Les Cahiers du GERAD ISSN: 0711–2440

Distributionally robust optimization for the multi-period multi-item lot-sizing problems under yield uncertainty

P. Metzker, S. Thevenin, Y. Adulyasak, A. Dolgui

G–2024–12

January 2024

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2024
– Bibliothèque et Archives Canada, 2024

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2024
– Library and Archives Canada, 2024

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Distributionally robust optimization for the multi-period multi-item lot-sizing problems under yield uncertainty

Paula Metzker [a, b]

Simon Thevenin [c]

Yossiri Adulyasak [a, b]

Alexandre Dolgui [c]

[a] *Department of Logistics and Operations Management, HEC Montréal, Montréal (Qc), Canada, H3T 2A7*

[b] *GERAD, Montréal (Qc), Canada, H3T 1J4*

[c] *IMT Atlantique, LS2N - UMR CNRS 6004, F-44307 Nantes, France*

paula.metzker-soares@hec.ca
simon.thevenin@imt-atlantique.fr
yossiri.adulyasak@hec.ca
alexandre.dolgui@imt-atlantique.fr

January 2024
Les Cahiers du GERAD
G–2024–12

Abstract : Yield uncertainty is an important issue in various industries such as agriculture, food, and textile where the production output is reliant on uncontrollable factors and fluctuating raw material quality. To systematically leverage data to deal with uncertainty in a cost-effective fashion, distributionally robust optimization combines the strengths of stochastic programming and robust optimization by optimizing the expected costs against an ambiguity set that defines possible distributions. In this work, we leverage a data-driven robust optimization framework and formulate a mixed-integer distributionally robust multi-item lot-sizing model with uncertain production yield to determine a robust production plan. To this end, we use a scenario-wise formulation that partitions the available data into scenarios that define different patterns influencing the quality of the product and production process. In addition, we apply the proposed approach to real-world data of a case study to demonstrate the effectiveness of the proposed framework in dealing with yield uncertainty. Our experimental results show that distributionally robust plans lead to more effective cost-saving strategies and decreased risk of stock-outs. Additionally, our findings suggest that the proposed model exhibits lower sensitivity to variations in production yield realizations and it is more proficient in incorporating historical data into the decision-making process. This results in a more effective response to challenges encountered within the production system under yield uncertainty.

Keywords : Lot-sizing, production planning, yield uncertainty, distributionally robust optimization.

Résumé : L'incertitude des rendements de production est un problème important dans diverses industries telles que l'agriculture, l'alimentation et le textile, où la production dépend de facteurs incontrôlables et de la qualité fluctuante des matières premières. Pour exploiter systématiquement les données afin de gérer l'incertitude de manière rentable, l'optimisation distributionnellement robuste combine les atouts de la programmation stochastique et de l'optimisation robuste en optimisant les coûts attendus par rapport à un ensemble d'ambiguïtés qui définit les distributions possibles. Dans ce travail, nous exploitons un cadre d'optimisation robuste basé sur les données et formulons un modèle de dimensionnement de lots multi-produits robuste sur le plan de la distribution et avec un rendement de production incertain pour déterminer un plan de production robuste. À cette fin, nous utilisons une formulation par scénarios qui divise les données disponibles en scénarios définissant différents modèles influençant la qualité du produit et du processus de production. De plus, nous appliquons l'approche proposée aux données réeles d'une étude de cas pour démontrer l'efficacité du cadre proposé pour traiter l'incertitude du rendement. Nos résultats expérimentaux montrent que des plans de distribution robustes conduisent à des stratégies de réduction des coûts plus efficaces et à une diminution du risque de rupture de stock. De plus, nos résultats suggèrent que le modèle proposé présente une sensibilité moindre aux variations des rendements de production et qu'il est plus efficace dans l'intégration des données historiques dans le processus de prise de décision. Il en résulte une réponse plus efficace aux défis rencontrés au sein du système productif dans un contexte d'incertitude sur les rendements de production.

Mots clés : Dimensionnement des lots, planification de la production, incertitude de rendement, optimisation distributionnellement robuste.

Acknowledgements: The authors wish to thank the Region Pays de la Loire (www.paysdelaloire.fr) and the Canada Research Chair in Supply Chain Analytics for the financial support of this research. The data used was provided by the ASSISTANT project (https://assistant-project.eu/) funded by the European Commission, under grant number 101000165, H2020–ICT38–2020, Artificial Intelligence for Industry.

1 Introduction

In today's fiercely competitive business environment, manufacturers across different industries, especially those where yield uncertainty is high, are struggling to make the most of their resources and production capabilities to ensure that they can meet the demand for high-quality products in a market that's constantly changing and volatile while keeping costs as low as possible. In such a situation, manufacturers seek to enhance their processes to improve quality and reduce waste. Despite advancements in technology for quality improvements and defect detection, many production environments still face challenges with high defect rates. This is particularly true for industries characterized by complex production systems, products with stringent specifications, or highly manual production processes. The quality of production in such industries is often influenced by external factors like climate changes, supply disruptions or delays, and fluctuations in raw material quality, as well as internal factors such as machinery maintenance, workforce availability and skill, and failures in the production process flow. As a result, the production yield rate becomes highly unpredictable in these industries, directly impacting the quantity of high-quality products obtained from production.

Production yield is a measure of a production system's performance and efficiency in meeting the quality specifications of the products. An accurate production yield estimation helps manufacturers to maintain the system under control, reducing its vulnerability to errors, malfunctions, irregularities, and loss of profit. The production yield is often estimated from the historical data and quality control data associated with the production process. However, these predictions are often inaccurate because of the complexity and various uncontrollable factors that can affect the quality of the production. Therefore, the losses in production quality are difficult to predict and measure, and their impacts can be highly damaging to the system [22]. For instance, in the olive oil industry, yield uncertainty can increase overall costs as low production yields force producers to buy olives from other farmers to fulfill demand and meet the contractual obligation [23]. Yield uncertainty can also affect the contractual arrangements in supply chains. If the production yield is low, it can result in increased insurance payouts, reduced capital investment, and difficulties in obtaining credit and subsidies [2]. Yield uncertainty in vaccine production can also have a direct impact on social welfare as it may result in weak immunity to new virus strains or mutations [13]. Therefore, optimization approaches that effectively and systematically hedge against such uncertainties are of utmost importance.

When faced with high levels of uncertainty, two commonly used approaches are stochastic programming (SP) and robust optimization (RO). Generally, SP optimizes based on the expected value when uncertainties follow a known probability distribution [6]. Traditional SP methods rely on a reliable estimate of the probability distribution of uncertain parameters. In practice, decision-makers in industry often struggle to provide accurate estimates for the probability distribution of the production yield due to a lack of quality data and their dependence on various factors such as weather conditions, operators, and raw material quality. RO is an alternative approach that does not rely on precise estimation of the underlying probability distribution. RO optimizes against the worst-case scenario within an uncertainty set, which can sometimes result in overly cautious and expensive solutions [3]. However, traditional robust optimization methods often use simplistic uncertainty representations which can result in solutions that are conservative and costly to implement. To address this issue, distributionally robust optimization (DRO) provides a robust perspective for stochastic programming [14]. More specifically, DRO incorporates the probabilistic concepts from SP and the robust perspective from RO [50]. DRO immunizes the system from the worst-case probability distribution described by an ambiguity set, which defines a family of probability distributions of the uncertain parameter. An ambiguity set can be created using partial stochastic information estimated from historical data [14]. As DRO is a recent optimization methodology, some challenges remain to be addressed. More specifically, there are restrictive assumptions that must be made to compute a tractable solution, the reformulation can be complex and affects the intractability of models [50]. Recently, [11] introduce an event-wise ambiguity set, and the authors provide an efficient reformulation for this ambiguity set. This formulation combines the scenario-tree method from SP with the bounded representation of uncertain parameters from RO. The event-wise formulation can be used to represent well-known ambiguity sets including generalized moment, mean absolute, Wasserstein distance and Wasserstein set, and clustering-based ambiguity set by making use of an additional random variable to generalize the definition of the uncertain parameter [11].

The paper introduces a distributionally robust model to address the multi-item lot-sizing problem (LSP) when dealing with yield uncertainty. The LSP aims to determine production setups and quantities to meet demand with quality goods while minimizing overall production and inventory costs [32]. Our contribution is fourfold. First, we propose a distributionally robust formulation for the multi-item LSP under production yield uncertainty. To this end, we leverage the scenario-wise ambiguity set based on the framework proposed by [11]. Second, using the framework presented in [48], we reformulate the distributionally robust model as a mixed integer linear program (MILP) that can be efficiently solved by commercial solvers. Third, through simulations, we compare the quality of the production plans obtained from the distributionally robust model with the production plans obtained from traditional stochastic programming and robust optimization models. Finally, we empirically validate the value of the robust solution using data from a case study to highlight the advantages of using DRO to deal with production yield uncertainty.

This work is organized as follows. Section 2 reviews the literature on distributionally robust optimization with a specific focus on its applications to lot-sizing problems and other related problems. Section 3 formally outlines the problem we are considering and introduces the scenario-wise distributionally robust optimization methodology as well as its MILP reformulations. Section 4 presents the experimental results on the quality of distributionally robust plans. Finally, Section 5 concludes this work and suggests some future research directions.

2 Literature review

This section first presents briefly the literature on lot-sizing under uncertainty, then the stochastic and robust studies on lot-sizing problems (LSPs) under yield uncertainty. Finally, we review the distributionally robust optimization theory and its application to the multi-item LSPs.

2.1 LSPs under yield uncertainty

The classical LSP has been first introduced in the 1950s by [47]. LSPs have attracted a wide range of research from management, production, operations, and mathematical optimization communities. We refer readers to [8] for the recent state-of-the-art methods on single-item LSPs, and to [33] for further information on LSPs. Although the literature on LSPs is vast, there are still some knowledge gaps on non-deterministic formulations. [1] and [8] indicate the predominance of studies on uncertain demand, whereas other uncertain parameters such as lead time, cost, capacity, and production yield, which can significantly affect the quality and efficiency of the manufacturing process and production plans, are not widely studied. Finally, a review of literature on lot-sizing problems under uncertain yield is presented by [51].

Non-deterministic multi-item LSP literature dates back to the 1990s, when [10] present a simulation approach to define production plans. [9] propose a MILP formulation to address capacitated LSP with uncertain demand, and the authors suggest a fix-and-relax heuristic strategy to determine a solution. [43] develops a column generation heuristic to solve a stochastic LSP under demand uncertainty. Since then, a large number of publications have considered non-deterministic LSPs, and there is abundant literature on stochastic programming and robust optimization for LSPs. Most of the studies consider the demand is unknown [44, 45, 27], and a different stream of research considers lead time uncertainties [46, 41]. However, other parameters may be unknown such as production capacity, process duration, and production yield, and they are only very few studies on the uncertainty of these parameters [39].

Research on production planning under yield uncertainty is rather scarce. [24] propose a stochastic program to handle yield uncertainty due to the quality of raw materials for a sawmill production plan. [35] propose a multi-item multi-echelon LSP for a manufacturing system that includes disassembly, refurnishing, and reassembly, all while dealing with uncertainties. As the stochastic MILP model is formulated with a scenario tree, the resulting model is very complex and suffers from a scalability issue. To overcome this drawback, [35] develop a branch-and-cut algorithm to compute good solutions within a reasonable timeframe. [31] propose a robust optimization model for the single item LSP under yield uncertainty. The authors investigate the robust formulation of the problem, and they propose a MILP model and a dynamic program to solve it. Then, [42] expends the previous work to the LSP where production decisions are adjustable to the realizations of yield uncertainty. The authors propose a MILP approximation and a column-and-constraint generation algorithm to compute an optimal adaptive robust production plan.

The literature on multi-item LSPs under yield uncertainty is in its early stages of development. To the best of our knowledge, no studies have yet explored the multi-item LSP under yield uncertainty, and this work aims to fill this knowledge gap via the distributionally robust optimization. This paper extends the lot-sizing model presented in [31, 42] to the multi-item variant, and we investigate the use of the distributionally robust optimization framework to enhance the quality of the solution in this challenging context.

2.2 Distributionally robust optimization

Although research on robust optimization has been first introduced in the 1950s with [38]'s work, [14] are among the first to formally establish a distributionally robust optimization framework. They did so by defining an ambiguity set based on moment information and proposing a semi-infinite programming method to solve it. In a later study, [50] provide a comprehensive review of convex DRO methodology, and they introduce standardized forms for convex ambiguity sets. The authors also outline the conditions which ensure the tractability of the distributionally robust models. [15] introduce the Wasserstein ambiguity set, which is both convex and allows for a tractable reformulation. [26] expand on previous research and present additional theory on the Wasserstein ambiguity set. Their work highlights the advantages of this approach both conceptually and computationally. They also demonstrate promising outcomes when utilizing the Wasserstein ambiguity set in conjunction with machine learning models. [5] propose a tractable adaptive DRO formulation based on second-order conic representable ambiguity sets. The authors provide tools to reformulate the distributionally robust models as an MILP that can be easily solved with commercial solvers. In the same vein, [11] introduce an event-wise formulation for the ambiguity sets, and provide a new modeling package, called RSOME, to help modelers reformulate and solve distributionally robust models. For an extensive review of DRO, we also refer interested readers to [36].

For the sake of tractability, distributionally robust models are often approximated using conic representations like second-order programs or linear reformulation [12]. These models take into account available distributional information to compute the expectation of the uncertainty either through an expected value formulation or in a chance constraint form [50]. For the latter, we can refer to the work of [30] where the authors propose a risk-averse perspective for a distributional robust model based on conditional value-at-risk (CVaR) constraints for a disassembly line balancing problem. [30] reformulate the CVaR constraints as second-order cone constraints and propose a cutting-plane algorithm to solve the reformulated robust model based on a reduced ambiguity set. [16] present a two-stage distributionally robust model for the COVID-19 testing facility territory design and capacity planning problem. The authors then develop a tractable reformulation and an adversarial approach to compute a robust solution. In the same spirit, [49] apply a second-order conic (SOC) programming approach to a two-stage distributionally robust model. The authors propose an approximation of the hard SOC model to determine a feasible robust solution.

[14] propose an ambiguity set constructed from the knowledge of the support of the moment of the distribution. [7] suggest that, when dealing with processes that involve random variables in a dynamic environment (such as the production yield rate that changes according to the process condition), it is important to consider a larger ambiguity radius and confidence region (such as the Wasserstein distance) to reduce disturbances in cases of errors or misestimation. The authors indicate that a large set of scenarios is necessary to accurately estimate the occurrence of time-varying random variables. [18] state that the Wasserstein ambiguity set can be used to effectively prevent disturbance and reduce the impacts caused by inaccurate distribution estimates or insufficient data. [37] propose an adversarial approach to incorporate estimations of the uncertainty distribution in a Wasserstein ambiguity set. The authors also discuss approaches to estimate these distributions in situations where they may vary over time. Finally, [29] report the effectiveness of the Wasserstein ambiguity set with a finite and small sample set. The authors demonstrate that this model delivers reliable out-of-sample performance and greater robustness compared to other models that do not consider distribution information.

The concept of distributionally robust optimization for inventory management problems has been studied since the 2010s. [19] propose a risk-averse distributionally robust multi-item newsvendor problem under uncertain demand. The authors approximate the distributionally robust model to a quadratic programming model, which yields a conservative but tractable formulation. Meanwhile, [21] present a reformulation approach based on the conditional value-at-risk (CVAR) to solve the multi-product assembly and the portfolio selection problem. They use a cutting plane algorithm to solve this robust model. [34] develop a reformulation by utilizing Lagrange multipliers and propose decomposition methods to solve the multi-product inventory problem under demand and supply uncertainties. Finally, [48] reformulate the multi-item newsvendor problem using event-wise affine decision rules and propose a column generation algorithm to determine the solution. Distributionally robust optimization has been applied to lot-sizing problems in recent years. However, the focus of the studies has been mostly on demand uncertainty. [17] present a distributionally robust model to mitigate prediction errors from demand prediction models for a two-stage lot-sizing problem. [17] also present adversarial approaches to handle demand uncertainty in this lot-sizing problem.

Our work differs from the aforementioned literature in various ways. To the best of our knowledge, this paper is the first to tackle the multi-item LSP under yield uncertainty. Since various operational conditions (such as changes in raw material quality and ambient temperature) can affect production yield, we adopted an event-wise ambiguity set in a DRO framework. This approach leads to a less conservative plan since one can incorporate available data and scenario analysis into the decision-making process.

Notation : We use boldface (e.g., $\mathbf{x}$) to denote vectors. The following definitions are also used in this paper :
a) The support function of a convex set $\mathcal{Q}$ describing a variable z is defined as $\delta^*(z|\mathcal{Q}) = \sup_{\boldsymbol{\xi} \in \mathcal{Q}} \boldsymbol{\xi}^\top \boldsymbol{z}$.
b) The perspective of a function $g(x) : \mathbb{R}^n \Rightarrow \mathbb{R}$ is a function $g(x,t) : \mathbb{R}^n \times \mathbb{R} \Rightarrow \mathbb{R} = tg(\frac{x}{t}),\ \forall t > 0$. The perspective of a function is a mathematical concept that often helps modelers provide non-intuitive behaviors or to demonstrate some properties associated with optimization problems [20].
c) The conjugate of a function $g(y)$ is a convex function $g^*(y)$, even if $g(x)$ is not convex. The conjugate is given by $g^*(y) = \sup_{x \in \boldsymbol{dom}\, g(x)} y^\top x - g(x)$, where $\boldsymbol{dom}$, the domain of a function, gives all the inputs to the function.

3 Distributionally robust multi-item lot-sizing problem under yield uncertainty

We consider a multi-item multi-period lot-sizing problem (LSP) with backorders and uncertain production yield. This is an extension of the single-item problem presented in [31]. We aim to determine the optimal production setups and lot sizes for a given set of items I within a finite planning horizon T.

For each item $i \in N = \{1, ..., |N|\}$ and each period $t \in T = \{1, ..., |T|\}$, the following parameters are given : setup cost s_{it}, unit production cost v_{it}, unit inventory cost h_{it}, unit backorder cost b_{it}, and demand d_{it}. The production yield $\tilde{\rho}_{it}$ for item i in period t is uncertain, and it is strictly positive ($0 < \tilde{\rho}_{it} \leq 1$). In addition, $M_{it} = \frac{\sum_{\tau \in T} d_{i\tau}}{\min_{\tau \leq t} \rho_\tau}$ is an upper bound on the production quantity. The LSP requires finding the production quantity X_{it} and the setup decision Y_{it} to meet demands and minimize costs. In the distributionally robust optimization paradigm, these decisions are made to minimize the worst-case expected inventory management cost concerning a set of probability distributions $\mathscr{F}$. We denote by $H_{it}(\tilde{\boldsymbol{\rho}})$ the inventory or backorder cost for item i and period t depending on the uncertain yield $\tilde{\boldsymbol{\rho}}$. We assume that the decision-maker follows a static strategy, where the production plan is decided before observing the realization of the uncertain production yield, and the decisions remain fixed over the entire planning horizon. The distributionally robust multi-item lot-sizing uncapacitated problem (DRLSP) can be formulated as follows :

$$\min \left(\sum_{i \in N} \sum_{t \in T} (s_{it} Y_{it} + v_{it} X_{it}) + \sup_{\mathbb{P} \in \mathscr{F}} \mathbb{E}_{\mathbb{P}} \Big[\sum_{i \in N} \sum_{t \in T} H_{it}(\tilde{\rho}_{it}) \Big] \right) \tag{1a}$$

$s.t.$

$$H_{it}(\tilde{\boldsymbol{\rho}}) \geq h_{it} \left[\sum_{\tau=1}^{t} (\tilde{\rho}_{i\tau} X_{i\tau} - d_{i\tau}) \right] \qquad \forall i \in N; t \in T; \tilde{\boldsymbol{\rho}} \in \mathbb{P} \tag{1b}$$

$$H_{it}(\tilde{\boldsymbol{\rho}}) \geq -b_{it} \left[\sum_{\tau=1}^{it} (\tilde{\rho}_{i\tau} X_{i\tau} - d_{i\tau}) \right] \qquad \forall i \in N; t \in T; \tilde{\boldsymbol{\rho}} \in \mathbb{P} \tag{1c}$$

$$X_{it} \leq M_{it} Y_{it} \qquad \forall i \in N; t \in T \tag{1d}$$

$$X_{it}, H_{it} \geq 0 \qquad \forall i \in N; t \in T \tag{1e}$$

$$Y_{it} \in \{0, 1\} \qquad \forall i \in N; t \in T \tag{1f}$$

where $\mathcal{F}$ is the ambiguity set which represents a family of the distributions, $\mathbb{P}$ is a distribution realized from this set, and $\tilde{\boldsymbol{\rho}}$ is the uncertain yield based on the distribution $\mathbb{P}$. Without a loss of generality, we assume that there is no stock or backorder at the beginning of the planning horizon. The objective function (1a) minimizes the total cost comprising the setup, unit production, and worst-case expected inventory and backorder costs. The model determines the inventory management and backlog costs $H_{it}(\rho)$ which account for the production yield. More specifically, constraints (1b) compute the inventory cost for item i in period t with the help of the cumulative amount of quality goods obtained for periods up to t. Similarly, constraints (1c) compute the backorder cost if the cumulative amount of quality goods is not enough to meet all demands up to t. The constraints (1d) are setup-forcing constraints that relate the lot sizes (X_{it}) to the setup decisions (Y_{it}). These constraints set the setup variable Y_{it} to 1 if any production for item i occurs in period t, and the setup remains inactive otherwise ($Y_{it} = 0$). To address the capacitated version of the problem, constraints (1d) should be modified to represent the resource availability by setting $M_{it} = \min\{C_{it}, M_{it}\}$, where C_{it} is the available capacity for item i in period t.

To obtain the optimal solution to the deterministic uncapacitated lot-sizing problem, one can solve the single-item model separately for each item. However, in the distributionally robust lot-sizing problem, an additional assumption is required, namely, that the sub-problem of finding the worst-case distribution is also separable per item. This means that the ambiguity set does not include constraints that link the distributions of different items.

The $\sup_{\mathbb{P} \in \mathcal{F}} \mathbb{E}_{\mathbb{P}}[.]$ represents the worst case expected cost over the ambiguity set $\mathcal{F}$. To solve problem (1) with a MILP solver, we transform it into a MILP distributionally robust model for LSP under yield uncertainty. This section provides our MILP reformulation. First, we define the scenario-wise ambiguity set. Second, we define the scenario-wise affine decision rule that computes the inventory management costs as a function of the realization of the random production yield. Finally, we rely on Slater's condition and Sion's minimax theorem to reformulate our DRO model as a robust model [11, 48].

3.1 Definition of the scenario-wise ambiguity set

The event-wise ambiguity set was first introduced by [11], and it incorporates the scenario tree representation from stochastic optimization with the affine decision rules to represent the uncertainties in the ambiguity set. For a formal definition of the event-wise ambiguity set and further information on its application, we refer the interested readers to [11].

As we consider a static decision framework, where all decisions are fixed before the realization of uncertain yield, the event ambiguity set has only one event with several scenarios that describe the uncertainty. In the static decision framework, the realization of the production yield impacts only the inventory management costs. Therefore, we reduced the event-wise ambiguity set to a scenario-wise ambiguity set where each scenario is a pattern that describes one possible behavior of the uncertain yield.

Let us assume a historical data set $\mathcal{H}$ of the production yield, which contains various measurements of production performance PP from the past (e.g., $\mathcal{H} = \{PP_1, PP_2, ..., PP_8\}$, where 8 production performances are available). From $\mathcal{H}$, we can identify different scenarios s that influence the production yield. For instance, scenarios can represent changes in temperature and changes in raw material. In this case, a scenario s_1 (denoted [normal, S1]) corresponds to the case where the ambient temperature is normal, and raw materials come from supplier S1. We have $s_2 = [low, A]$, $s_3 = [normal, B]$, $s_4 = [low, B]$. Therefore, we can partition $\mathcal{H}$ in S exclusive scenarios that contain at least one measurement each (e.g., $S\{s_1, s_2, s_3, s_4\}$, and $s_1 = \{PP_1, PP_6, PP_7\}$, $s_2 = \{PP_2, PP_4, PP_5\}$, $s_3 = \{PP_3, PP_8\}$, $s_4 = \{PP_2, PP_6, PP_8\}$, such that $s_1 \cup s_2 \cup s_3 \cup s_4 = S \subset \mathcal{H}$). Each scenario helps us to estimate the true value of our random variable (here the production yield $\tilde{\boldsymbol{\rho}}$). More precisely, from $\mathcal{H}$ we obtain a set $S = s_1, .., s_S$ of scenarios, where each scenario s represents a conditional moment information that the uncertain production yield follows.

A confidence set of distribution gives the range of values compatible with the data estimated with the distribution considered. To account for confidence sets, [50] redefines the random variable with the inclusion of an additional variable $\boldsymbol{m}$. This additional variable links the outcomes from different probability distributions with their respective confidence sets without imposing a condition on the confidence set [11]. Therefore, the scenario-wise model remains valid for different settings and definitions of the ambiguity set.

We redefine the random production yield as $(\tilde{\boldsymbol{\rho}}, \tilde{\boldsymbol{m}})$, where the primary random variablec $\tilde{\boldsymbol{\rho}} \in \mathbb{R}^{(N\times T)}$ gives the uncertain production yield, and the auxiliary random variable $\tilde{\boldsymbol{m}} \in \mathbb{R}^{(N\times T)_m}$ ensure that the scenario-wise model remains valid for different probability distributions without imposing additional or specific conditions on the confidence set. The space $\mathbb{R}^{(N\times T)_m}$ (resp. $\mathbb{R}^{(N\times T)_\rho}$) defines a sub-space from the space $\mathbb{R}^{(N\times T)}$ of the appropriated size to represent the auxiliary variable $\tilde{\boldsymbol{m}}$ (resp. $\tilde{\boldsymbol{\rho}}$). Note that the production yield is item-independent.

For each scenario s, we define the convex sets $\mathcal{Q}_s$ and $\mathcal{W}_s$. $\mathcal{Q}_s$ represents the expected value from the estimation of the random variable, while the support set $\mathcal{W}_s$ indicates the support of the random variable. $\mathcal{W}_s = \{(\boldsymbol{\rho}, \boldsymbol{m}) \in \mathbb{R}^{(N\times T)} + \mathbb{R}^{(N\times T)_m} | \underline{\underline{\boldsymbol{\rho}}} \leq \boldsymbol{\rho} \leq \overline{\overline{\boldsymbol{\rho}}}, \boldsymbol{g}_s(\boldsymbol{\rho}) \leq \boldsymbol{m}\}$ is the epigraph of $\boldsymbol{g}_s$ and relates the random variable $\boldsymbol{\rho}$ to the auxiliary variable $\tilde{\boldsymbol{m}}$. $\mathcal{W}_s$ also gives the lower bound $\underline{\underline{\boldsymbol{\rho}}}$ (resp. upper bound $\overline{\overline{\boldsymbol{\rho}}}$) of the production yield. We can define the scenario-wise ambiguity set $\mathscr{F}$ as follows :

$$\mathscr{F} = \left\{ \mathbb{P} \in \mathscr{P}_0 \;\middle|\; \begin{array}{ll} (\tilde{\boldsymbol{\rho}}, \tilde{\boldsymbol{m}}, \tilde{s}) \sim \mathbb{P}, \boldsymbol{p} \in \mathbb{P} & \\ \mathbb{E}_{\mathbb{P}}\Big[(\tilde{\boldsymbol{\rho}}, \tilde{\boldsymbol{m}}) \mid \widetilde{s} = s\Big] \in \mathcal{Q}_s & \forall s \in S \\ \mathbb{P}\Big[\Big((\tilde{\boldsymbol{\rho}}, \tilde{\boldsymbol{m}}) \in \mathcal{W}_s \mid \widetilde{s} = s\Big)\Big] = 1 & \forall s \in S \\ \mathbb{P}(\widetilde{s} = s) = p_s & \forall s \in S \end{array} \right\}$$

where $\mathscr{P}_0 \in \mathbb{R}^{(N\times T)+(N\times T)_m} \times S$ is the set of S distributions of the random yield for item $i \in N$ and period $t \in T$, $\mathbb{P}$ is one possible probability distribution for the random production yield, and p_s is the probability of scenario s, such that $\boldsymbol{p} > 0, \sum_{s\in S} p_s = 1$.

Before presenting the MILP reformulation for our problem based on the scenario-wise ambiguity set, we first describe the ambiguity set under each scenario within the scenario-wise ambiguity set in the following subsections.

3.1.1 Mean absolute ambiguity set

The mean absolute ambiguity set $\mathscr{F}_M$ [15, 48] defines the set of possible distributions based on the mean ($\bar{\boldsymbol{\rho}}_s$) and standard deviation ($\hat{\boldsymbol{\rho}}_s$) of the production yield from the lot sizes of each item i in a given period t and scenario $s \in S$. A distribution $\mathbb{P}_s \in \mathscr{F}_M$ can be represented by the support set $\mathcal{W}_s = \{(\boldsymbol{\rho}, \boldsymbol{m}) \in \mathbb{R}^{(N\times T)} + \mathbb{R}^{(N\times T)_m} \mid \underline{\underline{\boldsymbol{\rho}}} \leq \boldsymbol{\rho} \leq \overline{\overline{\boldsymbol{\rho}}},\ ||\boldsymbol{\rho} - \bar{\boldsymbol{\rho}}_s|| \leq \boldsymbol{m}\}\ \forall s \in S$. In addition, the expected value of the random variables takes values in the set $\mathcal{Q}_s = \{(\boldsymbol{q}_\rho, \boldsymbol{q}_m) | \boldsymbol{q}_\rho = \bar{\boldsymbol{\rho}}_s; \boldsymbol{q}_m = \hat{\boldsymbol{\rho}}_s\}$, and $p_s = \frac{1}{|S|}$. Assuming $\mathscr{P}_0 \in (\mathbb{R}^{(N\times T)+(N\times T)} \times S)$, the mean absolute ambiguity set for the uncertain production yield is given as follows :

$$\mathscr{F}_M = \left\{ \mathbb{P} \in \mathscr{P}_0 \;\middle|\; \begin{array}{ll} ((\tilde{\boldsymbol{\rho}}, \tilde{\boldsymbol{m}}), \tilde{s}) \sim \mathbb{P} & \\ \mathbb{E}_\mathbb{P}[\tilde{\boldsymbol{\rho}} \mid \tilde{s} = s] = \bar{\boldsymbol{\rho}}_s & \forall s \in S \\ \mathbb{E}_\mathbb{P}[\tilde{\boldsymbol{m}} \mid \tilde{s} = s] = \hat{\boldsymbol{\rho}}_s & \forall s \in S \\ \mathbb{P}\Big[(\tilde{\boldsymbol{\rho}}, \tilde{\boldsymbol{m}}) \in \mathcal{W}_s \mid \tilde{s} = s\Big] = 1 & \forall s \in S \\ \mathbb{P}(\tilde{s} = s) = \frac{1}{|S|} & \forall s \in S \end{array} \right\}$$

3.1.2 Wasserstein ambiguity set

A Wasserstein ambiguity set contains a collection of probability distributions that are at most θ distant from the available distributions built from past production yield behaviors. As explained by [15] and [48], the Wasserstein ambiguity set $\mathscr{F}_W$ considers an empirical distribution $\breve{\mathbb{P}}$ estimated from historical data, and $\mathscr{F}_W$ contains all probability distributions $\mathbb{P}$ whose Wasserstein distance with $\breve{\mathbb{P}}$ is lower than θ. θ represents the radius of the Wasserstein ball centered at the empirical distribution.

The Wasserstein ambiguity set can be represented as a scenario-wise ambiguity set. For each scenario s, $\mathcal{Q}_s$ is given by $\mathcal{Q}_s = \{(\boldsymbol{q}_\rho, q_m) | q_m = \theta\}$, while $\mathcal{W}_s$ is given by $\mathcal{W}_s = \Big\{(\boldsymbol{\rho}, m) \mid \boldsymbol{\rho} \in [\underline{\underline{\boldsymbol{\rho}}}, \overline{\overline{\boldsymbol{\rho}}}]; ||\boldsymbol{\rho} - \breve{\boldsymbol{\rho}}_s||_p \leq m\Big\}$. Note that p, in the Wasserstein support set, indicates the p-norm. Let us assume $\mathscr{P}_0 \in ((\mathbb{R}^{\mathrm{N\times T})+1} \times S)$. For the uncertain production yield, the Wasserstein ambiguity set is given as follows :

$$\mathscr{F}_W = \left\{ \mathbb{P} \in \mathscr{P}_0 \;\middle|\; \begin{array}{ll} ((\widetilde{\boldsymbol{\rho}}, \widetilde{m}), \widetilde{s}) \sim \mathbb{P} & \\ \mathbb{E}_\mathbb{P}\Big[\widetilde{m} \mid \widetilde{s} \in S\Big] = \theta & \\ \mathbb{P}\Big[(\widetilde{\boldsymbol{\rho}}, \widetilde{m}) \in \mathcal{W}_\mathrm{s}\Big] = 1 & \forall s \in S \\ \mathbb{P}(\widetilde{s} = s) = \frac{1}{|S|} & \forall s \in S \end{array} \right\}$$

3.2 Scenario-wise LSP under yield uncertainty reformulation

Under uncertain yield, the inventory and backorder cost $H_{it}(\widetilde{\rho}, \widetilde{m}, \widetilde{s})$ depend on the realization of the yield of each item i in each period t. We characterize $H_{it}(\widetilde{\rho}, \widetilde{m}, \widetilde{s})$ as an affine function of the uncertain yield with the following scenario-wise decision rule :

$$H_{it}(\boldsymbol{\rho}, \tilde{\boldsymbol{m}}, s) = H^0_{its} + \sum_{l\in T}\sum_{k\in N} H^{'}_{itkls}\rho_{kl} + \sum_{j\in(N\times T)_{|m|}} H^{''}_{itjs} m_j$$

where H^0_{its} represent the cost component that is free from disturbances due to uncertainty, $\boldsymbol{H}^{'}_{its}$ is the cost component that is a function to the realization of the random yield $\tilde{\boldsymbol{\rho}}$, and $\boldsymbol{H}^{''}_{its}$ is the cost component that is a function of the auxiliary variable $\tilde{\boldsymbol{m}}$ for item i in period t and for scenario s. Therefore, $\boldsymbol{H}^0 \in \mathbb{R}_+^{((N\times T)\times S)}$; $\boldsymbol{H}^{'} \in \mathbb{R}^{(N\times T+(N\times T)_{|\rho|}\times S)}$; $\boldsymbol{H}^{''} \in \mathbb{R}^{((N\times T)+(N\times T)_{|m|}\times S)}$. Note that $\boldsymbol{m}$ represents the conditional information for the confidence region of the ambiguity set. Therefore, for the mean absolute ambiguity set, $\tilde{\boldsymbol{m}}$ represents the deviation from the mean for each scenario, such that $\boldsymbol{m} \in \mathbb{R}^{((N\times T)+(N\times T)_{|m|}\times S)}$. For the Wasserstein ambiguity set, $\boldsymbol{m}$ represents the Wasserstein radius that should be considered in the confidence region. Since all the scenarios should respect the same condition on the Wasserstein distance, the vector $\tilde{\boldsymbol{m}}$ under a Wasserstein ambiguity set is reduced to only one variable $\tilde{m}$.

The formulation of the DRLSP under yield uncertainty with scenario-wise ambiguity set and the scenario-wise decision rule is given as problem (2) :

$$\min \left\{ \sum_{i\in N}\sum_{t\in T}\Big(s_{it}Y_{it} + v_{it}X_{it}\Big) + \sup_{\mathbb{P}\in\mathscr{F}} \mathbb{E}_{\mathbb{P}}\Big[\sum_{i\in N}\sum_{t\in T} H_{it}(\widetilde{\boldsymbol{\rho}}, \widetilde{\boldsymbol{m}}, \widetilde{\boldsymbol{s}})\Big]\right\} \tag{2a}$$

$s.t.$ (1d), (1f)

$$H_{it}(\boldsymbol{\rho},\boldsymbol{m},\boldsymbol{s}) \geq h_{it}\left[\sum_{\tau=1}^{t}(\boldsymbol{\rho}_{i\tau}X_{i\tau} - d_{i\tau})\right] \qquad \forall(\boldsymbol{\rho},\boldsymbol{m}) \in \mathcal{W}_s; i \in N; t \in T; s \in S \tag{2b}$$

$$H_{it}(\boldsymbol{\rho},\boldsymbol{m},\boldsymbol{s}) \geq -b_{it}\left[\sum_{\tau=1}^{it}(\boldsymbol{\rho}_{i\tau}X_{i\tau} - d_{i\tau})\right] \qquad \forall(\boldsymbol{\rho},\boldsymbol{m}) \in \mathcal{W}_s; i \in N; t \in T; s \in S \tag{2c}$$

$$H_{it}(\boldsymbol{\rho},\boldsymbol{m},\boldsymbol{s}) \geq 0 \qquad \forall(\boldsymbol{\rho},\boldsymbol{m}) \in \mathcal{W}_s; i \in N; t \in T; s \in S \tag{2d}$$

$$X_{it} \geq 0 \qquad \forall i \in N; t \in T \tag{2e}$$

which can be rewritten problem (3) :

$$\min \sum_{i\in N}\sum_{t\in T}\Big(s_{it}Y_{it} + v_{it}X_{it}\Big) + \sup_{\mathbb{P}\in\mathscr{F}} \mathbb{E}_{\mathbb{P}}\Big[\Phi\Big] \tag{3a}$$

$s.t.$ (1d), (1f), (2e)

$$H^0_{its} + \boldsymbol{H}^{'\top}_{its}\boldsymbol{\rho} + \boldsymbol{H}^{''\top}_{its}\boldsymbol{m} \geq h_{it}\left[\sum_{\tau=1}^{t}(\boldsymbol{\rho}_{i\tau}X_{i\tau} - d_{i\tau})\right] \qquad \forall(\boldsymbol{\rho},\boldsymbol{m}) \in \mathcal{W}_s; i \in N; t \in T; s \in S \tag{3b}$$

$$H^0_{its} + \boldsymbol{H}^{'\top}_{its}\boldsymbol{\rho} + \boldsymbol{H}^{''\top}_{its}\boldsymbol{m} \geq -b_{it}\left[\sum_{\tau=1}^{it}(\boldsymbol{\rho}_{i\tau}X_{i\tau} - d_{i\tau})\right] \qquad \forall(\boldsymbol{\rho},\boldsymbol{m}) \in \mathcal{W}_s; i \in N; t \in T; s \in S \tag{3c}$$

$$H^0_{its} + \boldsymbol{H}^{'\top}_{its}\boldsymbol{\rho} + \boldsymbol{H}^{'\top}_{its}\boldsymbol{m} \geq 0 \qquad \forall(\boldsymbol{\rho},\boldsymbol{m}) \in \mathcal{W}_s; i \in N; t \in T; s \in S \tag{3d}$$

where $\Phi = \sum_{i\in N}\sum_{t\in T}\Big(H^0_{it} + \boldsymbol{H}^{'}_{it}(\widetilde{s})^\top\boldsymbol{\rho} + \boldsymbol{H}^{''}_{it}(\widetilde{s})^\top\boldsymbol{m}\Big)$.

The rest of this section gives the reformulation of problem (3) as a MILP model. To reformulate (3) as a MILP, we follow the first theorem in [11] to transform the worst expectation into robust constraints, and we apply the first theorem in [48] to write the robust counterpart of all constraints subject to the production yield uncertainty.

The first theorem in [11] states that, if Slater's condition holds for the worst-case expectation problem, then the problem can be reformulated as a robust optimization problem. For our model, if Slater's condition is valid for $\sup_{\mathbb{P}\in\mathscr{F}} \mathbb{E}_{\mathbb{P}}\left[\sum_{i\in N}\sum_{t\in T} H_{it}(\widetilde{\boldsymbol{\rho}}, \widetilde{\boldsymbol{m}}, \widetilde{\boldsymbol{s}})\right]$, then the worst expectation for

$\sum_{i\in N}\sum_{t\in T} H_{it}(\widetilde{\boldsymbol{\rho}},\widetilde{\boldsymbol{m}},\widetilde{\boldsymbol{s}})$ can be formulated in a robust fashion as follows :

$$
\begin{aligned}
&\inf \gamma \\
&\gamma \geq \sum_{s\in S}\left(p_s\alpha_s + \boldsymbol{\mu}_s^{'\top}\boldsymbol{\beta}_s^{'} + \boldsymbol{\mu}_s^{''\top}\boldsymbol{\beta}_s^{''}\right) && \forall s\in S\\
&\alpha_s + \boldsymbol{\rho}^\top\boldsymbol{\beta}_s^{'} + \boldsymbol{m}^\top\boldsymbol{\beta}_s^{''} \geq \sum_{i\in N}\sum_{t\in T}\left(H^0_{its} + \boldsymbol{H}_{its}^{'}{}^\top\boldsymbol{\rho} + \boldsymbol{H}_{its}^{''}{}^\top\boldsymbol{m}\right) && \forall(\boldsymbol{\rho},\boldsymbol{m})\in\mathcal{W}_s; s\in S\\
&\gamma\in\mathbb{R},\ \boldsymbol{\alpha}\in\mathbb{R}^S && \forall s\in S\\
&\boldsymbol{\beta}_s^{'}\in\mathbb{R}^{(N\times T)},\ \boldsymbol{\beta}_s^{''}\in\mathbb{R}^{(N\times T)_m} && \forall s\in S
\end{aligned}
$$

Let us denote by $\boldsymbol{\mu}_s^{'}$ and $\boldsymbol{\mu}_s^{''}$ the expected values of the random variables $\tilde{\boldsymbol{\rho}}$ and $\tilde{\boldsymbol{m}}$ in the set $\mathcal{Q}_s$ for a scenario s ($\mu_s^{'},\mu_s^{''}\in\mathcal{Q}_s$). We can then redefine the worst expectation $\sup_{\mathbb{P}\in\mathscr{F}}\mathbb{E}_{\mathbb{P}}\left[\sum_{i\in N}\sum_{t\in T}H_{it}(\widetilde{\boldsymbol{\rho}},\widetilde{\boldsymbol{m}},\widetilde{\boldsymbol{s}})\right] = \lambda(\boldsymbol{p},\boldsymbol{\mu})$ as follows :

$$
\begin{aligned}
\lambda(\boldsymbol{p},\boldsymbol{\mu}) = \sup_{\mathbb{P}\in\mathscr{F}}\sum_{s\in S}p_s\mathbb{E}_{\mathbb{P}_s}&\left[\sum_{i\in N}\sum_{t\in T}H_{it}(\widetilde{\boldsymbol{\rho}},\widetilde{\boldsymbol{m}},s)\right]\\
&s.t.\\
&\mathbb{E}_{\mathbb{P}_s}[\tilde{\boldsymbol{\rho}}] = \boldsymbol{\mu}_s^{'} && \forall s\in S\\
&\mathbb{E}_{\mathbb{P}_s}[\tilde{\boldsymbol{m}}] = \boldsymbol{\mu}_s^{''} && \forall s\in S\\
&\mathbb{P}[(\widetilde{\boldsymbol{\rho}},\widetilde{\boldsymbol{m}})\in\mathcal{W}_s] = 1 && \forall s\in S
\end{aligned}
$$

The min-max theorem helps us to obtain the following dual from the supremum :

$$
\begin{aligned}
\lambda'(\boldsymbol{p},\boldsymbol{\mu}) = \inf\sum_{s\in S}&(\alpha_s + p_s\boldsymbol{\mu}^{'\top}\boldsymbol{\beta}^{'} + p_s\boldsymbol{\mu}^{''\top}\boldsymbol{\beta}^{''})\\
&s.t.\\
&\alpha_s + p_s(\boldsymbol{\rho}^\top\boldsymbol{\beta}^{'} + \boldsymbol{m}^\top\boldsymbol{\beta}^{''}) \geq H_{it}(\widetilde{\boldsymbol{\rho}},\widetilde{\boldsymbol{m}},\widetilde{\boldsymbol{s}}) && \forall(\boldsymbol{\rho},\boldsymbol{m})\in\mathcal{W}_s; s\in S\\
&\boldsymbol{\beta}_s^{'}\in\mathbb{R}^{(N\times T)},\boldsymbol{\beta}_s^{''}\in\mathbb{R}^{(N\times T)_m} && \forall s\in S\\
&\boldsymbol{\alpha}\in\mathbb{R}^S
\end{aligned}
$$

Since strong duality holds, $\lambda^* = \lambda'^*$. As $\boldsymbol{p\mu}$ is non-convex, we can replace μ_s with $\frac{\mu_s}{p_s}$ to obtain a convex representation of the infimum problem. We also replace α_s with $p_s\alpha_s$ (as $p_s > 0,\ p\in\mathbb{P}\ \forall s\in S$). Then, we divide the λ' by p_s which leads to the following reformulation of the infimum :

$$
\begin{aligned}
\lambda'(\boldsymbol{p},\boldsymbol{\mu}) = \inf(&\boldsymbol{p}^\top\boldsymbol{\alpha} + \boldsymbol{\mu}^{'\top}\boldsymbol{\beta}^{'} + \boldsymbol{\mu}^{''\top}\boldsymbol{\beta}^{''})\\
&s.t.\\
&\alpha_s + \boldsymbol{\rho}^\top\boldsymbol{\beta}_s^{'} + \boldsymbol{m}^\top\boldsymbol{\beta}_s^{''} \geq \sum_{i\in N}\sum_{t\in T}H_{it}(\widetilde{\boldsymbol{\rho}},\widetilde{\boldsymbol{m}},\widetilde{\boldsymbol{s}}) && \forall(\boldsymbol{\rho},\boldsymbol{m})\in\mathcal{W}_s; s\in S\\
&\boldsymbol{\beta}_s^{'}\in\mathbb{R}^{(N\times T)},\boldsymbol{\beta}_s^{''}\in\mathbb{R}^{(N\times T)_m} && \forall s\in S\\
&\boldsymbol{\alpha}\in\mathbb{R}^S
\end{aligned}
$$

Replacing λ with λ' in problem (3) leads to the following robust reformulation of our DRLSP model denoted as problem (4) :

$$
\min_{Y,X}\sum_{i\in N}\sum_{t\in T}\left(s_{it}Y_{it} + v_{it}X_{it}\right) + \gamma \tag{4a}
$$

$$\begin{aligned}
&s.t.\ (1d), (1f), (2e), (3b), (3c), (3d) \\
&\gamma \geq \sum_{s \in S} \left(p_s \alpha_s + {\boldsymbol{\mu}_s^{'}}^{\top} \boldsymbol{\beta}_s^{'} + {\boldsymbol{\mu}_s^{''}}^{\top} \boldsymbol{\beta}_s^{''} \right) && \forall p_s \in \mathbb{P}_s;\ \frac{\mu_s}{p_s} \in \mathcal{Q}_s; s \in S && (4b) \\
&\alpha_s + \boldsymbol{\rho}^{\top} \boldsymbol{\beta}_s^{'} + \boldsymbol{m}^{\top} \boldsymbol{\beta}_s^{''} \geq \sum_{i \in N} \sum_{t \in T} \left(H_{its}^{0} + {\boldsymbol{H}_{its}^{'}}^{\top} \boldsymbol{\rho} + {\boldsymbol{H}_{its}^{''}}^{\top} \boldsymbol{m} \right) && \forall (\boldsymbol{\rho}, \boldsymbol{m}) \in \mathcal{W}_s; s \in S && (4c) \\
&\boldsymbol{\beta}_s^{'} \in \mathbb{R}^{(N \times T)}, \boldsymbol{\beta}_s^{''} \in \mathbb{R}^{(N \times T)_m} && \forall s \in S && (4d) \\
&\gamma \in \mathbb{R}, \boldsymbol{\alpha} \in \mathbb{R}^{S} && && (4e)
\end{aligned}$$

Problem (4) contains an infinite set of constraints since the support $\mathcal{W}_s$ of the uncertain yield is not a finite set. To reformulate these constraints as linear robust counterpart constraints, we recall the first theorem in [48] that requires two steps. First, [48] apply Sion's minimax theorem (see [40]) to reverse the order of a sup and inf problem concerning a bounded distribution $\mathbb{P}$. Second, the authors rely on strong duality to derive robust counterpart reformulations of constraints subject to uncertainty.

Equations (5) give the robust counterpart reformulation of (4b). The second line separates the dual variables into individual problems, and it shows that only $\boldsymbol{\mu}$ dual variables are impacted by $\mathcal{Q}$, while all dual variables are dependent on $\mathbb{P}$. In the third line, we redefine the supremum problem in terms of the support function. In the fourth line, we redefine the support function as an infimum problem, and we gather the dual variables in the same supremum problem dependent on $\mathbb{P}$. In the fifth line, we exploit Sion's minimax theorem to reverse the order of the supremum and infimum problems. Finally, the last line gives the reformulation in terms of a support function for $\mathbb{P}$.

$$\begin{aligned}
&\sup_{\left\{\frac{\boldsymbol{\mu}_s}{p_s}\right\} \in \mathcal{Q}_s, s \in S, \boldsymbol{p} \in \mathbb{P}} \boldsymbol{\alpha}^{\top} \boldsymbol{p} + \boldsymbol{\beta}^{\top} \boldsymbol{\mu} \\
&= \sup_{\boldsymbol{p} \in \mathbb{P}} \boldsymbol{\alpha}^{\top} \boldsymbol{p} + \sum_{s \in S} p_s \sup_{\left\{\frac{\boldsymbol{\mu}_s}{p_s}\right\} \in \mathcal{Q}_s} \frac{\boldsymbol{\beta}_s^{\top} \boldsymbol{\mu}_s}{p_s} \\
&= \sup_{\boldsymbol{p} \in \mathbb{P}} \boldsymbol{\alpha}^{\top} \boldsymbol{p} + \sum_{s \in S} p_s \delta^* \left(\boldsymbol{\beta}_s | \mathcal{Q}_s \right) \qquad (5) \\
&= \sup_{\boldsymbol{p} \in \mathbb{P}} \left(\boldsymbol{\alpha}^{\top} \boldsymbol{p} + \boldsymbol{p} \inf_{\nu : \nu \geq \delta^* (\boldsymbol{\beta}_s | \mathcal{Q}_s)} \nu \right) \\
&= \inf_{\nu : \nu \geq \delta^* (\boldsymbol{\beta}_s | \mathcal{Q}_s)} \sup_{\boldsymbol{p} \in \mathbb{P}} \boldsymbol{p} (\boldsymbol{\alpha} + \nu) \\
&= \inf_{\nu : \nu \geq \delta^* (\boldsymbol{\beta} | \mathcal{Q})} \delta^* \left(\boldsymbol{\alpha} + \nu | \mathbb{P} \right)
\end{aligned}$$

The final linear reformulation is obtained by replacing the general support function with the specific linear support function of the considered ambiguity set.

For the remaining constraints subject to the uncertain production yield, we assume that Slater's condition holds, as does the strong duality. Slater's condition helps us to transform the worst expectation problem into a robust formulation. On the other side, the strong duality allows us to develop a tractable linear formulation for the obtained robust model. Thus, we derive equivalent robust counterpart formulations for these constraints. As a result, the constraints seek the supremum of the random variables over a support set [48]. The supremum problem sup is defined with the support functions of the support set. On the other hand, its dual (inf) is formulated using the conjugate of the epigraph g. Thus, we apply the strong duality to redefine the sup problem in terms of its dual inf. First, we demonstrate how to obtain the counterpart reformulation for constraint (4c). Then, we provide the reformulation for the remaining constraints in A.1.

We rewrite constraint (4c) to isolate the terms with the random variables on the right side, and we obtain the following reformulation :

$$\alpha_s - \sum_{i\in N}\sum_{t\in T} H^0_{its} \geq (\sum_{i\in N}\sum_{t\in T}\boldsymbol{H}^{'}_{its} - \boldsymbol{\beta}^{'}_s)^\top \boldsymbol{\rho} + (\sum_{i\in N}\sum_{t\in T}\boldsymbol{H}^{''}_{its} - \boldsymbol{\beta}^{''}_s)^\top \boldsymbol{m} \qquad \forall(\boldsymbol{\rho},\boldsymbol{m}) \in \mathcal{W}_s; s \in S$$

Note that the right side term can be rewritten as a convex optimization problem based on the support function of $\mathcal{W}_s$. If we exploit Slater's condition and the strong duality since the random variable $\tilde{\boldsymbol{\rho}}$ is bounded by $\underline{\underline{\boldsymbol{\rho}}}$ and $\overline{\overline{\boldsymbol{\rho}}}$ on $\mathcal{W}_s$, we can assume $\boldsymbol{\rho} = \frac{\overline{\overline{\boldsymbol{\rho}}}+\underline{\underline{\boldsymbol{\rho}}}}{2}$ and $\boldsymbol{v} = \boldsymbol{g}_s(\boldsymbol{\rho}) + 1$. Then, we define a new random variable $\boldsymbol{\xi}_j \in \mathcal{W}_s : \boldsymbol{\xi}_j = \boldsymbol{\rho}\ \forall j \in (N\times T)_{|m|}$ for any scenario $s \in S$. Finally, we obtain the following convex optimization problem :

$$\sup_{\kappa} \boldsymbol{\rho}^\top \boldsymbol{c}^1_s + \boldsymbol{m}^\top \boldsymbol{c}^2_s$$

where

$$\kappa = \{\boldsymbol{\rho}, \boldsymbol{m}, \{\boldsymbol{\xi}_j\}_{j\in(N\times T)_{|m|}} \in \mathbb{R},\ \boldsymbol{\rho} \in [\underline{\underline{\boldsymbol{\rho}}}, \overline{\overline{\boldsymbol{\rho}}}].$$

We also have

$$g_{js}(\xi_j) \leq m_j,$$
$$(\boldsymbol{\xi}_j) = \rho,\ \forall j \in (N\times T)_{|m|}\},$$
$$\boldsymbol{c}^1_s = \sum_{i\in N}\sum_{t\in T}\boldsymbol{H}^{'}_{its} - \boldsymbol{\beta}^{'}_s$$

and

$$\boldsymbol{c}^2_s = \sum_{i\in N}\sum_{t\in T}\boldsymbol{H}^{''}_{its} - \boldsymbol{\beta}^{''}_s.$$

As the strong duality holds, we can then reformulate this sup problem in terms of its inf dual problem through the Lagrangian duality. Thus, we obtain the following primal-dual problem :

$$\begin{array}{c} \textbf{primal} = \sup \boldsymbol{\rho}^\top \boldsymbol{c}^1_s + \boldsymbol{m}^\top \boldsymbol{c}^2_s \\ \overline{\overline{\boldsymbol{\rho}}} - \boldsymbol{\rho} \geq 0 \\ -(\underline{\underline{\boldsymbol{\rho}}} - \boldsymbol{\rho}) \geq 0 \\ m_j - g_{js}(\boldsymbol{\xi}_j) \geq 0 \\ \boldsymbol{\xi}_j - \boldsymbol{\rho} \geq 0 \\ \boldsymbol{\rho}, \boldsymbol{m}, \boldsymbol{\xi}_j \in \mathbb{R} \end{array} \quad \begin{array}{c} \boldsymbol{\eta}^1_s, \boldsymbol{\eta}^2_s, \boldsymbol{\lambda}^1_s \geq 0 \\ \boldsymbol{w}^1_{js} \in \mathbb{R} \end{array} \quad \xrightarrow[\text{dualized}]{} \quad \begin{array}{c} \inf fct \\ \sum_{j\in(N\times T)_{|m|}} \boldsymbol{w}^1_{js} = \boldsymbol{c}^1_s - \boldsymbol{\eta}^1_s + \boldsymbol{\eta}^2_s \\ \boldsymbol{c}^2_s + \boldsymbol{\lambda}^1_s = 0 \end{array}$$

where the dual's objective function is given by $fct = \overline{\overline{\boldsymbol{\rho}}}^\top \boldsymbol{\eta}^1_s - \underline{\underline{\boldsymbol{\rho}}}^\top \boldsymbol{\eta}^2_s + \sum_{j\in(N\times T)_{|m|}} \lambda^1_{js}(m_j - g_{js}(\boldsymbol{\xi}_j)) + {\boldsymbol{w}^1_{js}}^\top(\boldsymbol{\xi}_j - \boldsymbol{\rho})$. The dual variables are $\boldsymbol{\eta}^1$, $\boldsymbol{\eta}^2$, $\boldsymbol{\lambda}^1$ and $\boldsymbol{w}^1$. In addition, the equations indexed by j are defined for all $j \in (N\times T)_{|m|}$.

Finally, we rewrite the convex optimization problem as linear robust constraints based on the obtained dual problem and the epigraph in the support set $\mathcal{W}_s$ as given below. From the assumption that Slater's condition is valid, the strong duality holds. Thus, the optimal dual variables from the infimum yield an optimal supremum solution. We apply the minimax theorem to inverse the order of the sup and inf problems on the second line, we rewrite the dual objective function in terms of its conjugate function on the third line, and we isolate the sup problem that is now only dependent on the dual variables $\boldsymbol{w}$ and $\boldsymbol{\lambda}$ on the fourth line. Then, we replace the later conjugate function with its perspective on the fifth line given below.

$$\begin{aligned} &\sup_{\kappa} \boldsymbol{\rho}^\top \boldsymbol{c}^1_s + \boldsymbol{m}^\top \boldsymbol{c}^2_s \\ &= \sup_{\kappa}\inf_{\zeta} \boldsymbol{\rho}^\top \boldsymbol{c}^1_s + \boldsymbol{m}^\top \boldsymbol{c}^2_s + (\overline{\overline{\boldsymbol{\rho}}} - \boldsymbol{\rho})^\top \boldsymbol{\eta}^1_s - (\underline{\underline{\boldsymbol{\rho}}} - \boldsymbol{\rho})^\top \boldsymbol{\eta}^2_s + \sum_{j\in(N\times T)_{|m|}} \lambda^1_{js}(m_j - g_{js}(\boldsymbol{\xi}_j)) + {\boldsymbol{w}^1_{js}}^\top(\boldsymbol{\xi}_j - \boldsymbol{\rho}) \end{aligned}$$

$$= \inf_{\zeta} \sup_{\kappa} \boldsymbol{\rho}^\top \boldsymbol{c}_s^1 + \boldsymbol{m}^\top \boldsymbol{c}_s^2 + (\overline{\overline{\boldsymbol{\rho}}} - \boldsymbol{\rho})^\top \boldsymbol{\eta}_s^1 - (\underline{\underline{\boldsymbol{\rho}}} - \boldsymbol{\rho})^\top \boldsymbol{\eta}_s^2 + \sum_{j \in (N \times T)_{|m|}} \lambda_{js}^1 (m_j - g_{js}(\boldsymbol{\xi}_j)) + {\boldsymbol{w}_{js}^1}^\top (\boldsymbol{\xi}_j - \boldsymbol{\rho})$$

$$= \inf_{\zeta} \overline{\overline{\boldsymbol{\rho}}}^\top \boldsymbol{\eta}_s^1 - \underline{\underline{\boldsymbol{\rho}}}^\top \boldsymbol{\eta}_s^2 + \sum_{j \in (N \times T)_{|m|}} \sup_{\boldsymbol{\xi}_j} {\boldsymbol{w}_{js}^1}^\top (\boldsymbol{\xi}_j) - \lambda_{js}^1 g_{js}(\boldsymbol{\xi}_j)$$

$$= \inf_{\zeta} \overline{\overline{\boldsymbol{\rho}}}^\top \boldsymbol{\eta}_s^1 - \underline{\underline{\boldsymbol{\rho}}}^\top \boldsymbol{\eta}_s^2 + \sum_{j \in (N \times T)_{|m|}} \lambda_{js}^1 g_{js}^* \Big(\frac{\boldsymbol{w}_{js}^1}{\lambda_{js}^1} \Big)$$

where $\zeta = \{\boldsymbol{\eta}_s^1, \boldsymbol{\eta}_s^2, \boldsymbol{\lambda}_s^1 \geq 0;\ \boldsymbol{w}_{js}^1 \in \mathbb{R}\ \forall j \in (N \times T)_{|m|};\ \boldsymbol{c}_s^2 + \boldsymbol{\lambda}_s^1 = 0;\ \sum_{j \in (N \times T)_{|m|}} \boldsymbol{w}_{js}^1 = \boldsymbol{c}_s^1 - \boldsymbol{\eta}_s^1 + \boldsymbol{\eta}_s^2\}$ and $j \in (N \times T)_{|m|}$ whenever j index appears. As a result, we obtain the following reformulation for the sup convex model :

$$\sup_{\kappa} \boldsymbol{\rho}^\top \boldsymbol{c}_s^1 + \boldsymbol{m}^\top \boldsymbol{c}_s^2 = \inf_{\zeta} \overline{\overline{\boldsymbol{\rho}}}^\top \boldsymbol{\eta}_s^1 - \underline{\underline{\boldsymbol{\rho}}}^\top \boldsymbol{\eta}_s^2 + \sum_{j \in (N \times T)_{|m|}} \lambda_{js}^1 g_{js}^* \Big(\frac{\boldsymbol{w}_{js}^1}{\lambda_{js}^1} \Big)$$

where $\lambda_{js}^1 g_{js}^* \Big(\frac{\boldsymbol{w}_{js}^1}{\lambda_{js}^1} \Big)$ is the perspective function of the conjugate function of g_{js}^* on $(\boldsymbol{w}_{js}^1,\ \lambda_{js}^1)$ for all $j \in (N \times T)_{|m|}$.

Note that the superscript 1 on the dual variables $\boldsymbol{\lambda}$ and $\boldsymbol{w}$, (resp. the subscripts 1 and 2 on the coefficients $\boldsymbol{c}$ and the dual variables $\boldsymbol{\eta}$) indicates the terms associated to constraints (4c), while the subscript 2 (resp. 3 and 4) is associated to constraints (3b), subscript 3 (resp. 5 and 6) to constraints (3c), and subscript 4 (resp. 7 and 8) to constraints (3d).

Repeating the aforementioned reformulation technique to obtain a linear robust counterpart to all constraints subject to the uncertain production yield, we obtain the final MILP robust reformulation of problem (4), which is given as in Problem (6) :

$$\min \sum_{i \in N} \sum_{t \in T} \Big(s_{it} Y_{it} + v_{it} X_{it} \Big) + \gamma \tag{6a}$$

$s.t.$

$$\gamma \geq \delta^* \Big(\boldsymbol{\alpha} + \mathbf{1}\nu | \mathbb{P} \Big) \tag{6b}$$

$$\nu \geq \delta^* \Big(\boldsymbol{\beta} | \mathcal{Q} \Big) \tag{6c}$$

$$\alpha_s - \sum_{i \in N} \sum_{t \in T} H_{its}^0 \geq \sum_{i \in N} \sum_{t \in T} \overline{\overline{\rho}}_{it} \eta_{its}^1 - \sum_{i \in N} \sum_{t \in T} \underline{\underline{\rho}}_{it} \eta_{its}^2 + \sum_{j \in (N \times T)_{|m|}} \lambda_{js}^1 g_{js}^* \Big(\frac{\boldsymbol{w}_{js}^1}{\lambda_{js}^1} \Big) \quad \forall s \in S \tag{6d}$$

$$\lambda_{js}^1 = \beta_{js}^{''} - \sum_{i \in N} \sum_{t \in T} H_{itjs}^{''} \quad \forall j \in (N \times T)_m;\ s \in S \tag{6e}$$

$$\sum_{j \in (N \times T)_{|m|}} \boldsymbol{w}_{js}^1 = \sum_{i \in N} \sum_{t \in T} (-\beta_{its}^{'} - \eta_{its}^1 + \eta_{its}^2) \sum_{i \in N} \sum_{t \in T} (\sum_{k \in N} \sum_{l \in T} \boldsymbol{H}_{itkls}^{'}) \quad \forall s \in S \tag{6f}$$

$$H_{its}^0 + h_{it} \sum_{\tau=1}^{t} d_{i\tau} \geq \sum_{k \in N} \sum_{l \in T} (\overline{\overline{\rho}}_{kl} \eta_{itkls}^3 - \underline{\underline{\rho}}_{kl} \eta_{itkls}^4) + \sum_{j \in (N \times T)_{|m|}} \lambda_{itjs}^2 g_{js}^* \Big(\frac{\boldsymbol{w}_{itjs}^2}{\lambda_{itjs}^2} \Big) \quad \forall i \in N; t \in T; s \in S \tag{6g}$$

$$\lambda_{itjs}^2 = H_{itjs}^m \quad \forall i \in N; t \in T;\ j \in (N \times T)_m;\ s \in S \tag{6h}$$

$$\sum_{j\in(N\times T)_{|m|}} w^2_{itjs} = -H^{'}_{itkls} - \eta^3_{itkls} + \eta^4_{itkls} \qquad \forall i,k\in N; k\neq i;\ t,l\in T; s\in S \tag{6i}$$

$$\sum_{j\in(N\times T)_{|m|}} w^2_{itjs} = -H^{'}_{itkls} - \eta^3_{itkls} + \eta^4_{itkls} \qquad \forall i\in N; t,l\in T;\ l>t; s\in S \tag{6j}$$

$$\sum_{j\in(N\times T)_{|m|}} w^2_{itjs} = h_{it}X_{il} - H^{'}_{itkls} - \eta^3_{itkls} + \eta^4_{itkls} \qquad \forall i\in N; t,l\in T;\ l\leq t; s\in S \tag{6k}$$

$$H^0_{its} + b_{it}\sum_{\tau=1}^{t}(\rho_{i\tau}X_{i\tau} - d_{i\tau}) \geq \sum_{k\in N}\sum_{l\in T}(\overline{\overline{\rho}}_{kl}\eta^5_{itkls} - \underline{\underline{\rho}}_{kl}\eta^6_{itkls}) + \sum_{j\in(N\times T)_{|m|}} \lambda^2_{itjs} g^*_{js}\left(\frac{\boldsymbol{w}^3_{itjs}}{\lambda^2_{itjs}}\right) \qquad \forall i\in N; t\in T;\ s\in S \tag{6l}$$

$$\lambda^3_{itjs} = H^m_{itjs} \qquad \forall i\in N; t\in T;\ j\in(N\times T)_m;\ s\in S \tag{6m}$$

$$\sum_{j\in(N\times T)_{|m|}} w^3_{itjs} = -H^{'}_{itkls} - \eta^5_{itkls} + \eta^6_{itkls} \qquad \forall i,k\in N; k\neq i;\ t,l\in T; s\in S \tag{6n}$$

$$\sum_{j\in(N\times T)_{|m|}} w^3_{itjs} = -H^{'}_{itkls} - \eta^5_{itkls} + \eta^6_{itkls} \qquad \forall i\in N; t,l\in T;\ l>t; s\in S \tag{6o}$$

$$\sum_{j\in(N\times T)_{|m|}} w^3_{itjs} = -b_{it}X_{il} - H^{'}_{itkls} - \eta^5_{itkls} + \eta^6_{itkls} \qquad \forall i\in N; t,l\in T;\ l\leq t; s\in S \tag{6p}$$

$$H^0_{its} \geq \sum_{k\in N}\sum_{l\in T}(\overline{\overline{\rho}}_{kl}\eta^7_{itkls} - \underline{\underline{\rho}}_{kl}\eta^8_{itkls}) + \sum_{j\in(N\times T)_{|m|}} \lambda^2_{itjs} g^*_{js}\left(\frac{\boldsymbol{w}^4_{itjs}}{\lambda^2_{itjs}}\right) \qquad \forall s\in S \tag{6q}$$

$$\lambda^4_{itjs} = H^m_{itjs} \qquad \forall i\in N; t\in T;\ j\in(N\times T)_m;\ s\in S \tag{6r}$$

$$\sum_{j\in(N\times T)_{|m|}} \boldsymbol{w}^4_{itjs} = -\boldsymbol{H}^{'}_{its} - \boldsymbol{\eta}^7_{its} + \boldsymbol{\eta}^8_{its} \qquad \forall s\in S \tag{6s}$$

$$X_{it} \leq M_{it}Y_{it} \qquad \forall i\in N; t\in T \tag{6t}$$

$$X_{it}, H^0_{its}, \boldsymbol{\eta}^1_s, \boldsymbol{\eta}^2_s, \boldsymbol{\eta}^3_{its}, \boldsymbol{\eta}^4_{its}, \boldsymbol{\eta}^5_{its}, \boldsymbol{\eta}^6_{its}, \boldsymbol{\eta}^7_{its}, \boldsymbol{\eta}^8_{its}, \lambda^1_{js}, \lambda^2_{itjs} \geq 0 \qquad \forall i\in N; t\in T;\ j\in(N\times T)_m;\ s\in S \tag{6u}$$

$$\boldsymbol{\alpha}\in\mathbb{R}^S, \boldsymbol{\beta}^{'}_s\in\mathbb{R}^{(N\times T)\times S};\quad \boldsymbol{\beta}^{''}_s\in\mathbb{R}^{(N\times T)\times S}_m, \boldsymbol{H}^{'}_{its}, \boldsymbol{H}^{''}_{its}, \gamma, \boldsymbol{w}^1_{js}, \boldsymbol{w}^2_{itjs}, \boldsymbol{w}^3_{itjs}, \boldsymbol{w}^4_{itjs}\in\mathbb{R} \qquad \forall i\in N; t\in T;\ j\in(N\times T)_m;\ s\in S \tag{6v}$$

$$Y_{it}\in\{0,1\} \qquad \forall i\in N; t\in T \tag{6w}$$

where : constraints (4b) are reformulated as constraints (6b)–(6c) ; constraints (4c) are reformulated as constraints (6d)–(6f) ; constraints (3b) are reformulated as constraints (6g)–(6k) ; constraints (3c) are reformulated as constraints (6l)–(6p) ; and constraints (3d) are reformulated as constraints (6q)–(6s).

4 Computational experiments and discussion

The objectives of the computational experiments are : (i) to evaluate the performance of distributionally robust optimization (DRO) models for the LSP with uncertain production yield compared to the solutions obtained from traditional robust optimization and stochastic programming models ; (ii) to present an in-depth investigation of distributionally robust plans for the LSP with uncertain production yield in terms of costs and quality, as well as computational efficiency ; (iii) to demonstrate the performance of the DRO in a real-world setting through a case study.

We consider the following models in the experiments :

1. RO, the static robust optimization model presented in Appendix A.2
2. SP_N, the two stage stochastic programming model presented in Appendix A.3 that assumes the production yield follows a normal distribution
3. SP_U, the two stage stochastic programming model presented in Appendix A.3 that assumes the production yield follows a uniform distribution
4. $MDRO$, the mean absolute distributionally robust model given in Appendix A.4
5. $WDRO$, the Wasserstein distributionally robust model given in Appendix A.5

4.1 Instance generation and simulation approach

The experiments are performed with instances generated following the standard approach in the literature on LSPs as in [31]. Each parameter was drawn from a uniform distribution, such that the production cost, holding stock cost, demand, nominal value, and maximum deviation of the uncertain production yield supports correspond to the following intervals : $v_{it} \in U(30, 50)$, $h_{it} \in U(1, 10)$, $d_{it} \in U(450, 780)$, $\bar{\rho}^{'}_{it} \in U(0.7, 0.9)$, and $\hat{\rho}^{'}_{it} \in U(0.01, 0.1)$, respectively. The setup costs for each item i are computed with the time between order formula $s_{it} = \frac{\bar{D}_{it} \cdot TBO^2 \cdot h_t}{2}$, where $\bar{D}_{it}$ represents the average demand for item i in periods up to t and the time between order (TBO) indicates the duration between a customer's orders. We also set the inventory and backorder levels at the beginning of the horizon to zero.

To represent data of historical yields, we randomly generated 100 historical production yields for each item. These vectors are drawn from a uniform distribution with support $[\bar{\rho}^{'}_{it} - \hat{\rho}^{'}_{it}; \bar{\rho}^{'}_{it} + \hat{\rho}'_{it}]$ for each item i in period t. From this generated data set, a K-means clustering algorithm is applied to partition the data in K different scenarios that share some distributional information in terms of the production yield rate [28]. For each cluster s (which represents a scenario or pattern of the uncertain production yield), we compute the average production yield $\bar{\rho}_{its}$, standard deviation of the production yield $\hat{\rho}_{its}$, and finally, upper and lower bounds on the production yield given by $\bar{\bar{\rho}}_{its}$ and $\underline{\underline{\rho}}_{its}$, respectively, for all item $i \in N$, all period $t \in T$ and all scenario $s \in S$ that should be considered on the ambiguity set. From the generated data set, we also take the average production yield $\bar{\rho}^{''}_{it}$, standard deviation of the production yield $\hat{\rho}^{''}_{it}$, and finally, upper and lower bounds on the production yield given by $\bar{\bar{\rho}}^{''}_{it}$ and $\underline{\underline{\rho}}^{''}_{it}$ for all item $i \in N$, all period $t \in T$ to represent the uncertainty set for the robust model or probability distributions for the stochastic programs considered in this section.

The ambiguity sets are created from the generated historical data, and so does the uncertainty set and parameters to define the distributions for the stochastic models. For $WDRO$, each centroid of the obtained clusters represents the empirical distribution for the respective scenario. For $MDRO$, we compute the mean and standard deviation of the production yield in each cluster. We consider that a completely robust production plan can be computed with the most conservative uncertainty set from the robust optimization, i.e. the box uncertainty set. This set is equivalent to the budgeted uncertainty set proposed by [4] when the budget is set to $\Gamma_t = t$ for each period t. The production yield is mapped in terms of the nominal value and maximum deviation of the production yield calculated from the average value over all the scenarios. For the stochastic programs, we consider two possible estimations for the probability distribution of the uncertain yield. While SP_U uses 100 scenarios drawn from a uniform distribution based on the maximum and minimum values of the production yield measured from the generated data set, SP_N samples scenarios from a normal distribution based on the average and standard deviation of the production yield obtained from the generated data. The algorithms were implemented in Python 3.6, and the MILP and the RSOME models are solved with CPLEX version 12.10. We use the KMeans function from the Python scikit-learn 1.2.2 library to partition the generated data [25]. The experiments were run on Intel(R) Gold 6148/2.4GHz processors with 92G of RAM. All the models for all the instances were solved to optimality.

To investigate the quality of the distributionally robust optimization solutions, we evaluate the models' performance through a Monte Carlo simulation with $|\Omega| = 5000$ scenarios, where each scenario ω gives a possible production yield rate ρ_{it}^{ω} for each item i in each period t, and the distribution of scenarios follows the same probability of occurrence as the scenarios in the ambiguity set. Considering a case of good estimation of the production yield, the true realization ρ_{it}^{ω} follows a uniform distribution with support $[\underline{\underline{\rho}}_{it}^{''};\overline{\overline{\rho}}_{it}^{''}]$. When we consider the case of misspecification of the distribution of the production yield, we assume that ρ_{it}^{ω} follows a normal distribution with the average production yield $\bar{\rho}_{it}^{''}$ and standard deviation of the production yield $\hat{\rho}_{it}^{''}$. We evaluate each model by solving the deterministic model for each scenario ω with the setups and lot size decisions fixed to the values obtained from the optimization step.

The rest of this section is organized as follows. First, Section 4.3 defines the best parameters for the ambiguity sets, that is, the number of scenarios, and the Wasserstein radius for the Wasserstein ambiguity set. Then, Section 4.3 reports the simulation results based on the production plans for all the considered models. Next, Section 4.4 presents the results of an industrial case study.

4.2 Definition of the ambiguity set

To design a sufficiently good ambiguity set, we estimate the best parameters to compute the DRO models based on the average expected cost and the computational time. For that, we consider the same data set for an instance with 5 items and 12 production periods, and we analyze the impact of these parameters on the solution. For both DRO models, we partition the generated data in 8 scenarios (that is $S \in \{1, ..., 8\}$. Considering the $WDRO$ model, we also show how the Wasserstein distance impacts the expected cost and the computational time. We consider a Wasserstein radius $\theta \in \{0.01, 0.1, 0.5, 1\}$, where the higher the radius, the more conservative the ambiguity set is.

Figure 1 shows the average expected cost and computation time of the $MDRO$ and $WDRO$ models for different numbers of scenarios in their respective ambiguity sets. Although the results under an ambiguity set of one unique scenario give the lowest expected average cost and computation time, this ambiguity set disregards the possibility of the different and independent possible patterns describing the uncertain parameter. Thus, we evaluate the best number of scenarios that can define different and independent scenarios yielding a low expected cost and an acceptable computational time. Therefore, Figure 1 shows that three scenarios are enough to quickly calculate a satisfactory $MDRO$ plan whose average cost is the lowest, and different and independent scenarios are taken into account separately. The figure also indicates that a good $WDRO$ plan can be defined over the same amount of scenarios as the MRO plan. Furthermore, the figure shows that although low θ lead to a slightly lower expected cost, no significant difference is observed. Therefore, we can set the Wasserstein distance to its most conservative value ($\theta = 1$), without a significant increase in the expected costs.

4.3 Performance of the models

This section reports the performance and quality of production plans for the considered LSP under production yield uncertainty. Based on the results presented in Section 4.2 we consider 3 scenarios for the ambiguity sets considered here, and we set the Wasserstein radius to 1. For the simulation, we

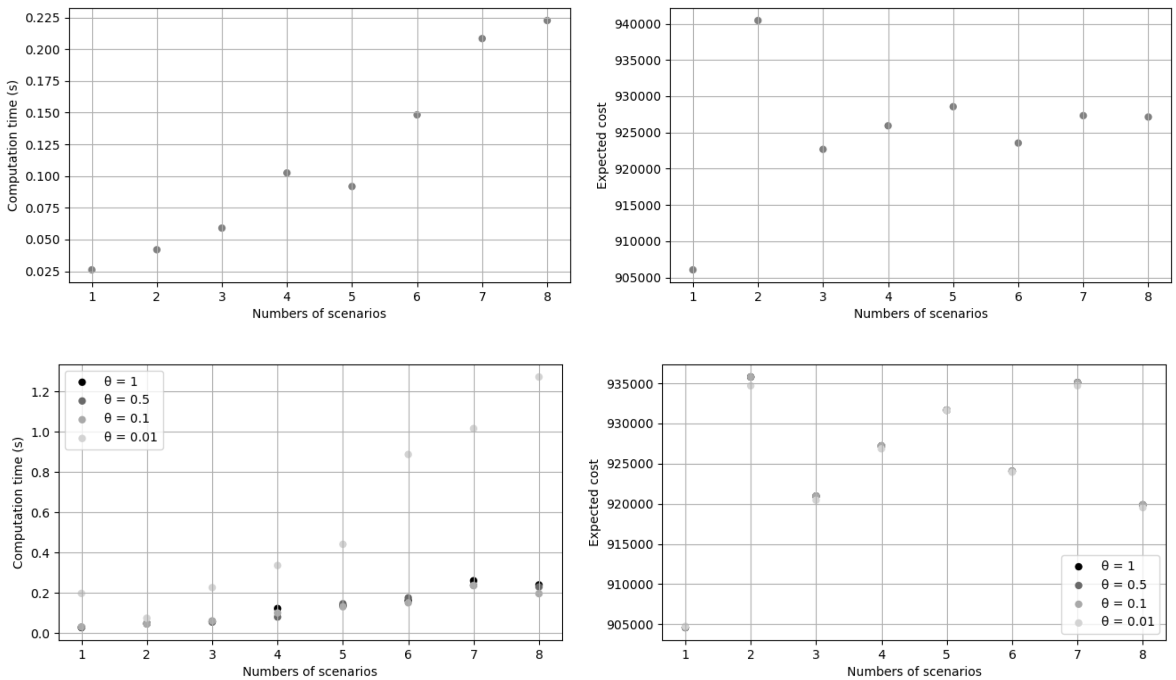


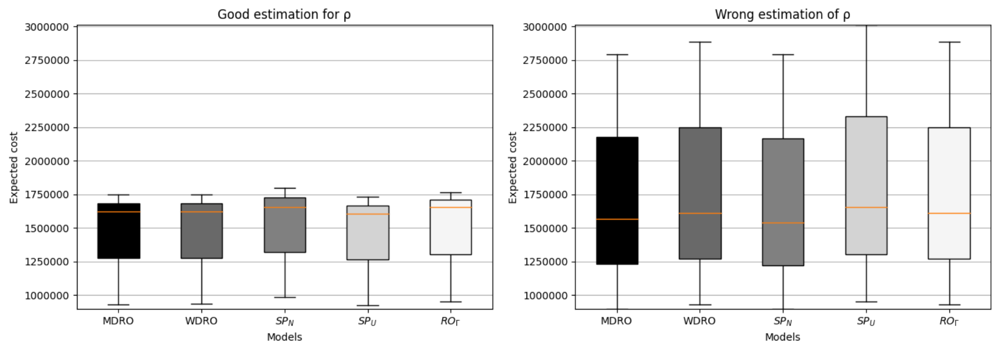
Figure 1 – Setting the parameters for the DRO models

consider two simulation frameworks : an in-sample and an out-of-sample context. The later simulation framework investigates a situation when the underlying yield uncertainty can differ from the distribution used to generate the data. The first considers that the generated data is a good representation of the future value of the production yield. When simulating with a uniform distribution, SP_U gives an expected cost whose distribution is a good estimation of the uncertainty, while SP_N proposes the production plan when the stochastic program follows a wrong estimation of the uncertainty. When simulating with a normal distribution, the SP_U plan is based on a wrong estimation of the uncertainty, and SP_N gives the production plan that follows a good estimation of the uncertainty.

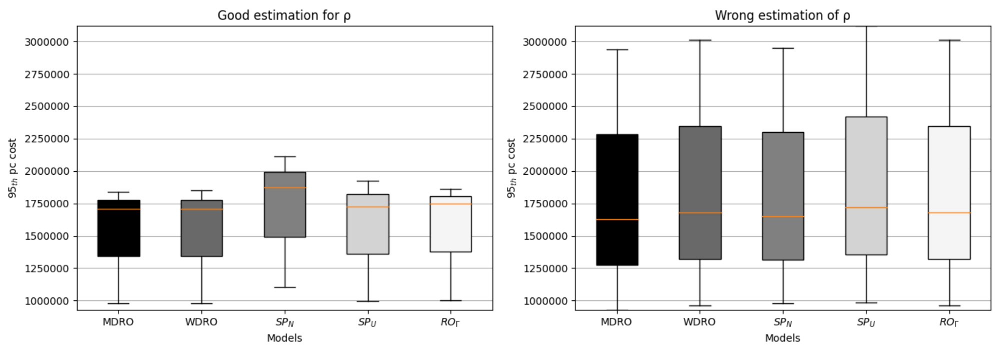
We present some charts to represent the impact of the production yield on the production plan considering the two simulation frameworks. In each figure, we represent the in-sample simulation on the left side, while the right side corresponds to the out-of-sample simulation. Figure 2 reports on the impact of the realization of the production yield on the costs. The DRO models outperform other methods (that is SP and RO models) since they result in the lowest average costs. As $MDRO$ is more conservative than $WDRO$, it presents the lowest expected costs based on average scenarios (even in the case of misspecification of the production yield distribution). $MDRO$ yields the lowest costs in the more pessimistic scenarios (here represented by the 95^{th} and 99^{th} percentile costs and also the cases of misspecification of the uncertain distribution.

Table 1 and Table 2 give the numerical results represented in Figure 2 for the in-sample and out-of-sample simulation, respectively. In these tables, we compare the methods based on the average computational time (column $Time$, in seconds) from the optimization, and the expected value (column $Exp.Cost$) of each solution approaches evaluated in the simulation, along with the 95^{th} and 99^{th} percentile cost (p.c.), where the 99^{th} percentile cost gives the approximate behavior of the models for an adverse context. We also indicate the coefficient of variation CV of the costs, which gives the percentage of variability of the costs. Thus, CV gives the ratio of the standard deviation to the mean, where a high CV indicates costs widely dispersed from the average expected cost.

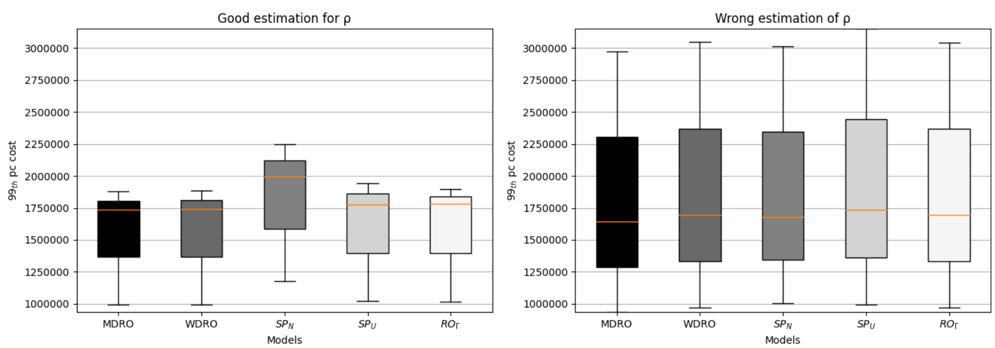
Analyzing the in-sample simulation, Table 1 confirms that DRO models outperform other approaches. Although the expected value obtained with SP_U is 1% (resp. 1.1%) lower than the respec-



(a) Average expected cost



(b) Average 95^{th} percentile cost



(c) Average 99^{th} percentile cost

Figure 2 – Cost distributions of the production plans

tive costs obtained with $WDRO$ (resp. DRO), $WDRO$ (resp. $MDRO$) reduces the average expected cost by 1.7% (resp. 1.6%) compared to the RO plan, and by 3.2% (resp. 3.1%) compared to SP_N. In addition, $WDRO$ (resp. $MDRO$) reduces the 95^{th} percentile costs by 2%, 2.6% and 12.5% (resp. 1.8%, 2.4% and 12.2%) compared to RO, SP_U and SP_N, respectively, with a reduction in terms of 99^{th} percentile costs of about 1.9%, 2.8% and 17.5% (resp. 1.7%, 2.6% and 17.3%). The DRO models are more robust to adverse events (and less impacted by variations on the realization of the uncertain yield) with a CV of 2.92% for both models, followed by RO (CV of 2.97%), SP_U (CV of 4.95%), and finally SP_U (CV of 8.11%). Therefore, in terms of average costs and a controlled environment (the in-sample simulation), $WDRO$ followed by $MDRO$ present the best production plans, with the lowest average costs and highest cost savings when compared to the other methodologies.

Table 1 – Performance of the models when the simulation follows a uniform distribution

Model	Exp. Cost	95^{th} p.c.	99^{th} p.cc	Time	CV
$WDRO$	1,432,243	1,508,904	1,536,600	0.37	2.92%
$MDRO$	1,433,744	1,512,217	1,540,049	0.45	2.92%
SP_N	1,477,526	1,697,263	1,806,253	0.18	8.11%
SP_U	1,417,972	1,548,363	1,579,840	0.16	4.95%
RO	1,456,837	1,538,831	1,565,853	0.01	2.97%

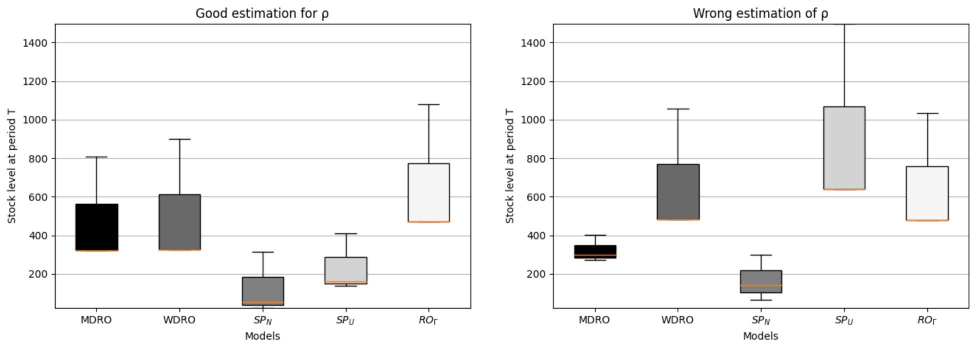
Table 2 reports the numerical results in an adverse production environment represented by the out-of-sample simulation. In this case, the production plan given by SP_N follows the same distribution as the simulation, while SP_U represents the case of misspecification of the production yield distribution in the optimization stage. Although the expected value obtained with SP_N is 0.4% (resp. 3.6%) lower than the respective costs obtained with $WDRO$ (resp. DRO), $WDRO$ (resp. $MDRO$) reduces the average expected cost by 3.4% (resp. 0.1%) compared to the RO plan, and by 6.9% (resp. 3.5%) compared to SP_U. In addition, $WDRO$ (resp. $MDRO$) reduces the 95^{th} percentile costs by 3% and 6.2% (resp. 0.1% and 3.1%) compared to RO and SP_U, respectively, with a reduction in terms of 99^{th} percentile costs of about 2.9% and 5.9% (resp. 0.1% and 3%). In the same line, $WDRO$ reduces the 95^{th} and 99^{th} percentile costs by 1.6% and 2.6%, respectively, when compared to SP_N. Since $WDRO$ is the more conservative DRO plan, it proposes a feasible plan in the worst-case perspective of the distribution. That justifies the overly conservatism of $MDRO$ compared to SP_N, yielding to higher 95^{th} and 99^{th} percentile costs, with an increase of about 1.3% and 0.2%, respectively. These results confirm the robustness and effectiveness of the DRO models, especially the $WDRO$ model to mitigate uncertainties and minimize the overall costs in a fuzzy or unexpected environment simulated with the out-of-sample framework.

Table 2 – Performance of the models when the simulation follows a normal distribution

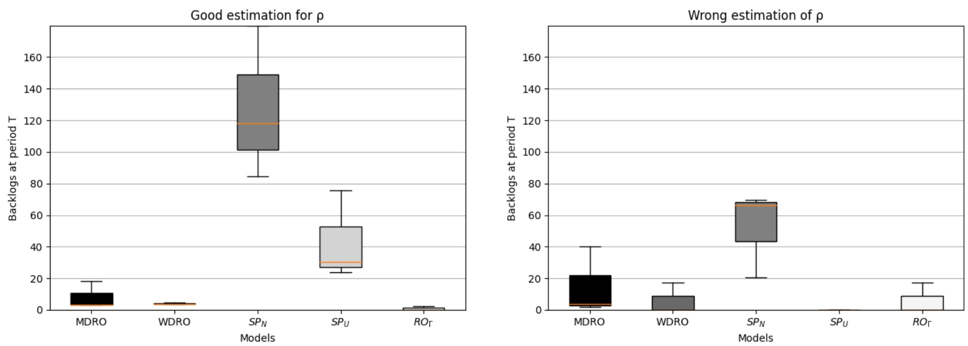
Model	Exp. Cost	95^{th} p.c.	99^{th} p.cc	Time	CV
$WDRO$	1,750,526	1,830,390	1,850,714	0.36	2.55%
$MDRO$	1,808,320	1,884,443	1,902,370	0.47	2.47%
SP_N	1,743,162	1,859,418	1,898,707	0.13	3.91%
SP_U	1,871,989	1,943,210	1,959,890	0.10	2.31%
RO	1,809,762	1,885,929	1,903,950	0.01	2.47%

Figure 3 reports the inventory and backorder levels of the production plans at the end of the production horizon, for each method throughout the simulation. It shows that a robust strategy generally results in the largest stock to protect against yield uncertainties to avoid stockouts. On the other hand, the stochastic programs keep fewer items in stock compared to the DRO and RO. Compared to RO, Figure 3 shows that DRO models achieve a better balance in inventory management activity. As $WDRO$ reduces the conservatism of a robust plan by considering the available distributional information, it leads to a lower inventory level, which implies a small increase in stockouts compared to RO. However, like $MDRO$, it achieves a good balance in inventory management cost, which leads to lower average costs despite a slightly higher amount of stockouts compared to RO. Figure 3 also shows that SP models increase the risks of stock-outs to reduce the inventory and production costs, while DRO plans produce a higher quantity of products which leads to higher demand satisfaction and lower backorders. In addition, since fewer items are kept in stock, or suffer from stockouts, in a case of misspecification of yield uncertain distribution (which is observed by the low variability in the inventory and backorder levels at the end of the production horizon), the strategy adopted by the DRO models achieves good cost savings.

Figure 3 shows that DRO is less impacted by misestimation and variabilities in yields compared to the stochastic programs. DRO models have larger lot sizes, which implies a larger stock of goods and backorder levels as sporadic and as low as possible. DRO models and RO models have the lowest



(a) Stock level at the end of the production horizon



(b) Stockout at the end of the production horizon

Figure 3 – Inventory levels of the simulated production plans

backorder levels and the largest stock of goods, even when the actual yield realizations are significantly different from those extracted from the data. DRO models propose a strategy that balances between robust optimization and stochastic programming methods, which generally leads to a higher stock level than the solutions from SP and makes the occurrence of backorders more frequent than the solutions from the traditional RO framework. In addition, DRO models remain robust despite the changes in the uncertainty yield distribution.

Since DRO models use data-driven ambiguity sets, they provide more robust and cost-effective production plans. Contrarily to SP, DRO models are not sensitive to the risk of misestimation of the probability, since they provide a plan that performs better for contexts not considered within the stochastic program formulation. Therefore, DRO solutions are still robust when the production yield suddenly changes behavior. Therefore, the experimental results indicate that DRO plans have a better cost-cutting strategy compared with robust optimization or stochastic program solutions, while they are more immunized from errors in predictions.

4.4 Case study

The case study relies on data provided by a manufacturer of industrial equipment. We consider a specific factory that produces a large number of variants of a sensitive part required in the finished product. These parts have a very strict tolerance to ensure the longevity of the equipment sold to the customer. To avoid deviations from specifications, the manufacturer currently tests all parts. When a part does not meet the specifications, it is most often discarded. A discarded part cannot be replaced by the production of a new one, because it would require ordering material from suppliers. The material is specific to each part, and because the manufacturer produces a large number of variants, it does not keep an unnecessary amount of material in stock.

In this case study, we have the historical production yield rate data for 2 products in four different scenarios ($S = 4$), namely : s_1 when a machine is down ; s_2 when a machine has a low battery ; s_3 when the temperature is normal ; and s_4 when the temperature is low. We consider only the $MDRO$ since it can deal with these four different scenarios considering their conditional information independently. Therefore, DRO considers the mean absolute ambiguity set. From the available data, for each scenario $s \in S$ we gather the available sample set which allows us to measure the average production yield $\bar{\rho}_{its}$, the standard deviation of the production yield $\hat{\rho}_{its}$ for all item $i \in N$ and period t. These data become the parameters of each scenario in our scenario-wise ambiguity set for $MDRO$. In addition, we compute the upper and lower output yield limits given by $\bar{\bar{\rho}}_{its}$ and $\underline{\underline{\rho}}_{its}$, for all elements $i \in N$, in period t. These limits define the uncertainty set for the RO model and the support of the uniform probability distribution for the SP model. Note that the uncertain yield in the three models (RO, SP, and $MDRO$) has a bounded support that corresponds to the minimum and maximum measure of production yields from the available data. To facilitate the analysis, we simulate the different models in a space close to a uniform distribution. We compare the performance of the models with the simulation results. We follow the same simulation process as presented in Section 4.1. We solve each model to optimality and simulate 1000 scenarios to analyze the performance of the obtained production plans.

Table 3 reports the costs resulting from the simulation. In the table are given the overall costs, standard deviation, and coefficient variation as well as computational time for 20 instances with 2 items and 12 periods. We used the same table structure presented in Section 4.3, with which the models are compared based on the average computational time from the optimization, and the expected value along with the 95^{th} and 99^{th} percentile cost (p.c.) and the coefficient of variation CV of the costs evaluated in the simulation framework.

Table 3 – Performance of the models in a real application

Model	Exp. Cost	95^{th} p.c.	99^{th} p.cc	Time	CV
DRO	3,702,977	4,256,385	4,595,437	3.62	10%
SP	3,741,098	4,707,514	5,213,221	0.84	21%
RO	4,257,385	4,639,200	4,754,106	0.02	8%

Table 3 shows that DRO models outperform the stochastic program and the robust optimization models for all criteria but the computational time, and the coefficient of variation for RO. Although RO models are solved faster, DRO leads to more significant cost savings (since the expected costs and even the 99^{th} percentile costs with DRO are lower than the costs with RO) and higher robustness (since its coefficient of variation is the lowest) in an average situation. The expected cost of DRO represents a cost saving of 38.1K compared to SP, and around 554.4K compared to RO, which corresponds to a decrease of approximately 1% and 15% respectively in the average costs. DRO also has a reduction of 11% and 9% (which correspond to a cost saving of approximately 451.1K and 382.8K) on the 95^{th} percentile costs when compared to RO and SP respectively. In terms of 99^{th} percentile costs, the reduction on the average cost corresponds to 13% (about 617.8K) and 3% (about 158.7K) of the average costs compared to RO and SP, respectively. Regarding the computational time, Table 3 shows that the solver solves all models in less than a second. Therefore, scalability is not an issue. In addition, the coefficient of variation of DRO is 10%, while the respective CV for the robust model is 8% and 21% for the SP model. This is not surprising, as a distribution-dependent model (SP) is more subject to large variations due to poor uncertainty estimation while a fully robust model ($R0$) seeks to immunize the system from all types of uncertainties and with a feasible solution that remains robust to any realization of uncertainty. As a result, the DRO model leads to a more cost-effective strategy and a lower risk sensitivity compared to models highly dependent on good estimation of uncertainties or highly conservative models.

Table 4 reports the characteristics of the simulated production plans. Column $\|X\|$ gives the number of items produced, while column $\|I\|$ (resp. $\|B\|$) shows the cumulative stock (resp. backorder) over the production planning, column $\|I_T\|$ (resp. column $\|B_T\|$) gives the number of items kept on stock

Table 4 – Characteristic of the production plans

Model	$\|X\|$	$\|I\|$	$\|B\|$	$\|I_T\|$	$\|B_T\|$	$\%\|Y\|$
DRO	52,758	35,787	1,013	2,237	110	64%
SP	51,887	27,794	1,415	1,565	223	72%
RO	57,021	50,535	555	5,844	-	75%

(resp. stockouts) at the end of the production horizon, and, finally, column $\|Y\|$ gives the frequency of setup over the entire production horizon. DRO produces more than SP, which reduces the stockout at the end of the production horizon, and it maintains a larger level of stocks over the production horizon to balance the inventory management activities and to mitigate unexpected costs due to malfunctions in quality assurance. Compared to RO, DRO reduces the amount to produce, which leads to a lower inventory level at the expense of the possibility of stockouts (which is completely avoided at the end of the production horizon with the plan given by RO). However, DRO plans still provide better inventory flow than RO as DRO minimizes total costs with a balance between maintaining an inventory level and accepting stockouts. Therefore, it is clear that DRO adopts a strategy to maintain a sufficient amount of inventory to avoid cost overruns due to stockouts and to be able to satisfy all demands regardless of dysfunctions in production performance.

5 Conclusion

This paper introduces a distributionally robust formulation for multi-item multi-period lot-sizing under yield uncertainty. We consider the case where the probability distribution of the uncertain yield takes value in a scenario-wise ambiguity set, and we reformulate the problem as a MILP. The resulting model can easily be solved with a MILP solver. Our experimental results show that the distributionally robust LSP model provides sufficiently robust production plans. These plans combine a good cost-cutting strategy and a lower risk sensitivity to the variation of the production yield. Other advantages of the event-wise distributionally robust optimization over the other considered approaches are that the construction of the ambiguity set is data-driven, and it is free from strong assumptions about uncertain parameter patterns. Further investigation is still needed to improve the quality of the distributionally robust formulations and to reduce their sensitivity to large disturbances on the uncertain production yield for adverse scenarios. In addition, the distributionally robust models can suffer from scalability issues for very large instances and large ambiguity sets (i.e. when we increase the number of scenarios and the size of the production horizon of the number of items to be processed). Therefore, addressing large-size instances is envisaged, where a decomposition approach could help modelers compute good distributionally robust production plans and obtain better bounds with less computational effort. In addition, an extension of this distributionally robust model with a clustering-based ambiguity set is envisaged to bring the proposed model closer to real-world applications and cases where a probability distribution is difficult to estimate or is correlated for different items.

A Supplementary material

A.1 Counterpart reformulation for DRLSP constraints

In this section, we present the reformulation of the remaining constraints subject to uncertainties from problem (4) in a robust fashion. For that, we assume that Slater's condition and the strong duality hold on each constraint, which helps us define the constraint in a robust counterpart form.

A.1.1 Reformulation for constraints (3b)

If Slater's condition holds on constraints (3b), then constraints (3b) is equivalent to

$$H^0_{its} + h_{it}d_{i\tau} \geq \sup_{\Lambda} \boldsymbol{\rho}^\top \boldsymbol{c}^3_{its} + \boldsymbol{m}^\top \boldsymbol{c}^4_{its} \qquad \begin{array}{r}\forall(\boldsymbol{\rho},\boldsymbol{m}) \in \mathcal{W}_s;\\ i \in N; t \in T;\\ s \in S\end{array}$$

where $\Lambda = \boldsymbol{\rho}, \boldsymbol{m}, \{\boldsymbol{\xi}_j\}_{j\in(N\times T)_{|m|}}, \boldsymbol{\rho} \in [\underline{\underline{\boldsymbol{\rho}}}, \overline{\overline{\boldsymbol{\rho}}}], g_{js}(\xi_j) \leq m_j, (\boldsymbol{\xi}_j) = \rho,\ \forall j \in (N \times T)_{|m|}$. Note that $\boldsymbol{c}^3_{its}$ is given in (7) and $c^4_{itjs} = -H^{''}_{itjs}$.

$$c^3_{itkls} = \begin{cases} h_{it}X_{il} - H^{'}_{itils}, & \text{if } i = k, l \leq t \\ -H^{'}_{itils}, & \text{otherwise} \end{cases} \tag{7}$$

If we reformulate the supremum in terms of the dual infimum for a scenario $s \in S$, we obtain :

$$\sup_{\Lambda} \boldsymbol{\rho}^\top \boldsymbol{c}^3_{its} + \boldsymbol{m}^\top \boldsymbol{c}^4_{its} = \inf_{\Lambda'} \overline{\overline{\boldsymbol{\rho}}}^\top \boldsymbol{\eta}^3_{its} - \underline{\underline{\boldsymbol{\rho}}}^\top \boldsymbol{\eta}^4_{its} + \sum_{j\in(N\times T)_{|m|}} \lambda^2_{js} g^*_{js}\left(\frac{\boldsymbol{w}^2_{js}}{\lambda^2_{js}}\right)$$

where $\Lambda' = \boldsymbol{\eta}^3_{its}, \boldsymbol{\eta}^4_{its}, \boldsymbol{\lambda}^2_{its} \geq 0, \{\boldsymbol{w}^2_{js}\}_{j\in(N\times T)_{|m|}}, \boldsymbol{c}^4_{its} + \boldsymbol{\lambda}^2_{its} = 0, \sum_{j\in(N\times T)_{|m|}} \boldsymbol{w}^2_{js} = \boldsymbol{c}^3_{its} - \boldsymbol{\eta}^3_{its} + \boldsymbol{\eta}^4_{its}$. Note that $\boldsymbol{\xi}_j = \boldsymbol{\rho}$, $\boldsymbol{m} = \boldsymbol{g}_s(\boldsymbol{\rho}) + 1$, and $\lambda^2_{js} g^*_{js}\left(\frac{\boldsymbol{w}^2_{js}}{\lambda^2_{js}}\right)$ is the perspective function of the conjugate function of g^*_{js} on $(\boldsymbol{w}^2_{js}, \lambda^2_{js})$ for $j \in (N \times T)_{|m|}$. As a consequence, we can replace constraints (3b) with constraints (6g)–(6k) in our final reformulation.

A.1.2 Reformulation for constraints (3c)

If Slater's condition holds on constraints (3c), then constraints (3c) is equivalent to

$$H^0_{its} - b_{it}d_{i\tau} \geq \sup_{\Lambda} \boldsymbol{\rho}^\top \boldsymbol{c}^5_{its} + \boldsymbol{m}^\top \boldsymbol{c}^6_{its} \qquad \begin{array}{r}\forall(\boldsymbol{\rho},\boldsymbol{m}) \in \mathcal{W}_s;\\ i \in N; t \in T;\\ s \in S\end{array}$$

where $\Lambda = \boldsymbol{\rho}, \boldsymbol{m}, \{\boldsymbol{\xi}_j\}_{j\in(N\times T)_{|m|}}, \boldsymbol{\rho} \in [\underline{\underline{\boldsymbol{\rho}}}, \overline{\overline{\boldsymbol{\rho}}}], g_{js}(\xi_j) \leq m_j, (\boldsymbol{\xi}_j) = \rho,\ \forall j \in (N \times T)_{|m|}$. Note that $\boldsymbol{c}^5_{its}$ is given in (8) and $c^6_{itjs} = -H^{''}_{itjs}$.

$$c^5_{itkls} = \begin{cases} -b_{it}X_{il} - H^{'}_{itils}, & \text{if } i = k, l \leq t \\ -H^{'}_{itils}, & \text{otherwise} \end{cases} \tag{8}$$

If we reformulate the supremum in terms of the dual infimum for a scenario $s \in S$, we obtain :

$$\sup_{\Lambda} \boldsymbol{\rho}^\top \boldsymbol{c}^5_{its} + \boldsymbol{m}^\top \boldsymbol{c}^6_{its} = \inf_{\Lambda 4} \overline{\overline{\boldsymbol{\rho}}}^\top \boldsymbol{\eta}^5_{its} - \underline{\underline{\boldsymbol{\rho}}}^\top \boldsymbol{\eta}^6_{its} + \sum_{j\in(N\times T)_{|m|}} \lambda^3_{js} g^*_{js}\left(\frac{\boldsymbol{w}^3_{js}}{\lambda^3_{js}}\right)$$

where $\Lambda' = \boldsymbol{\eta}^5_{its}, \boldsymbol{\eta}^6_{its}, \boldsymbol{\lambda}^3_{its} \geq 0, \{\boldsymbol{w}^3_{js}\}_{j\in(N\times T)_{|m|}}, \boldsymbol{c}^6_{its} + \boldsymbol{\lambda}^3_{its} = 0, \sum_{j\in(N\times T)_{|m|}} \boldsymbol{w}^3_{js} = \boldsymbol{c}^5_{its} - \boldsymbol{\eta}^5_{its} + \boldsymbol{\eta}^6_{its}$. Note that $\boldsymbol{\xi}_j = \boldsymbol{\rho}$, $\boldsymbol{m} = \boldsymbol{g}_s(\boldsymbol{\rho}) + 1$, and $\lambda^3_{js} g^*_{js}\left(\frac{\boldsymbol{w}^3_{js}}{\lambda^3_{js}}\right)$ is the perspective function of the conjugate function of g^*_{js} on $(\boldsymbol{w}^3_{js}, \lambda^3_{js})$ for $j \in (N \times T)_{|m|}$. As a consequence we can reformulate constraints (3c) as constraints (6l)–(6p) in our final reformulation.

A.1.3 Reformulation for constraints (3d)

If Slater's condition holds on constraints (3d), then constraints (3d) is equivalent to

$$H^0_{its} \geq \sup_{\Lambda} \boldsymbol{\rho}^\top \boldsymbol{c}^7_{its} + \boldsymbol{m}^\top \boldsymbol{c}^8_{its} \qquad \forall(\boldsymbol{\rho}, \boldsymbol{m}) \in \mathcal{W}_s; \; i \in N; t \in T; \; s \in S$$

where $\Lambda = \boldsymbol{\rho}, \boldsymbol{m}, \{\boldsymbol{\xi}_j\}_{j\in(N\times T)_{|m|}}, \boldsymbol{\rho} \in [\underline{\underline{\boldsymbol{\rho}}}, \overline{\overline{\boldsymbol{\rho}}}], g_{js}(\xi_j) \leq m_j, (\boldsymbol{\xi}_j) = \rho, \;\; \forall j \in (N \times T)_{|m|}$. Note that $\boldsymbol{c}^7_{itkls} = H^{'}_{itkls}$ and $c^8_{itjs} = -H^{''}_{itjs}$.

If we reformulate the supremum in terms of the dual infimum for a scenario $s \in S$, we obtain :

$$\sup_{\Lambda} \boldsymbol{\rho}^\top \boldsymbol{c}^7_{its} + \boldsymbol{m}^\top \boldsymbol{c}^8_{its} = \inf_{\Lambda'} \overline{\overline{\boldsymbol{\rho}}}^\top \boldsymbol{\eta}^7_{its} - \underline{\underline{\boldsymbol{\rho}}}^\top \boldsymbol{\eta}^8_{its} + \sum_{j\in(N\times T)_{|m|}} \lambda^4_{js} g^*_{js}\Big(\frac{\boldsymbol{w}^4_{js}}{\lambda^4_{js}}\Big)$$

where $\Lambda' = \boldsymbol{\eta}^7_{its}, \boldsymbol{\eta}^8_{its}, \boldsymbol{\lambda}^4_{its} \geq 0, \{\boldsymbol{w}^4_{js}\}_{j\in(N\times T)_{|m|}}, \boldsymbol{c}^8_{its} + \boldsymbol{\lambda}^4_{its} = 0, \sum_{j\in(N\times T)_{|m|}} \boldsymbol{w}^4_{js} = \boldsymbol{c}^7_{its} - \boldsymbol{\eta}^7_{its} + \boldsymbol{\eta}^8_{its}$. Note that $\boldsymbol{\xi}_j = \boldsymbol{\rho}$, $\boldsymbol{m} = \boldsymbol{g}_s(\boldsymbol{\rho}) + 1$, and $\lambda^4_{js} g^*_{js}\Big(\frac{\boldsymbol{w}^4_{js}}{\lambda^4_{js}}\Big)$ is the perspective function of the conjugate function of g^*_{js} on $(\boldsymbol{w}^4_{js}, \, \lambda^4_{js})$ for $j \in (N \times T)_{|m|}$. As a consequence, we can reformulate constraints (3d) as constraints (6q)–(6s) in our final reformulation.

A.2 The robust LSP with yield uncertainty model

Based on the robust single-item LSP with uncertain production yield proposed by [31], the multi-item lot-sizing problem under yield uncertainty is given as follows :

$$\begin{aligned}
&\min \sum_{i\in N} \sum_{t\in T} s_{it} Y_t + v_{it} X_{it} + H_{it} && \\
&s.t. : && \\
&H_{it} \geq h_{it} \max_{\widetilde{\rho}\in\mathcal{U}_{it}} \left[\sum_{\tau=1}^{t} (\widetilde{\rho}_{i\tau} X_{i\tau} - d_{i\tau})\right] && \forall t \in T; i \in N \\
&H_t \geq -b_{it} \max_{\widetilde{i\rho}\in\mathcal{U}_{it}} \left[\sum_{\tau=1}^{t} (\widetilde{i\rho}_{\tau} X_{i\tau} - d_{i\tau})\right] && \forall t \in T; i \in N \\
&X_{it} \leq M_{it} Y_{it} && \forall t \in T; i \in N \\
&X_{it}, H_{it} \geq 0 && \forall t \in T; i \in N \\
&Y_{it} \in \{0, 1\} && \forall t \in T; i \in N
\end{aligned}$$

where $\tilde{\boldsymbol{\rho}}$ mapped by an affine rule that bounds its realization to a range centered on its nominal value $\bar{\boldsymbol{\rho}}$ and spread by its maximum deviation $\hat{\boldsymbol{\rho}}$. The uncertain production yield is represented in the budgeted uncertainty set $\mathcal{U}_{it}$ where a budget Γ controls the size of the uncertainty set according to the decision maker's sensitivity to risk. Thus, for each item i in each period t, the uncertain production yield $\tilde{\rho}_{it}$ belongs to $\mathcal{U}_{it}$ that is given by $\mathcal{U}_{it} = \{\mathbf{-1} \leq \boldsymbol{Z}^t_i \leq \mathbf{1} : \sum_{\tau=1}^{t} |Z^t i\tau| \leq \Gamma_t\}$.

A.3 The stochastic programming LSP with yield uncertainty model

Similar to the model proposed in the appendix from [31], we propose a scenario-based stochastic program to represent the multi-item lot-sizing problem under yield uncertainty. To represent the static strategy, we rely on the two-stage formulation, where only the inventory and backorder levels react to the different scenarios.

We consider a set Ω of possible yield scenarios, where each scenario ω has a probability p_ω of realization. ρ^ω_{it} is the realization of the uncertain yield for item i in period t and scenario ω. The two-stage stochastic program for the LSP with uncertain yield is given as follows :

$$
\begin{aligned}
&\min \sum_{\omega\in\Omega} p_\omega \sum_{t\in T}\sum_{i\in N} s_{it}Y_{it} + v_{it}X_{it} + h_{it}I^\omega_{it} + b_{it}B^\omega_{it} \\
&s.t. : \\
&I^\omega_{it} - B^\omega_{it} = I^\omega_{it-1} - B^\omega_{it-1} + \rho^\omega_{it}X_{it} - d_{it} && \forall t\in T; i\in N; s\in\Omega \\
&X_{it} \le M_{it}Y_{it} && \forall t\in T; i\in N \\
&X_{it}, I^\omega_t, B^\omega_t \ge 0 && \forall t\in T; i\in N; s\in\Omega \\
&Y_{it}\in\{0,1\} && \forall t\in T; i\in N
\end{aligned}
$$

To better represent the realization of the uncertain yield, we need to generate as many scenarios as possible. However, to avoid the drawbacks of scalability issues, such as prohibitive computational time, we generate the scenarios with a Monte Carlo approach. Thus, the yield rate is randomly drawn from a uniform distribution with support $[\bar\rho_{it} - \hat\rho_{it}; \bar\rho_{it} + \hat\rho_{it}]$ over 200 scenarios.

A.4 Mixed-Integer Mean Absolute distributionally robust formulation

This section presents the mean absolute distributionally robust formulation (MDRLSP), where the mean absolute ambiguity set $\mathscr{F}_M$ given in Section 3.1.1 replaces the general scenario-wise ambiguity set $\mathscr{F}$ in problem (6). As the support function for $\mathscr{F}_M$ is given by $\delta^*\Big((\boldsymbol{z}_\rho, \boldsymbol{z}_m)|\mathcal{Q}_k\Big) = \bar{\boldsymbol{\rho}}_s^\top \boldsymbol{z}_\rho + \hat{\boldsymbol{\rho}}_s^\top \boldsymbol{z}_m$, MDRLSP in problem (9) :

$$
\begin{aligned}
&\min \sum_{i\in N}\sum_{t\in T}\Big(s_{it}Y_{it} + v_{it}X_{it}\Big) + \gamma \\
&s.t. \\
&\gamma \ge \frac{1}{S}\sum_{s\in S}\sum_{i\in N}\sum_{t\in T}\bar\rho_{its}\beta'_{its} + \frac{1}{S}\sum_{s\in S}\sum_{i\in N}\sum_{t\in T}\hat\rho_{its}\beta''_{its} + \frac{1}{S}\sum_{s\in S}\alpha_s && && \text{(9a)} \\
&\alpha_s \ge \sum_{i\in N}\sum_{t\in T}\Big(\overline{\overline{\rho}}_{it}\eta^1_{its} - \underline{\underline{\rho}}_{it}\eta^2_{its}\Big) - \sum_{i\in N}\sum_{t\in T}(\eta^1_{its} - \eta^2_{its})\bar\rho_{its} \\
&\qquad + \sum_{i,k\in N}\sum_{l,t\in T} H'_{itkls}\bar\rho_{its} + \sum_{i\in N}\sum_{t\in T}\big(H^0_{its} - \beta'_{its}\bar\rho_{its}\big) && \forall s\in S && \text{(9b)} \\
&w^1_{kls} \le -\sum_{i\in N}\sum_{t\in T} H''_{itkls} + \beta''_{kls} && \forall k\in N; l\in T; s\in S && \text{(9c)} \\
&-w^1_{kls} \le \sum_{i\in N}\sum_{t\in T} H'_{itkls} + \eta^1_{kls} - \eta^2_{lks} && \forall k\in N; l\in T; s\in S && \text{(9d)} \\
&-w^1_{kls} \le \sum_{i\in N}\sum_{t\in T} H'_{itkls} - \eta^1_{kls} + \eta^2_{lks} && \forall k\in N; l\in T; s\in S && \text{(9e)} \\
&H^0_{its} \ge -h_{it}\sum_{l=1}^{t}(d_{i\tau})\sum_{k=1}^{N}\sum_{l=1}^{T}\overline{\overline{\rho}}_{kl}\eta^3_{itkls} - \sum_{k=1}^{N}\sum_{l=1}^{T}\underline{\underline{\rho}}_{kl}\eta^4_{itkls} \\
&\qquad - \sum_{k=1}^{N}\sum_{l=1}^{T} H'_{itkls}\bar\rho_{kls} - \sum_{k=1}^{N}\sum_{l=1}^{T}\eta^3_{itkls}\bar\rho_{kls} && \forall i\in N; t\in T; s\in S && \text{(9f)} \\
&\qquad + \sum_{k=1}^{N}\sum_{l=1}^{T}\eta^4_{itkls}\bar\rho_{kls} + \sum_{l=1}^{t} h_{it}X_{il}\bar\rho_{ils} \\
&w^2_{itkls} - H''_{itkls} \le 0 && \forall i,k\in N; t,l\in T; s\in S && \text{(9g)} \\
&-H'_{itkls} \le w^2_{itkls} + \eta^3_{itkls} - \eta^4_{itkls} && \forall i,k\in N; k\ne i; t,l\in T; s\in S && \text{(9h)}
\end{aligned}
$$

$$-H'_{itils} \leq w^2_{itils} + \eta^3_{itils} - \eta^4_{itils} \qquad \forall i \in N; t,l \in T; l > t; s \in S \tag{9i}$$

$$-H'_{itils} \leq w^2_{itils} + \eta^3_{itils} - \eta^4_{itils} - h_{it}X_{il} \qquad \forall i \in N; t,l \in T; l \leq t; s \in S \tag{9j}$$

$$H'_{itkls} \leq w^2_{itkls} - \eta^3_{itkls} + \eta^4_{itkls} \qquad \forall i,k \in N; k \neq i; t,l \in T; s \in S \tag{9k}$$

$$H'_{itils} \leq w^2_{itils} - \eta^3_{itils} + \eta^4_{itils} \qquad \forall i \in N; t,l \in T; l > t; s \in S \tag{9l}$$

$$+H'_{itils} \leq w^2_{itils} - \eta^3_{itils} + \eta^4_{itils} + h_{it}X_{il} \qquad \forall i \in N; t,l \in T; l \leq t; s \in S \tag{9m}$$

$$\begin{aligned} H^0_{its} \geq\ & b_{it}\sum_{l=1}^{t}(d_{i\tau}) + \sum_{k=1}^{N}\sum_{l=1}^{T}\overline{\overline{\rho}}_{kl}\eta^5_{itkls} - \sum_{k=1}^{N}\sum_{l=1}^{T}\underline{\underline{\rho}}_{kl}\eta^6_{itkls} \\ & - \sum_{k=1}^{N}\sum_{l=1}^{T}H'_{itkls}\bar{\rho}_{kls} - \sum_{k=1}^{N}\sum_{l=1}^{T}\eta^5_{itkls}\bar{\rho}_{kls} \\ & + \sum_{k=1}^{N}\sum_{l=1}^{T}\eta^6_{itkls}\bar{\rho}_{kls} - \sum_{l=1}^{t}(b_{it}X_{il})\bar{\rho}_{ils} \end{aligned} \qquad \forall i \in N; t \in T; s \in S \tag{9n}$$

$$w^3_{itkls} - H''_{itkls} \leq 0 \qquad \forall i,k \in N; t,l \in T; s \in S \tag{9o}$$

$$-H'_{itkls} \leq w^3_{itkls} + \eta^5_{itkls} - \eta^6_{itkls} \qquad \forall i,k \in N; k \neq i; t,l \in T; s \in S \tag{9p}$$

$$-H'_{itils} \leq w^3_{itils} + \eta^5_{itils} - \eta^6_{itils} \qquad \forall i \in N; t,l \in T; l > t; s \in S \tag{9q}$$

$$-w^3_{itils} \leq -H'_{itils} + \eta^5_{itils} - \eta^6_{itils} + b_{it}X_{il} \qquad \forall i \in N; t,l \in T; l \leq t; s \in S \tag{9r}$$

$$H'_{itkls} \leq w^3_{itkls} - \eta^5_{itkls} + \eta^6_{itkls} \qquad \forall i,k \in N; k \neq i; t,l \in T; s \in S \tag{9s}$$

$$H'_{itils} \leq w^3_{itils} - \eta^5_{itils} + \eta^6_{itils} \qquad \forall i \in N; t,l \in T; l > t; s \in S \tag{9t}$$

$$-w^3_{itils} \leq -H'_{itils} - \eta^5_{itils} + \eta^6_{itils} - b_{it}X_{il} \qquad \forall i \in N; t,l \in T; l \leq t; s \in S \tag{9u}$$

$$\begin{aligned} H^0_{its} \geq\ & \sum_{k=1}^{i-1}\sum_{l=1}^{T}\overline{\overline{\rho}}_{kl}\eta^7_{itkls} - \sum_{k=1}^{i-1}\sum_{l=1}^{T}\underline{\underline{\rho}}_{kl}\eta^8_{itkls} - \sum_{k=1}^{i-1}\sum_{l=1}^{T}H'_{itkls}\bar{\rho}_{kls} \\ & - \sum_{k=1}^{i-1}\sum_{l=1}^{T}\eta^7_{itkls}\bar{\rho}_{kls} + \sum_{k=1}^{i-1}\sum_{l=1}^{T}\eta^8_{itkls}\bar{\rho}_{kls} \end{aligned} \qquad \forall i \in N; t \in T; s \in S \tag{9v}$$

$$w^4_{itkls} - H''_{itkls} \leq 0 \qquad \forall i,k \in N; t,l \in T; s \in S \tag{9w}$$

$$-H'_{itkls} \leq w^4_{itkls} + \eta^7_{itkls} - \eta^8_{itkls} \qquad \forall i,k \in N; t,l \in T; s \in S \tag{9x}$$

$$H'_{itkls} \leq w^4_{itkls} - \eta^7_{itkls} + \eta^8_{itkls} \qquad \forall i,k \in N; t,l \in T; s \in S \tag{9y}$$

$$X_{it} \leq M_{it}Y_{it} \qquad \forall i \in N; t \in T \tag{9z}$$

$$X_{it}, \boldsymbol{\eta}^1_s, \boldsymbol{\eta}^2_s, \boldsymbol{\eta}^3_{its}, \boldsymbol{\eta}^4_{its}, \boldsymbol{\eta}^5_{its}, \boldsymbol{\eta}^6_{its}, \boldsymbol{\eta}^7_{its}, \boldsymbol{\eta}^8_{its} \geq 0 \qquad \forall i \in N; t \in T; s \in S$$

$$H^0_{its}, \boldsymbol{H}'_{its}, \boldsymbol{H}''_{its}, \boldsymbol{w}^1_s \in \mathbb{R}, \boldsymbol{w}^2_{its}, \boldsymbol{w}^3_{its}, \boldsymbol{w}^4_{its} \in \mathbb{R}, \gamma, \boldsymbol{\beta} \in \mathbb{R}^{(N\times T)}, \boldsymbol{\alpha} \in \mathbb{R}^S \qquad \forall i \in N; t \in T; s \in S$$

$$Y_{it} \in \{0,1\} \qquad \forall i \in N; t \in T$$

A.5 Mixed-Integer Wasserstein distributionally robust formulation

If the Wasserstein ambiguity set $\mathscr{F}_W$ given in Section 3.1.2 replaces the general scenario-wise ambiguity set $\mathscr{F}$ in problem (6), we obtain the Wasserstein distributionally robust formulation (WDRLSP). The support function for $\mathscr{F}_W$ is given by

$$\delta^*\Big((\boldsymbol{z}_\rho, z_m)|\mathcal{Q}_k\Big) = \begin{cases} \theta z_m, \text{ if } \boldsymbol{z}_\rho = \boldsymbol{0} \\ \infty, \text{ otherwise} \end{cases}$$

As a result, the WDRLSP is given in (10) :

$$\min \sum_{i\in N}\sum_{t\in T}\Big(s_{it}Y_{it} + v_{it}X_{it}\Big) + \gamma$$

$s.t.$

$$\gamma \geq \frac{1}{S}\sum_{s\in S} + \theta\beta \tag{10a}$$

$$\alpha_s \geq \sum_{i\in N}\sum_{t\in T}\left(\overline{\overline{\rho}}_{it}\eta^1_{its} - \underline{\underline{\rho}}_{it}\eta^2_{its}\right) - \sum_{i\in N}\sum_{t\in T}\left(\eta^1_{its} - \eta^2_{its}\right)\hat{\rho}_{its} + \sum_{i,k\in N}\sum_{l,t\in T} H'_{itkls}\hat{\rho}_{its} + \sum_{i\in N}\sum_{t\in T} H^0_{its} \qquad \forall s\in S \tag{10b}$$

$$w^1_{kls} \leq -\sum_{i\in N}\sum_{t\in T} H''_{its} + \beta \qquad \forall s\in S \tag{10c}$$

$$-w^1_{kls} \leq \sum_{i\in N}\sum_{t\in T} H'_{itkls} + \eta^1_{kls} - \eta^2_{lks} \qquad \forall k\in N; l\in T; s\in S \tag{10d}$$

$$-w^1_{kls} \leq \sum_{i\in N}\sum_{t\in T} H'_{itkls} - \eta^1_{kls} + \eta^2_{lks} \qquad \forall k\in N; l\in T; s\in S \tag{10e}$$

$$H^0_{its} \geq -h_{it}\sum_{l=1}^{t}(d_{i\tau})\sum_{k=1}^{N}\sum_{l=1}^{T}\overline{\overline{\rho}}_{kl}\eta^3_{itkls} - \sum_{k=1}^{N}\sum_{l=1}^{T}\underline{\underline{\rho}}_{kl}\eta^4_{itkls} - \sum_{k=1}^{N}\sum_{l=1}^{T} H'_{itkls}\hat{\rho}_{kls} - \sum_{k=1}^{N}\sum_{l=1}^{T}\eta^3_{itkls}\hat{\rho}_{kls} + \sum_{k=1}^{N}\sum_{l=1}^{T}\eta^4_{itkls}\hat{\rho}_{kls} + \sum_{l=1}^{t} h_{it}X_{il}\hat{\rho}_{ils} \qquad \forall i\in N; t\in T; s\in S \tag{10f}$$

$$H^0_{its} \geq b_{it}\sum_{l=1}^{t}(d_{i\tau}) + \sum_{k=1}^{N}\sum_{l=1}^{T}\overline{\overline{\rho}}_{kl}\eta^5_{itkls} - \sum_{k=1}^{N}\sum_{l=1}^{T}\underline{\underline{\rho}}_{kl}\eta^6_{itkls} - \sum_{k=1}^{N}\sum_{l=1}^{T} H'_{itkls}\hat{\rho}_{kls} - \sum_{k=1}^{N}\sum_{l=1}^{T}\eta^5_{itkls}\hat{\rho}_{kls} + \sum_{k=1}^{N}\sum_{l=1}^{T}\eta^6_{itkls}\hat{\rho}_{kls} - \sum_{l=1}^{t}(b_{it}X_{il})\hat{\rho}_{ils} \qquad \forall i\in N; t\in T; s\in S \tag{10g}$$

$$w^2_{its} - H''_{its} \leq 0 \qquad \forall i\in N; t\in T; s\in S \tag{10h}$$

$$-H'_{itkls} \leq w^2_{its} + \eta^3_{itkls} - \eta^4_{itkls} \qquad \forall i,k\in N; k\neq i; t,l\in T; s\in S \tag{10i}$$

$$-H'_{itils} \leq w^2_{its} + \eta^3_{itils} - \eta^4_{itils} \qquad \forall i\in N; t,l\in T; l>t; s\in S \tag{10j}$$

$$-H'_{itils} \leq w^2_{its} + \eta^3_{itils} - \eta^4_{itils} - h_{it}X_{il} \qquad \forall i\in N; t,l\in T; l\leq t; s\in S \tag{10k}$$

$$H'_{itkls} \leq w^2_{its} - \eta^3_{itkls} + \eta^4_{itkls} \qquad \forall i,k\in N; k\neq i; t,l\in T; s\in S \tag{10l}$$

$$H'_{itils} \leq w^2_{its} - \eta^3_{itils} + \eta^4_{itils} \qquad \forall i\in N; t,l\in T; l>t; s\in S \tag{10m}$$

$$+H'_{itils} \leq w^2_{its} - \eta^3_{itils} + \eta^4_{itils} + h_{it}X_{il} \qquad \forall i\in N; t,l\in T; l\leq t; s\in S \tag{10n}$$

$$w^3_{its} - H''_{its} \leq 0 \qquad \forall i\in N; t\in T; s\in S \tag{10o}$$

$$-H'_{itkls} \leq w^3_{its} + \eta^5_{itkls} - \eta^6_{itkls} \qquad \forall i,k\in N; k\neq i; t,l\in T; s\in S \tag{10p}$$

$$-H'_{itils} \leq w^3_{its} + \eta^5_{itils} - \eta^6_{itils} \qquad \forall i\in N; t,l\in T; l>t; s\in S \tag{10q}$$

$$-w^3_{its} \leq -H'_{itils} + \eta^5_{itils} - \eta^6_{itils} + b_{it}X_{il} \qquad \forall i\in N; t,l\in T; l\leq t; s\in S \tag{10r}$$

$$H'_{itkls} \leq w^3_{its} - \eta^5_{itkls} + \eta^6_{itkls} \qquad \forall i,k\in N; k\neq i; t,l\in T; s\in S \tag{10s}$$

$$H'_{itils} \leq w^3_{its} - \eta^5_{itils} + \eta^6_{itils} \qquad \forall i\in N; t,l\in T; l>t; s\in S \tag{10t}$$

$$-w^3_{its} \leq -H'_{itils} - \eta^5_{itils} + \eta^6_{itils} - b_{it}X_{il} \qquad \forall i\in N; t,l\in T; l\leq t; s\in S \tag{10u}$$

$$H^0_{its} \geq \sum_{k=1}^{i-1}\sum_{l=1}^{T} \overline{\overline{\rho}}_{kl}\eta^7_{itkls} - \sum_{k=1}^{i-1}\sum_{l=1}^{T} \underline{\underline{\rho}}_{kl}\eta^8_{itkls} - \sum_{k=1}^{i-1}\sum_{l=1}^{T} H^{'}_{itkls}\hat{\rho}_{kls} - \sum_{k=1}^{i-1}\sum_{l=1}^{T} \eta^7_{itkls}\hat{\rho}_{kls} + \sum_{k=1}^{i-1}\sum_{l=1}^{T} \eta^8_{itkls}\hat{\rho}_{kls} \quad \forall i \in N; t \in T; s \in S \quad (10v)$$

$$w^4_{its} - H^{''}_{its} \leq 0 \quad \forall i \in N; t \in T; s \in S \quad (10w)$$

$$-H^{'}_{itkls} \leq w^4_{its} + \eta^7_{itkls} - \eta^8_{itkls} \quad \forall i,k \in N; t,l \in T; s \in S \quad (10x)$$

$$H^{'}_{itkls} \leq w^4_{its} - \eta^7_{itkls} + \eta^8_{itkls} \quad \forall i,k \in N; t,l \in T; s \in S \quad (10y)$$

$$X_{it} \leq M_{it}Y_{it} \quad \forall i \in N; t \in T \quad (10z)$$

$$X_{it}, \boldsymbol{\eta}^1_s, \boldsymbol{\eta}^2_s, \boldsymbol{\eta}^3_{its}, \boldsymbol{\eta}^4_{its}, \boldsymbol{\eta}^5_{its}, \boldsymbol{\eta}^6_{its}, \boldsymbol{\eta}^7_{its}, \boldsymbol{\eta}^8_{its} \geq 0 \quad \forall i \in N; t \in T; s \in S$$

$$H^0_{its}, \boldsymbol{H}^{'}_{its}, H^{''}_{its}, w^1_s \in \mathbb{R}, w^2_{its}, w^3_{its}, w^4_{its} \in \mathbb{R}, \gamma, \boldsymbol{\beta} \in \mathbb{R}^{(N\times T)}, \boldsymbol{\alpha} \in \mathbb{R}^S \quad \forall i \in N; t \in T; s \in S$$

$$Y_{it} \in \{0,1\} \quad \forall i \in N; t \in T$$

Références

[1] Mohamed Ali Aloulou, Alexandre Dolgui, and Mikhail Y Kovalyov. A bibliography of non-deterministic lot-sizing models. International Journal of Production Research, 52(8) :2293–2310, 2014.

[2] Edward Anderson and Marta Monjardino. Contract design in agriculture supply chains with random yield. European Journal of Operational Research, 277(3) :1072–1082, 2019.

[3] Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. Robust optimization. Princeton University Press, 2009.

[4] Dimitris Bertsimas and Melvyn Sim. The price of robustness. Operations Research, 52(1) :35–53, 2004.

[5] Dimitris Bertsimas, Melvyn Sim, and Meilin Zhang. Adaptive distributionally robust optimization. Management Science, 65(2) :604–618, 2019.

[6] John R Birge and Francois Louveaux. Introduction to stochastic programming. Springer Science & Business Media, 2011.

[7] Dimitris Boskos, Jorge Cortés, and Sonia Martínez. Data-driven ambiguity sets with probabilistic guarantees for dynamic processes. IEEE Transactions on Automatic Control, 66(7) :2991–3006, 2020.

[8] Nadjib Brahimi, Nabil Absi, Stéphane Dauzère-Pérès, and Atle Nordli. Single-item dynamic lot-sizing problems : An updated survey. European Journal of Operational Research, 263(3) :838–863, 2017.

[9] Paolo Brandimarte. Multi-item capacitated lot-sizing with demand uncertainty. International Journal of Production Research, 44(15) :2997–3022, 2006.

[10] MD Byrne. Multi-item production lot sizing using a search simulation approach. Engineering Costs and Production Economics, 19(1–3) :307–311, 1990.

[11] Zhi Chen, Melvyn Sim, and Peng Xiong. Robust stochastic optimization made easy with rsome. Management Science, 66(8) :3329–3339, 2020.

[12] Zhi Chen and Peng Xiong. Rsome in python : An open-source package for robust stochastic optimization made easy. INFORMS Journal on Computing, 2023.

[13] Soo-Haeng Cho. The optimal composition of influenza vaccines subject to random production yields. Manufacturing & Service Operations Management, 12(2) :256–277, 2010.

[14] Erick Delage and Yinyu Ye. Distributionally robust optimization under moment uncertainty with application to data-driven problems. Operations Research, 58(3) :595–612, 2010.

[15] Peyman Mohajerin Esfahani and Daniel Kuhn. Data-driven distributionally robust optimization using the wasserstein metric : Performance guarantees and tractable reformulations. Mathematical Programming, 171(1) :115–166, 2018.

[16] Zhenghao Fan and Xiaolei Xie. A distributionally robust optimisation for covid-19 testing facility territory design and capacity planning. International Journal of Production Research, 60(13) :4229–4252, 2022.

[17] Hailei Gong, Yanzi Zhang, and Zhi-Hai Zhang. Training demand prediction models by decision error for two-stage lot-sizing problems. IEEE Transactions on Automation Science and Engineering, 2023.

[18] Astghik Hakobyan and Insoon Yang. Wasserstein distributionally robust motion control for collision avoidance using conditional value-at-risk. IEEE Transactions on Robotics, 38(2) :939–957, 2021.

[19] Grani A Hanasusanto, Daniel Kuhn, Stein W Wallace, and Steve Zymler. Distributionally robust multi-item newsvendor problems with multimodal demand distributions. Mathematical Programming, 152(1) :1–32, 2015.

[20] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. Convex analysis and minimization algorithms I : Fundamentals, volume 305. Springer science & business media, 2013.

[21] Ripeng Huang, Shaojian Qu, Xiaoguang Yang, and Zhimin Liu. Multi-stage distributionally robust optimization with risk aversion. Journal of Industrial & Management Optimization, 17(1) :233, 2021.

[22] Karl Inderfurth and Stephanie Vogelgesang. Concepts for safety stock determination under stochastic demand and different types of random production yield. European Journal of Operational Research, 224(2) :293–301, 2013.

[23] Burak Kazaz. Production planning under yield and demand uncertainty with yield-dependent cost and price. Manufacturing & Service Operations Management, 6(3) :209–224, 2004.

[24] Masoumeh Kazemi Zanjani, Mustapha Nourelfath, and Daoud Ait-Kadi. A multi-stage stochastic programming approach for production planning with uncertainty in the quality of raw materials and demand. International Journal of Production Research, 48(16) :4701–4723, 2010.

[25] Oliver Kramer and Oliver Kramer. Scikit-learn. Machine learning for evolution strategies, pages 45–53, 2016.

[26] Daniel Kuhn, Peyman Mohajerin Esfahani, Viet Anh Nguyen, and Soroosh Shafieezadeh-Abadeh. Wasserstein distributionally robust optimization : Theory and applications in machine learning. In Operations research & management science in the age of analytics, pages 130–166. Informs, 2019.

[27] Yuchen Li, Francisco Saldanha-da Gama, Ming Liu, and Zaoli Yang. A risk-averse two-stage stochastic programming model for a joint multi-item capacitated line balancing and lot-sizing problem. European Journal of Operational Research, 304(1) :353–365, 2023.

[28] Aristidis Likas, Nikos Vlassis, and Jakob J Verbeek. The global k-means clustering algorithm. Pattern recognition, 36(2) :451–461, 2003.

[29] Linyu Liu, Shiji Song, Zhuolin Wang, and Yuli Zhang. Data-driven distributionally robust optimization for railway timetabling problem. IEEE Transactions on Automation Science and Engineering, 2022.

[30] Xin Liu, Feng Chu, Feifeng Zheng, Chengbin Chu, and Ming Liu. Distributionally robust and risk-averse optimisation for the stochastic multi-product disassembly line balancing problem with workforce assignment. International Journal of Production Research, 60(6) :1973–1991, 2022.

[31] Paula Metzker, Simon Thevenin, Yossiri Adulyasak, and Alexandre Dolgui. Robust optimization for lot-sizing problems under yield uncertainty. Computers & Operations Research, 149 :106025, 2023.

[32] DNP Murthy and L Ma. Mrp with uncertainty : a review and some extensions. International Journal of Production Economics, 25(1–3) :51–64, 1991.

[33] Yves Pochet and Laurence A Wolsey. Production planning by mixed integer programming. Springer Science & Business Media, 2006.

[34] Ruozhen Qiu, Yue Sun, and Minghe Sun. A distributionally robust optimization approach for multi-product inventory decisions with budget constraint and demand and yield uncertainties. Computers & Operations Research, 126 :105081, 2021.

[35] Franco Quezada, Céline Gicquel, Safia Kedad-Sidhoum, and Dong Quan Vu. A multi-stage stochastic integer programming approach for a multi-echelon lot-sizing problem with returns and lost sales. Computers & Operations Research, 116 :104865, 2020.

[36] Hamed Rahimian and Sanjay Mehrotra. Distributionally robust optimization : A review. arXiv preprint arXiv :1908.05659, 2019.

[37] Allen Z Ren and Anirudha Majumdar. Distributionally robust policy learning via adversarial environment generation. IEEE Robotics and Automation Letters, 7(2) :1379–1386, 2022.

[38] Herbert Scarf. A min-max solution of an inventory problem. Studies in the mathematical theory of inventory and production, 1958.

[39] Kseniya Schemeleva, Xavier Delorme, Alexandre Dolgui, and Frédéric Grimaud. Multi-product sequencing and lot-sizing under uncertainties : A memetic algorithm. Engineering Applications of Artificial Intelligence, 25(8) :1598–1610, 2012.

[40] Maurice Sion. On general minimax theorems. Pacific Journal of mathematics, 8(1) :171–176, 1958.

[41] Ilhem Slama, Oussama Ben-Ammar, Simon Thevenin, Alexandre Dolgui, and Faouzi Masmoudi. Stochastic program for disassembly lot-sizing under uncertain component refurbishing lead times. European Journal of Operational Research, 303(3) :1183–1198, 2022.

[42] Paula Metzker Soares, Simon Thevenin, Yossiri Adulyasak, and Alexandre Dolgui. Adaptive robust optimization for lot-sizing under yield uncertainty. European Journal of Operational Research, 2023.

[43] Horst Tempelmeier. A column generation heuristic for dynamic capacitated lot sizing with random demand under a fill rate constraint. Omega, 39(6) :627–633, 2011.

[44] Simon Thevenin, Yossiri Adulyasak, and Jean-françois Cordeau. Material requirements planning under demand uncertainty using stochastic optimization. Production and Operations Management, 30(2) :475–493, 2021.

[45] Simon Thevenin, Yossiri Adulyasak, and Jean-François Cordeau. Stochastic dual dynamic programming for multiechelon lot sizing with component substitution. INFORMS Journal on Computing, 34(6) :3151–3169, 2022.

[46] Simon Thevenin, Oussama Ben-Ammar, and Nadjib Brahimi. Robust optimization approaches for purchase planning with supplier selection under lead time uncertainty. European Journal of Operational Research, 303(3) :1199–1215, 2022.

[47] Harvey M Wagner and Thomson M Whitin. Dynamic version of the economic lot size model. Management Science, 5(1) :89–96, 1958.

[48] S Wang and E Delage. A column generation scheme for distributionally robust multi-item newsvendor problems. Les Cahiers du GERAD ISSN, 711 :2440, 2021.

[49] Zhuolin Wang, Keyou You, Shiji Song, and Yuli Zhang. Second-order conic programming approach for wasserstein distributionally robust two-stage linear programs. IEEE Transactions on Automation Science and Engineering, 19(2) :946–958, 2021.

[50] Wolfram Wiesemann, Daniel Kuhn, and Melvyn Sim. Distributionally robust convex optimization. Operations Research, 62(6) :1358–1376, 2014.

[51] Candace Arai Yano and Hau L Lee. Lot sizing with random yields : A review. Operations Research, 43(2) :311–334, 1995.