formulations and Methodologies.* PhD thesis, University of Notre Dame, Indiana, United States of America, 2004.

[3] P. Augerat. *Approche polyédrale du problème de tournées de vehicules.* PhD thesis, Institut National Polytechnique de Grenoble, 1995.

[4] C. Baudrit, I. Couso, and D. Dubois. Joint propagation of probability and possibility in risk analysis: Towards a formal framework. *International Journal of Approximate Reasoning*, 45(1):82–105, 2007.

[5] C. Baudrit, D. Dubois, and D. Guyonnet. Joint propagation and exploitation of probabilistic and possibilistic information in risk assessment. *IEEE Trans. Fuzzy Systems*, 14(5):593–608, 2006.

[6] D. Bertsimas, D. B. Brown, and C. Caramanis. Theory and applications of robust optimization. *SIAM Review*, 53(3):464–501, 2011.

[7] J. R. Birge and F. Louveaux. *Introduction to Stochastic Programming.* Springer-Verlag, New York, 1997.

[8] L. D. Bodin, B. L. Golden, A. A. Assad, and M. O. Ball. Routing and scheduling of vehicles and crews: The state of the art. *Computers and operations research*, 10(2):63–212, 1983.

[9] J. Brito, J. A. Moreno, and J. L. Verdegay. Fuzzy optimization in vehicle routing problems. In *Proceedings of the Joint 2009 International Fuzzy Systems Association World Congress and 2009 European Society of Fuzzy Logic and Technology Conference*, Lisbon, Portugal, 2009.

[10] A. Charnes, W. W. Cooper, and G. H. Symonds. Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil. *Management Science*, 4(3):235 – 263, 1958.

[11] J. Q. Chen, W. L. Li, and T. Murata. Particle swarm optimization for vehicle routing problem with uncertain demand. In *4th IEEE International Conference on Software Engineering and Service Science (ICSESS)*, Beijing, China, 2013. IEEE.

[12] C. H. Christiansen and J. Lysgaard. A branch-and-price algorithm for the capacitated vehicle routing problem with stochastic demands. *Operations Research Letters*, 35:773–781, 2007.

[13] I. Couso and S. Moral. Independence concepts in evidence theory. *International Journal of Approximate Reasoning*, 51(7):748 – 758, 2010.

[14] I. Couso, S. Moral, and P. Walley. A survey of concepts of independence for imprecise probabilities. *Risk, Decision and Policy*, 5(2):165 – 181, 2000.

[15] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *The annals of mathematical statistics*, 38(2):325–339, 1967.

[16] T. Denoeux. Analysis of evidence-theoretic decision rules for pattern classification. *Pattern recognition*, 30(7):1095–1107, 1997.

[17] T. Denoeux. Extending stochastic ordering to belief functions on the real line. *Information Sciences*, 179(9):1362 – 1376, 2009.

[18] T. Denoeux. 40 years of Dempster-Shafer theory. *International Journal of Approximate Reasoning*, 79(C):1–6, 2016.

[19] S. Destercke. Independence and 2-monotonicity: Nice to have, hard to keep. *International Journal of Approximate Reasoning*, 54:478–490, 2013.

[20] S. Destercke and I. Couso. Ranking of fuzzy intervals seen through the imprecise probabilistic lens. *Fuzzy Sets and Systems*, 278(Supplement C):20 – 39, 2015.

[21] S. Destercke and D. Dubois. Idempotent conjunctive combination of belief functions: Extending the minimum rule of possibility theory. *Information Sciences*, 181(18):3925–3945, 2011.

[22] S. Destercke, M-H. Masson, and M. Poss. Cautious label ranking with label-wise decomposition. *European Journal of Operational Research*, 246(3):927–935, 2015.

[23] M. Dror, G. Laporte, and P. Trudeau. Vehicle routing with stochastic demands: Properties and solution frameworks. *Transportation science*, 23(3):166–176, 1989.

[24] D. Dubois and H. Prade. A set-theoretic view of belief functions: logical operations and approximations by fuzzy sets. *International Journal of General Systems*, 12(3):193–226, 1986.

[25] D. Dubois and H. Prade. Random sets and fuzzy interval analysis. *Fuzzy Sets and Systems*, 42:87 – 101, 1991.

[26] D. Dubois and H. Prade. *Decision-making Process: Concepts and Methods*, chapter Formal Representations of Uncertainty, pages 85–156. ISTE, London, UK, 2009.

[27] C. Gauvin, G. Desaulniers, and M. Gendreau. A branch-cut-and-price algorithm for the vehicle routing problem with stochastic demands. *Computers and Operations Research*, 50:141–153, 2014.

[28] M. Gendreau, G. Laporte, and R. Séguin. Stochastic vehicle routing. *European Journal of Operations Research*, 88:3–12, 1996.

[29] M. Gendreau, G. Laporte, and R. Séguin. A tabu search heuristic for the vehicle routing problem with stochastic demands and customers. *Operations Research*, 44:469–477, 1996.

[30] F. Glover. Tabu search-part 1. *ORSA Journal on Computing*, 1(3):190–206, 1989.

[31] F. Glover. Tabu search-part 2. *ORSA Journal on Computing*, 2(1):4–32, 1990.

[32] H. Harmanani, D. Azar, N. Helal, and W. Keirouz. A simulated annealing algorithm for the capacitated vehicle routing problem. In *26th International Conference on Computers and their Applications*, New Orleans USA, 2011.

[33] N. Helal, F. Pichon, D. Porumbel, D. Mercier, and É. Lefèvre. The capacitated vehicle routing problem with evidential demands: a belief-constrained programming approach. In Jiřina Vejnarová and Václav Kratochvíl, editors, *Belief Functions: Theory and Applications, Proc. of the Fourth International Conference, BELIEF 2016*, volume 9861 of *Lecture Notes in Computer Science*, pages 212–221, Prague, Czech Republic, 2016. Springer.

[34] N. Helal, F. Pichon, D. Porumbel, D. Mercier, and É. Lefèvre. CVRPED benchmarks. `https://www.lgi2a.univ-artois.fr/spip/en/benchmarks/the-capacitated-vehicle-routing-problem-with-evidential-demands`, 2017. Accessed: 2017-08-25.

[35] N. Helal, F. Pichon, D. Porumbel, D. Mercier, and É. Lefèvre. A recourse approach for the capacitated vehicle routing problem with evidential demands. In Antonucci et al., editor, *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, volume 10369 of *Lecture Notes in Artificial Intelligence*, pages 190–200, Lugano, Switzerland, 2017. Springer.

[36] J. H. Holland. *Adaptation in natural and artificial systems.* University of Michigan Press, Ann Arbor, Michigan, U.S.A., 1975. (Second edition, 1992, MIT Press, Cambridge, Massachusetts).

[37] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.

[38] G. Laporte. The vehicle routing problem: An overview of exact and approximate algorithms. *European Journal of Operational Research*, 59(3):345–358, 1992.

[39] G. Laporte, F. Louveaux, and L. van Hamme. An integer l-shaped algorithm for the capacitated vehicle routing problem with stochastic demands. *Operations Research*, 50:415–423, 2002.

[40] H. Masri and F. Ben Abdelaziz. Belief linear programming. *International Journal of Approximate Reasoning*, 51:973–983, 2010.

[41] Z. P. Mourelatos and J. Zhou. A design optimization method using evidence theory. *Journal of Mechanical Design*, 128:901–908, 2006.

[42] Y. Peng and J. Chen. Vehicle routing problem with fuzzy demands and the particle swarm optimization solution. In *2010 International Conference on Management and Service Science (MASS)*, Wuhan, China, 2010. IEEE.

[43] F. Pichon, D. Dubois, and T. Denoeux. Relevance and truthfulness in information correction and fusion. *International Journal of Approximate Reasoning*, 53(2):159–175, 2012.

[44] E. Quaeghebeur, K. Shariatmadar, and G. De Cooman. Constrained optimization problems under uncertainty with coherent lower previsions. *Fuzzy Sets and Systems*, 206:74–88, 2012.

[45] J. K. Sengupta. A generalization of some distribution aspects of chance-constrained linear programming. *International Economic Review*, 11(2):287–304, 1970.

[46] Vehicle Routing Data sets. `http://www.coin-or.org/SYMPHONY/branchandcut/VRP/data/index.htm`. Accessed: 2016-03-20.

[47] G. Shafer. *A mathematical theory of evidence.* Princeton University Press, 1976.

[48] R. K. Srivastava, K. Deb, and R. Tulshyan. An evolutionary algorithm based approach to design optimization using evidence theory. *Journal of Mechanical Design*, 135(8):081003 – 081003–12, 2013.

[49] I. Sungur, F. Ordónez, and M. Dessouky. A robust optimization approach for the capacitated vehicle routing problem with demand uncertainty. *IIE Transactions*, 40:509–523, 2008.

[50] D. Teodorovic and G. Pavkovic. The fuzzy set theory approach to the vehicle routing problem when demand at nodes is uncertain. *Fuzzy Sets and Systems*, 82(3):307 – 317, 1996.

[51] P. Toth and D. Vigo. *The Vehicle Routing Problem*, chapter An Overview of Vehicle Routing Problems, pages 1–26. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2002.

[52] M. C. M. Troffaes. Decision making under uncertainty using imprecise probabilities. *International Journal of Approximate Reasoning*, 45(1):17–29, 2007.

[53] P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall/CRC Monographs on Statistics and Applied Probability. Taylor and Francis, 1991.

[54] L.A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 100(Supplement 1):9 – 34, 1999.

Appendix A. Proof of Proposition 1

Suppose $m_1^X \preceq m_2^X$. For any Q, $1 \leq Q \leq K$, we have then

$$
\begin{aligned}
Bel_1^X(x \in A_{1,Q}) &= \sum_{\overline{a} \leq Q} m_1^X(A) \\
&= \sum_{\overline{a} \leq Q} \sum \{R(A,B)m_2^X(B)|A \leq_{lo} B\} \\
&= \sum \{R(A,B)m_2^X(B)|A \leq_{lo} B, \overline{a} \leq Q\} \\
&= \sum \{R(A,B)m_2^X(B)|A \leq_{lo} B, \overline{a} \leq Q, \overline{b} \leq Q\} \\
&\quad + \sum \{R(A,B)m_2^X(B)|A \leq_{lo} B, \overline{a} \leq Q, \overline{b} > Q\} \\
&= \sum \{R(A,B)m_2^X(B)|A \leq_{lo} B, \overline{b} \leq Q\} \\
&\quad + \sum \{R(A,B)m_2^X(B)|A \leq_{lo} B, \overline{a} \leq Q, \overline{b} > Q\} \\
&= \sum_{\overline{b} \leq Q} m_2^X(B) \sum_{A \leq_{lo} B} R(A,B) + \sum_{\overline{b} > Q} m_2^X(B) \sum_{A \leq_{lo} B, \overline{a} \leq Q} R(A,B) \\
&= Bel_2^X(x \in A_{1,Q}) + \sum_{\overline{b} > Q} m_2^X(B) \sum_{A \leq_{lo} B, \overline{a} \leq Q} R(A,B) \\
&\geq Bel_2^X(x \in A_{1,Q}). \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad (A.1)
\end{aligned}
$$

The proof is similar to show that $Pl_1^X(x \in A_{1,Q}) \geq Pl_2^X(x \in A_{1,Q})$.

That the implication in Equation (22) is strict is indicated by the following counter-example. Let $X = \{x_1, \ldots, x_8\}$, $Q = 5$ and m_1^X and m_2^X be two mass functions defined as

$$m_1^X(\{x_3, x_4, x_5\}) = 0.8,\ m_1^X(\{x_1, x_3, x_4\}) = 0.2,$$
$$m_2^X(\{x_2, x_3, x_5\}) = 0.7,\ m_2^X(\{x_6, x_8\}) = 0.3.$$

We have

$$Bel_1^X(x \in A_{1,5}) = 1 \geq Bel_2^X(x \in A_{1,5}) = 0.7,$$
$$Pl_1^X(x \in A_{1,5}) = 1 \geq Pl_2^X(x \in A_{1,5}) = 0.7.$$

However, to have $m_1^X \preceq m_2^X$, it must be the case that the mass $m_2^X(\{x_2, x_3, x_5\}) = 0.7$ can be shared among the focal sets of m_1^X that are smaller (according to $\leq_{lo}$) than $\{x_2, x_3, x_5\}$, which is impossible since m_1^X has only one focal set ($\{x_1, x_3, x_4\}$) that is smaller than $\{x_2, x_3, x_5\}$ and this focal set has mass 0.2.

Appendix B. Proof of Proposition 2

Let us consider a set $\mathcal{C} = \{R_1, \ldots, R_m\}$ composed of m routes R_k, $k = 1, \ldots, m$, such that it is not known whether this set respects the belief-constraints (26) and (27), but it is known that it respects all the other constraints of the CVRPED-BCP.

It is clear that for any $\underline{\beta}$ and $\overline{\beta}$, as Q increases (starting from 1), it reaches necessarily a value at which constraints (26) and (27) are satisfied, and thus at which $\mathcal{C}$ becomes a solution to the CVRPED-BCP. Hence, for $Q' \geq Q$, the set of solutions to the CVRPED-BCP associated with value Q is included in or equal to the set of solutions to the CVRPED-BCP associated with value Q'.

Appendix C. Proof of Proposition 3

Let us consider a set $\mathcal{C} = \{R_1, \ldots, R_m\}$ composed of m routes R_k, $k = 1, \ldots, m$, such that it is not known whether this set respects the belief-constraints (26), but it is known that it respects all the other constraints of the CVRPED-BCP, in particular constraints (27).

It is clear that for any Q, as $\underline{\beta}$ increases from $\overline{\beta}$ to 1, it reaches necessarily a value at which constraints (26) are satisfied, and thus at which $\mathcal{C}$ becomes a solution to the CVRPED-BCP. Hence, for $\underline{\beta}' \geq \underline{\beta}$, the set of solutions to the CVRPED-BCP associated with value $\underline{\beta}$ is included in or equal to the set of solutions to the CVRPED-BCP associated with value $\underline{\beta}'$.

Appendix D. Proof of Proposition 5

Let R denote a route containing N clients. Without lack of generality, assume that the i-th client on R is the client i. Let $m_{\sum}^{\Theta_R}$ denote the MF defined on $\Theta_R := \{1, 2, \ldots, N \cdot Q\}$ and representing the sum of the customer demands

on R when the demand of client i is known in the form of MF m^{Θ_i}, and let $m_{\Sigma^+}^{\Theta_R}$ denote the MF representing the sum of the customer demands on R when the demand of client i is known in the form of MF $m_+^{\Theta_i}$.

Using a similar proof to that of [25, Proposition 3] (with operation $*$ instanciated to addition $+$, specialisation $\sqsubseteq$ replaced by ranking $\preceq$ and set inclusion $\subseteq$ replaced by lattice ordering $\leq_{lo}$), it is direct to show that $m^{\Theta_i} \preceq m_+^{\Theta_i}, i = 1, \ldots, N \Rightarrow m_{\Sigma}^{\Theta_R} \preceq m_{\Sigma^+}^{\Theta_R}$.

Let Bel and Bel_+ (resp. Pl and Pl_+) denote the belief functions (resp. plausibility functions) associated to $m_{\Sigma}^{\Theta_R}$ and $m_{\Sigma^+}^{\Theta_R}$, respectively. From Proposition 1, we have then

$$Bel(\sum_{i=1}^{N} d_i \leq Q) \geq Bel_+(\sum_{i=1}^{N} d_i \leq Q), \tag{D.1}$$

$$Pl(\sum_{i=1}^{N} d_i \leq Q) \geq Pl_+(\sum_{i=1}^{N} d_i \leq Q). \tag{D.2}$$

The proposition follows from the fact that Equations (D.1) and (D.2) hold for any route.

Appendix E. Proof of Proposition 6

Let $B_i \subseteq \Omega_i := \{0,1\}^{i-1}$, $i = 2, \ldots, N$, denote the set of possible failure situations that may occur at the i-th customer on route R, i.e.,

$$B_i = f(A_1 \times \cdots \times A_i), \tag{E.1}$$

with $A_\ell := [\![\underline{A_\ell}; \overline{A_\ell}]\!]$, $\ell = 1, \ldots, i$.

Let h_i be the function from Θ^i to $\mathbb{N}^*$ defined by $h_i(\theta_1, \ldots, \theta_i) = q_i$, with q_i defined by (44). In other words, h_i provides the load in the vehicle after serving the i-th customer given that customer demands are $(\theta_1, \ldots, \theta_i)$.

Remark that any $\omega^i \in B_i$ may be obtained by several vectors $(\theta_1, \ldots, \theta_i) \in A_1 \times \cdots \times A_i$. As a consequence, when it is known that the failure situation ω^i has occurred at the i-th customer, then the load in the vehicle after serving the i-th customer is known only in the form of a set L_{ω^i} such that

$$L_{\omega^i} = \left\{ h_i(\theta_1, \ldots, \theta_i) \mid \quad \forall (\theta_1, \ldots, \theta_i) \in A_1 \times \cdots \times A_i, f(\theta_1, \ldots, \theta_i) = \omega^i \right\}. \tag{E.2}$$

Consider the tree built according to Algorithm 1 and remove all its nodes below level i. Call $Tree_i$ the resulting tree. Then, for a given leaf of $Tree_i$, by concatenating in a vector the Boolean failure variable r_ℓ at level ℓ, $\ell = 2, \ldots, i$, written on the path from the root to the leaf, we obtain the binary failure situation vector $t^i = (r_2, r_3, \ldots, r_i) \in \Omega_i$ and this leaf contains also an interval LT_{t^i} of integers representing imprecise knowledge about the vehicle load after serving the i-th customer when t^i has occurred. Besides, all the leaves of $Tree_i$ yield the subset $BT_i \subseteq \Omega_i$.

We will now show by induction that for $i = 2, \ldots, N$, we have: $B_i = BT_i$ and $\forall \omega^i \in B_i$, $L_{\omega^i} = LT_{t^i}$ for $t^i \in BT_i$ such that $t^i = \omega^i$. Note that from the definition of the addition of two intervals of integers I_1 and I_2, i.e., $I_1 + I_2 = \{x_1 + x_2 | x_1 \in I_1, x_2 \in I_2\}$, we have $\forall x \in I_1 + I_2$, $\exists x_1 \in I_1, x_2 \in I_2$ such that $x_1 + x_2 = x$.

- Consider first the case $i = 2$, hence $\Omega_2 = \{\omega_1^2, \omega_2^2\}$ with $\omega_1^2 = (0)$ and $\omega_2^2 = (1)$. In such case, either $B_2 = \{\omega_1^2\}$ or $B_2 = \{\omega_2^2\}$ or $B_2 = \{\omega_1^2, \omega_2^2\}$.

  - If $B_2 = \{\omega_1^2\}$, then it implies that $\overline{A_1} + \overline{A_2} \leq Q$ and clearly $L_{\omega_1^2} = A_1 + A_2$. Besides, if $\overline{A_1} + \overline{A_2} \leq Q$, then according to Algorithm 1 we have $BT_2 = \{\omega_1^2\}$ and $LT_{\omega_1^2} = A_1 + A_2$.

  - If $B_2 = \{\omega_2^2\}$, then it implies that $\underline{A_1} + \underline{A_2} > Q$ and clearly $L_{\omega_2^2} = A_1 + A_2 - Q$. Besides, if $\underline{A_1} + \underline{A_2} > Q$, then according to Algorithm 1 we have $BT_2 = \{\omega_2^2\}$ and $LT_{\omega_2^2} = A_1 + A_2 - Q$.

  - If $B_2 = \{\omega_1^2, \omega_2^2\}$, then it implies that $\exists(\theta_1, \theta_2) \in A_1 \times A_2$ such that $f(\theta_1, \theta_2) = \omega_1^2$, and thus $\exists(\theta_1, \theta_2) \in A_1 \times A_2$ such that $\theta_1 + \theta_2 \leq Q$, and it also implies $\exists(\theta_1, \theta_2) \in A_1 \times A_2$ such that $f(\theta_1, \theta_2) = \omega_2^2$, and thus $\exists(\theta_1, \theta_2) \in A_1 \times A_2$ such that $\theta_1 + \theta_2 > Q$. In particular, it implies that $\underline{A_1} + \underline{A_2} \leq Q < \overline{A_1} + \overline{A_2}$. Hence, since for $\theta_1 + \theta_2 \leq Q$, we have $q_2 = \theta_1 + \theta_2$, and for $\theta_1 + \theta_2 > Q$, we have $q_2 = \theta_1 + \theta_2 - Q$, we obtain that $L_{\omega_1^2} = [\![\underline{A_1} + \underline{A_2}; Q]\!]$ and $L_{\omega_2^2} = [\![1; \overline{A_1} + \overline{A_2} - Q]\!]$. Besides, if $\underline{A_1} + \underline{A_2} \leq Q < \overline{A_1} + \overline{A_2}$, then according to Algorithm 1 we have $BT_2 = \{\omega_1^2, \omega_2^2\}$ and $LT_{\omega_1^2} = [\![\underline{A_1} + \underline{A_2}; Q]\!]$ and $LT_{\omega_2^2} = [\![1; \overline{A_1} + \overline{A_2} - Q]\!]$.

- Suppose that for $i < N$ we have: $B_i = BT_i$ and $\forall \omega^i \in B_i$, $L_{\omega^i} = LT_{t^i}$ for $t^i \in BT_i$ such that $t^i = \omega^i$. Let us show that it holds for $i + 1$.

  From the preceding assumption, we have $\forall \omega^i = (r_1^i, \ldots, r_i^i) \in B_i$ that L_{ω^i} is the interval LT_{t^i}, i.e., $L_{\omega^i} = [\![\underline{L_{\omega^i}}; \overline{L_{\omega^i}}]\!] = [\![\underline{LT_{t^i}}; \overline{LT_{t^i}}]\!]$ for $t^i \in BT_i$ such that $t^i = \omega^i$. In addition, we have $\forall \omega^i = (r_1^i, \ldots, r_i^i) \in B_i$:

  - Either $\overline{L_{\omega^i}} + \overline{A_{i+1}} \leq Q$, in which case the failure situation ω^i at the i-th customer will induce a failure situation $\omega^{i+1} \in B_{i+1}$ at the $i+1$-th customer such that $\omega^{i+1} = (r_1^i, \ldots, r_i^i, 0)$ and $L_{\omega^{i+1}} = L_{\omega^i} + A_{i+1}$. In addition, $\overline{L_{\omega^i}} + \overline{A_{i+1}} \leq Q$ is equivalent to $\overline{LT_{t^i}} + \overline{A_{i+1}} \leq Q$, in which case the leaf of $Tree_i$ associated to t^i will induce according to Algorithm 1 the leaf of $Tree_{i+1}$ with associated vector $t^{i+1} = \omega^{i+1}$ and interval $LT_{t^{i+1}} = LT_{t^i} + A_{i+1}$.

  - Or $\underline{L_{\omega^i}} + \underline{A_{i+1}} > Q$, in which case the failure situation ω^i at the i-th customer will induce a failure situation $\omega^{i+1} \in B_{i+1}$ at the $i + 1$-th customer such that $\omega^{i+1} = (r_1^i, \ldots, r_i^i, 1)$ and $L_{\omega^{i+1}} = L_{\omega^i} + A_{i+1} - Q$. In addition, $\underline{L_{\omega^i}} + \underline{A_{i+1}} > Q$ is equivalent to $\underline{LT_{t^i}} + \underline{A_{i+1}} > Q$, in which case the leaf of $Tree_i$ associated to t^i will induce according to Algorithm 1 the leaf of $Tree_{i+1}$ with associated vector $t^{i+1} = \omega^{i+1}$ and interval $LT_{t^{i+1}} = LT_{t^i} + A_{i+1} - Q$.

– Or $\underline{L_{\omega^i}} + \underline{A_{i+1}} \le Q < \overline{L_{\omega^i}} + \overline{A_{i+1}}$, in which case the failure situation ω^i at the i-th customer will induce a failure situation $\omega_L^{i+1} \in B_{i+1}$ at the $i+1$-th customer such that $\omega_L^{i+1} = (r_1^i, \ldots, r_i^i, 0)$ and $L_{\omega_L^{i+1}} = [\![\underline{L_{\omega^i}} + \underline{A_{i+1}}; Q]\!]$ since for $q_i + \theta_{i+1} \le Q$ we have $q_{i+1} = q_i + \theta_{i+1}$. It will also induce a failure situation $\omega_R^{i+1} \in B_{i+1}$ at the $i+1$-th customer such that $\omega_R^{i+1} = (r_1^i, \ldots, r_i^i, 1)$ and $L_{\omega_R^{i+1}} = [\![1; \overline{L_{\omega^i}} + \overline{A_{i+1}} - Q]\!]$ since for $q_i + \theta_{i+1} > Q$ we have $q_{i+1} = q_i + \theta_{i+1} - Q$. In addition, $\underline{L_{\omega^i}} + \underline{A_{i+1}} \le Q < \overline{L_{\omega^i}} + \overline{A_{i+1}}$ is equivalent to $\underline{L_{t^i}} + \underline{A_{i+1}} \le Q < \overline{L_{t^i}} + \overline{A_{i+1}}$, in which case the leaf of $Tree_i$ associated to t^i will induce according to Algorithm 1 the leaf of $Tree_{i+1}$ with associated vector $t_L^{i+1} = \omega_L^{i+1}$ and interval $LT_{t_L^{i+1}} = [\![\underline{LT_{t^i}} + \underline{A_{i+1}}; Q]\!]$. It will also induce the leaf of $Tree_{i+1}$ with associated vector $t_R^{i+1} = \omega_R^{i+1}$ and interval $LT_{t_R^{i+1}} = [\![1; \overline{LT_{t^i}} + \overline{A_{i+1}} - Q]\!]$.

Appendix F. Proof of Proposition 7

The proof of Proposition 7 relies on the following lemma.

Lemma 1. *Let m^Ω and m'^Ω be two MF representing uncertainty about the recourses on a given route R, such that $m^\Omega \sqsubseteq m'^\Omega$. Let $C_E^*(R)$ and $C_E'^*(R)$ denote the upper expected costs of R under m^Ω and m'^Ω, respectively. We have $C_E^*(R) \le C_E'^*(R)$.*

Proof. Let $C_P^*(R)$ and $C_P'^*(R)$ denote the upper expected penalty costs of some route R, under m^Ω and m'^Ω (denoted for simplicity m and m' in this proof), respectively.

We have

$$C_P'^*(R) = \sum_{B \subseteq \Omega} m'(B) \max_{\omega \in B} g(\omega), \tag{F.1}$$

and

$$C_P^*(R) = \sum_{A \subseteq \Omega} m(A) \max_{\omega \in A} g(\omega). \tag{F.2}$$

Since $m \sqsubseteq m'$,

$$
\begin{aligned}
C_P^*(R) &= \sum_{A \subseteq \Omega} \left(\left(\sum_{B \subseteq \Omega} S(A, B) m'(B) \right) \max_{\omega \in A} g(\omega) \right) \\
&= \sum_{B \subseteq \Omega} m'(B) \left(\sum_{A \subseteq \Omega} S(A, B) \max_{\omega \in A} g(\omega) \right).
\end{aligned}
\tag{F.3}
$$

Since $S(A, B) = 0, \forall A \not\subseteq B$, we can replace the condition of the second sum from $A \subseteq \Omega$ to $A \subseteq B$:

$$C_{\mathrm{P}}^*(R) = \sum_{B \subseteq \Omega} m'(B) \left(\sum_{A \subseteq B} S(A, B) \max_{\omega \in A} g(\omega) \right). \qquad \text{(F.4)}$$

In addition, for any $A, B \subseteq \Omega$ such that $A \subseteq B$, we have

$$\max_{\omega \in A} g(\omega) \;\; \leq \;\; \max_{\omega \in B} g(\omega),$$

hence

$$C_{\mathrm{P}}^*(R) = \sum_{B \subseteq \Omega} m'(B) \left(\sum_{A \subseteq B} S(A, B) \max_{\omega \in A} g(\omega) \right)$$

$$\leq \sum_{B \subseteq \Omega} m'(B) \left(\sum_{A \subseteq B} S(A, B) \max_{\omega \in B} g(\omega) \right)$$

$$= \sum_{B \subseteq \Omega} m'(B) \max_{\omega \in B} g(\omega)$$

$$= C_P'^*(R), \qquad \text{(F.5)}$$

where we used the fact that S is stochastic so that $\max_{\omega \in B} g(\omega) = \sum_{A \subseteq B} S(A, B) \max_{\omega \in B} g(\omega)$

for any $B \subseteq \Omega$. Using $C_{\mathrm{P}}'^*(R) \geq C_{\mathrm{P}}^*(R)$ we obtain

$$C_{\mathrm{P}}'^*(R) + C(R) \geq C_{\mathrm{P}}^*(R) + C(R) ,$$

which means

$$C_{\mathrm{E}}'^*(R) \geq C_{\mathrm{E}}^*(R) .$$

$\square$

Proposition 7 may then be proved as follows.

Let R denote a route containing N clients. Without lack of generality, assume that the i-th client on R is the client i. Let m^Ω denote the MF representing uncertainty about recourses on R when the demand of client i, $i = 1, \ldots, N$ is known in the form of MF m^{Θ_i}, and let $m_\star^\Omega$ denote the MF representing uncertainty about recourses on R when the demand of client i is known in the form of MF $m_\star^{\Theta_i}$.

Using a similar proof to that of [25, Proposition 3] (with operation $*$ replaced by function f defined in Section 3.2.2), it is direct to show that $m^{\Theta_i} \sqsubseteq m_\star^{\Theta_i}, i = 1, \ldots, n \Rightarrow m^\Omega \sqsubseteq m_\star^\Omega$. From Lemma 1, we obtain then $C_{\mathrm{E}}^*(R) \leq C_{\mathrm{E}}^{\star*}(R)$.

Considering that the optimal solution with mass functions $m_\star^{\Theta_i}$ consists of a set S of routes $\{R_1, \ldots, R_m\}$, we have then $C_{\mathrm{E}}^{\star*}(R_k) \geq C_{\mathrm{E}}^*(R_k)$ for each $k \in \{1, \ldots, m\}$, which yields

$$\hat{C}_{Rec}^\star = \sum_{k=1}^m C_{\mathrm{E}}^{\star*}(R_k) \geq \sum_{k=1}^m C_{\mathrm{E}}^*(R_k) \geq \hat{C}_{Rec}. \qquad \text{(F.6)}$$

Appendix G. The simulated annealing algorithm for the CVRPED

The pseudo-code of our simulated annealing algorithm is provided by Algorithm 2. It starts by generating an initial configuration (candidate solution) $\mathcal{C}$ using the `initial_config(...)` routine, when the initial temperature of the system T is at its highest value. Afterwards, T is progressively decreased until reaching the freezing temperature $freez$, while a sequence of iterations tot_iter are performed for each T. Throughout each iteration $iter$, a neighbourhood configuration C^* of the current configuration $\mathcal{C}$ is generated, and the variation in the cost Δ_{cost} is computed. In other words, each configuration represents an intermediate solution that has a different cost which is computed using the cost method, and Δ_{cost} is equal to the difference between the new cost of the neighbourhood configuration C^*_{cost} and the current cost of the current configuration $\mathcal{C}_{cost}$. If the cost decreases then the move to the new cost is accepted (lines $16-18$).

However, if Δ_{cost} is positive then the move is accepted or rejected with a probability that equals $e^{-\frac{\Delta_{cost}}{T}}$. Effectively, the probability of accepting inferior solutions is a function of the temperature T and the change in cost Δ_{cost}. We repeat this whole process for a total number of trials tot_tr and the algorithm finally returns the best configuration ever visited.

The set of parameters controlling Algorithm 2 were experimentally determined using Augerat set A instances of the CVRP [46], in which vertices coordinates were randomly constructed [3]. Specifically, the initial temperature T was set to 5000 and was decreased by a temperature reduction multiplier κ that was set to 0.82 until reaching a freezing temperature $freez$ that equals 1. The total number of iterations tot_iter was regulated to 30000, while the total number of trials of the algorithm tot_tr was determined to be 5. The results with our algorithm varied between 1% and 12% from the optimal solutions of the CVRP, with an average running time under 30 minutes, for all instances. This algorithm is an adaptation of the algorithm introduced in [32] for the CVRP.

Let us remark that a configuration in Algorithm 2 is a set of routes that can be generated either by the `initial_config(...)` or the `neighbourhood_configuration(...)` routines. Besides, the `cost(...)` routine evaluates the objective value of the configuration. All these routines depend on a "modelling technique parameter" MD that specifies whether we are solving the CVPRED-BCP or the CVRPED-recourse model. In particular, if $MD = \text{``}BCP''$ the `cost(...)` routine evaluates the objective value of the configuration in the BCP model. The objective function is thus the (classical) total travelled distance of the routes (1). If $MD=$"recourse", the `cost(...)` routine corresponds to using the recourse objective function, that aims at minimising the total upper expected cost (42). The descriptions of the routines `initial_config(...)` and `neighbourhood _configuration(...)` given the value of parameter MD, are provided in Section 4.1.

Algorithm 2 Simulated annealing algorithm

Input: initial temperature T, temperature reduction multiplier κ, freezing temperature $freez$, total number of iterations tot_iter, total number of trials tot_tr, modelling technique MD (BCP or recourse)

Output: Best solution ever visited $Best_C$

1: $Best_{cost} = \infty$
2: **for** $tr = 0$ **to** tot_tr **do**
3:    **if** $tr == 0$ **then**
4:       $\mathcal{C} = $ initial_config(greedy, MD)   ▷ greedy generation
5:    **else**
6:       $\mathcal{C} = $ initial_config(random, MD)   ▷ random generation
7:    **end if**
8:    $\mathcal{C}_{cost} = $cost($\mathcal{C}$, MD)
9:    $TBest_C = \mathcal{C}$
10:    $TBest_{cost} = \mathcal{C}_{cost}$
11:    **repeat**
12:       **for** $iter = 0$ **to** tot_iter **do**
13:          $C^* = $ neighbourhood_configuration($\mathcal{C}$, MD)
14:          $C^*_{cost} = $ cost(C^*, MD)
15:          $\Delta_{cost} = C^*_{cost} - \mathcal{C}_{cost}$
16:          **if** $(\Delta_{cost} < 0)$ **then**
17:             $\mathcal{C} = C^*$
18:             $\mathcal{C}_{cost} = C^*_{cost}$
19:             **if** $C^*_{cost} < TBest_{cost}$ **then**
20:                $TBest_C = C^*$
21:                $TBest_{cost} = C^*_{cost}$
22:             **end if**
23:          **else if** $rnd \leq e^{-\frac{\Delta_{cost}}{T}}$ **then**   ▷ rnd is a random number in $[0,1]$
24:             $\mathcal{C} = C^*$
25:             $\mathcal{C}_{cost} = C^*_{cost}$
26:          **end if**
27:       **end for**
28:       $T = \kappa \cdot T$
29:    **until** $(T == freez)$
30:    **if** $(TBest_{cost} < Best_{cost})$ **then**
31:       $Best_C = TBest_C$
32:       $Best_{cost} = TBest_{cost}$
33:    **end if**
34: **end for**

Appendix H. Examples for the neighbourhood operators Fix_minimum and Replace_highest_average

Example 7. *(Fix_minimum) Suppose configuration $\mathcal{C}$ consists of the set of three routes $\mathcal{C} = \{(0, 3, 6, 10, 0), (0, 1, 5, 8, 4, 9, 7, 0), (0, 2, 0)\}$. Assume that the smallest distances between consecutive clients are those between clients $< 3, 6 >$, $< 6, 10 >$, $< 1, 5 >$, $< 4, 9 >$ and $< 7, 0 >$. This means that customers $3, 6, 1, 4$ and 7 cannot be removed from their routes by the **fix_minimum** operator. Consequently, **fix_minimum** selects five random customers excluding $3, 6, 1, 4, 7$ and the depot. For instance customer 5 is selected and removed from the second route in $\mathcal{C}$ and inserted randomly in one of the three available routes in $\mathcal{C}$, while respecting all problem constraints. After selecting the new route for client 5, it is inserted at the position with the resulting smallest distance to client 5. We repeat the same process to move the four other random customers which in this example cannot be other than customers $2, 5, 8, 9$ and 10.*

Example 8. *(Replace_highest_average) Suppose a configuration $\mathcal{C}$ that consists of the set of routes $\mathcal{C} = \{(0, 1, 3, 8, 0), (0, 2, 5, 4, 9, 10, 0), (0, 6, 7, 0)\}$. Suppose 1, 8, 5, 9 and 7 are the clients having the five highest average distances (separating each one of these clients from its neighbours). Then, these clients are removed from their routes, and we will have $\mathcal{C} = \{(0, 3, 0), (0, 2, 4, 10, 0), (0, 6, 0)\}$. Afterwards, each removed client is inserted randomly in one of the available routes. The position where to insert a customer is chosen, such that: i) the problem constraints are respected; and ii) the new position of that customer on the chosen route has the smallest average distance, if compared to all other possible positions of this client on the chosen route. For instance, suppose the first route in $\mathcal{C}$ was chosen randomly for client 1. We know that inserting client 1 on that route does not violate the belief-constraints, and it can be inserted either before client 3 or right after client 3. Suppose also that inserting client 1 right after client 3 has a smallest average distance, than if client 1 was inserted right before client 3. Then, client 1 is inserted right after client 3. This same process is repeated for the remaining clients 8, 5, 9 and 7.*