Stochastic programming in master production scheduling: overcoming barriers to industry application

Alexandre Forel, Martin Grunow

Technical University of Munich, Advanced Optimization in a Networked Economy (AdONE), Munich, Germany

Abstract

This paper stems from the observation that companies still rely on deterministic rolling-horizon planning despite a substantial body of literature on stochastic planning models. To foster practical applications, we identify barriers that limit the widespread application of stochastic programming in master production scheduling and develop a framework to overcome them. Our solutions include modelling uncertainty from available data, reflecting planning processes in the optimisation model and evaluating its performance accurately. A two-stage stochastic model with production recourse is introduced to improve planning flexibility, stability and communicability. It is applied on a real-world case study with large product portfolio, complex production processes and uncertain seasonal demand. Out-of-sample rolling-horizon simulations show that well-defined stochastic models can provide high demand satisfaction and low inventory costs while improving planning stability. In particular, planning nervousness can be reduced by 40% and raw-material nervousness by 80% compared to our industry partner's current production scheduling solution.

Keywords: stochastic programming, recourse, raw-material ordering, planning stability

1. Introduction

Even when demand is highly uncertain, companies still rely on deterministic rolling-horizon planning and rule-of-thumbs for safety-stock calculations in master production scheduling (Meistering and Stadtler, 2017). Yet, recent applications of stochastic programming have shown impressive results in controlled simulation environments (Gruson et al., 2021; Thevenin et al., 2021). Stochastic programming can accurately determine the volume and timing of safety stocks based on probabilistic uncertainty models. Further, they reflect the flexibility of rolling-horizon planning through recourse decisions that adapt to uncertainty as it unfolds. When demand is dynamic and forecasts have poor accuracy, planning flexibility is critical to ensure that demand can be met. Nonetheless, stochastic programming is far from widely applied in practice despite promising complementarity with rolling-horizon planning . This observation is especially surprising considering the breadth of existing research on stochastic programming. It suggests that existing models still contain important shortcomings that prevent their application. In particular, there appears to be a lack of discussion on how to translate models from academic settings, that rely on simplifying assumptions, to real-world problems and their complexity.

In this paper, we study how to overcome barriers facing practitioners when applying stochastic programming to master production scheduling. First, we identify barriers that still prevent the application of stochastic programming. We distinguish barriers relating to the identification and representation of uncertainty, relating to the development of stochastic planning models that fit existing planning structures, and relating to the computational challenges of evaluating model performance. Second, we propose a decision framework to set up stochastic models in master production scheduling. We present novel strategies to overcome the barriers such as aggregating products into optimal families, which increases planning stability and allows flexible production recourse. A two-stage stochastic model is developed to integrate the above strategies and determine a master production schedule that provides high demand satisfaction for low inventory costs and ensures planning stability on both the production and raw-material levels. We demonstrate our approach on a real-world case study in the agrochemical industry and evaluate its performance through out-of-sample rolling-horizon simulations.

The remainder of this paper is organised as follows. In Section 2, barriers limiting the application of stochastic planning models are presented and related to existing literature. In Section 3, the real-world case study is introduced. We discuss the specific form of the barriers identified in the previous section and we provide an overview of our strategies to overcome them. In Section 4, uncertainty models are derived from available historical data and used to construct scenario trees. In Section 5, stochastic planning models are developed to improve planning flexibility, stability and communicability. In Section 6, sensitivity analyses are conducted to tune the model and a comparison with industry benchmarks is presented. In Section 7, we summarise our work and propose directions for future research.

2. The barriers of applying stochastic programming in master production scheduling

This section details barriers that prevent the widespread application of stochastic planning models. We categorise the barriers in three groups relating to modelling uncertainty, the planning environment and the numerical challenges.

2.1. Modelling uncertainty from data

In practice, probability distributions are not available to describe demand uncertainty. Instead, uncertainty models have to be constructed from available data.

Barrier 1 **(Data scarcity)**. *Data is essential when setting up stochastic models, yet it is especially scarce in master production scheduling.*

Master production scheduling derives tactical decisions typically aggregated on a monthly granularity with planning horizons between 6 months and 2 years. Data sets cover only several years of historical data at most, hence only a limited number of observations are available. Further, the relevance of older data is limited by product life cycles and changes in market conditions (Chopra and Meindl, 2013). If demand is dynamic, for instance if there is a yearly seasonality, only few observations of the entire demand process are available in the data set. Yet, data is essential to measure the uncertainty of the planning environment and to evaluate model performance in simulations. Data scarcity is an ever-present problem for planners who require quantitative methods to support their decisions. This limits the application of sophisticated planning techniques. In particular, recent developments in data-driven operations research (Mišić and Perakis, 2020) may not be applicable.

Barrier 2 **(Uncertainty definition)**. *Identifying the nature, number and stationarity or lack thereof of uncertain processes influencing demand is critical.*

There are two main methods to characterise demand uncertainty. Most common is the assumption that demand itself can be modelled as an uncertain process. This method relies on past demand observations to predict future demand. It may be especially relevant when there is (a) a stationary demand process, (b) seasonality, or (c) if demand follows a type of auto-regressive process (Klabjan et al., 2013). For instance, Ban (2020) considers seasonal goods whose demand realises over a long season. Demand periods within the season are assumed correlated but observations of the full season are assumed independent identically distributed. Li and Disney (2017) model demand as a simple first-order auto-regressive process. In effect, the first approach assumes that demand is a seasonal but stationary process, whereas the second approach assumes that demand evolves over time but is locally stationary. The second method to characterise demand uncertainty focuses on the error caused by inaccurate forecasts. The key stochastic process is then the forecast error, which can be modelled as a probability distribution (Prak et al., 2017; Trapero et al., 2019). Both approaches ultimately provide uncertainty models for the demand over the planning horizon. However, the resulting uncertainty models vary drastically when measuring either demand or forecast uncertainty from data. Using wrong assumptions may lead to severely inaccurate models with long-lasting consequence. This first step is fundamental when applying stochastic models from data but is often overlooked in the literature. Practical guidelines describing methods to identify and measure uncertainty from limited data are still missing.

Barrier 3 **(Uncertainty model)**. *It is not clear when to use past data directly and when to estimate distributions, which is challenging with scarce data.*

Once the uncertain processes are defined and samples have been measured from historical data, the question arises of whether to use these samples directly in a data-driven fashion (Kleywegt et al., 2002) or to assume that they are observations of an underlying probability distribution. Creating scenario trees directly from data allows to capture correlation between products and period while avoiding distribution assumptions. However, it may fail to generalise from the data set and lead to overfitting.

In the literature, demand is commonly assumed to follow a known distribution, often normal, which is fitted to the data (Silver et al., 2016). These distributions can be sampled to create scenario trees over the horizon (Heitsch and Römisch, 2009; Homem-de-Mello and Bayraksan, 2014). Still, there is no guarantee that demand follows a probability distribution. Further, distribution parameters cannot be estimated precisely from scarce data and are subject to estimation error (Prak and Teunter, 2019). In a multi-dimensional setting, estimation error plays an even larger role since the number of observations may be much smaller than the number of parameters to estimate. An alternative to estimating probability distribution is to use distribution-free methods such as robust optimisation (Bertsimas et al., 2018a) or distributionally robust optimisation (Ben-Tal et al., 2013; Wiesemann et al., 2014). Yet, these methods have been applied to problems for which hundreds of observations are available and may not be suitable to problems with scarce data.

Barrier 2 and 3 are closely related but focus on different problems. Barrier 2 describes the challenges in identifying the source of uncertainty and obtaining relevant samples from past data whereas Barrier 3 discusses the different methods to process the samples.

Contributions. The strategies to overcome the barriers are typically problem specific. We propose two approaches to measure uncertainty from limited data based on seasonal demand uncertainty and forecast error respectively. For both uncertainty definitions, we compare the use of empirical and estimated probability distributions. The models are evaluated in simulations using real-world data. We show that accurately defining uncertainty is critical to ensure high demand satisfaction. In fact, deterministic models with simple rule-of-thumbs for safety stock but accurate uncertainty definition outperform stochastic models with wrong uncertainty definition. Thus, Barrier 2 is found more critical for performance than Barrier 3, which is remarkable considering that existing literature mostly focuses on the latter barrier at the expense of the former.

2.2. Reflecting the planning process

Stochastic models can reduce costs through recourse decisions that adapt to uncertainty as it unfolds. However, they also need to respect the constraints of the planning processes. The interaction of recourse models with planning flexibility, communicability and stability remains understudied.

Barrier 4 (Flexibility representation). *Stochastic programming models must be designed to properly represent the planning flexibility resulting from rolling-horizon planning processes.*

Since scenario-based stochastic programming can introduce recourse variables that adapt to the uncertain process as it unfolds (King and Wallace, 2012), it can capture the flexibility of rolling-horizon planning. Flexibility in production planning has been studied in early works by Escudero et al. (1993) and Brandimarte (2006) who compare different recourse structures in lot-sizing problems. Recently, Tavaghof-Gigloo and Minner (2020) propose a heuristic to integrate re-planning opportunities in a single-stage stochastic model by reducing safety stock levels when capacity is unlimited. Yet, recourse decisions should also improve costs when capacity is tight thanks lower safety stocks and better prioritisation of products over the horizon. Hence, how to best match the flexibility of stochastic programming (i.e. the definition of stages and recourse decisions) to the flexibility of the production environment is also an open question.
Quantifying the value of recourse in rolling-horizon planning has only be done partially. Existing works conduct static comparison of two-stage and multi-stage formulations in problems such as production planning with demand and yield uncertainty (Kazemi Zanjani et al., 2010), and lot-sizing and scheduling (Hu and Hu, 2018). Static evaluations ignore the rolling-horizon implementation of planning models, which provides flexible re-planning opportunities to stochastic models without recourse even if they are not explicitly modelled. A notable exception has been proposed by Stephan et al. (2010), who accurately measure the value of multi-stage models in capacity planning problems by using a rolling two-stage benchmark. Hence, practitioners cannot estimate the value of applying recourse models in rolling horizon.

Barrier 5 **(Communicability)**. *Scenario-independent reference plans need to be communicated to upstream and downstream members of the supply chain.*

Recourse models typically ignore the communicability requirement of rolling-horizon planning, which is essential throughout the supply chain. Contrary to deterministic or stochastic models without recourse, there is no unique plan obtained when solving a model with recourse. Instead, a tree of decisions is derived over the planning horizon that merely represents what-if statements. However, unconditional production plans need to be communicated to downstream parts of the supply chain to coordinate production schedules as well as the distribution and sale of finished goods.

In the same vein, raw-material orders are communicated to upstream parts of the supply chain to coordinate production and purchasing activities. Considerations of raw-material ordering and availability in production planning problems are rare and seem restricted to settings in which raw materials exhibit specific properties. For instance, Cunha et al. (2018) determine raw-material purchases with quantity-based discounts. Bollapragada et al. (2015) investigate the stochastic optimisation of procurement and production decisions in a make-to-order environment with supply uncertainty. More generally, Kanyalkar and Adil (2010) develop a two-stage stochastic model for the procurement, production and distribution including raw materials but consider a simple product structure with a single raw material. New formulations are thus needed to ensure communicability of a reference plan while allowing the flexibility of stochastic models with recourse. Further, scenario-based multi-stage solutions are typically not implementable in practice unless the true uncertainty distribution is discrete and completely captured in the scenario tree. Thevenin et al. (2021) investigate this issue by proposing several methods to determine a production policy from scenario-based multi-stage solutions. Yet, it is not discussed how to translate the obtained policy into a reference plan that provides long-term visibility.

Barrier 6 **(Plan stability)**. *Reference plans should be stable in rolling horizon with only limited changes between successive review periods, which may restrict the flexibility of recourse decisions.*

Significant plan changes create nervousness, which hinders supply chain performance, leads to loss of confidence, confusion through the supply chain and ultimately higher costs (Atadeniz and Sridharan, 2020). Seminal works analyse the nervousness resulting from lot-sizing heuristics in single-level (Carlson et al., 1979; Sridharan et al., 1988) or multi-level environments (Blackburn et al., 1986; Ho, 1989; Zhao et al., 2001). They develop strategies to mitigate nervousness such as freezing periods or penalising plan changes. Recent research studies the nervousness resulting from optimal planning models. Lin and Uzsoy (2016)

compare chance-constraint formulations to capture demand uncertainty and their impact on planning stability. Herrera et al. (2016) integrate different nervousness penalty costs in the objective function to identify a balance between stability and operational costs. Meistering and Stadtler (2017) propose a stabilised-cycle strategy that allows changes in production decisions only when necessary to reach the target service level
Existing nervousness mitigation strategies are based on restricting planning flexibility, which may reduce planning performance when short-term uncertainty is high. While stochastic models should derive optimal production volumes despite the limited flexibility, it is not clear how they would perform when distributions are not known but modelled from data. Further, since freezing periods inherently prohibit recourse opportunities, the trade-off between traditional nervousness reduction methods and stochastic programming with recourse remains open.

Contributions. We note that existing stochastic models with recourse do not evaluate the resulting nervousness, since reference plans are not determined in existing stochastic programming models. By providing reference plans when solving stochastic models with recourse, we can bridge the gap between research on planning stability and stochastic programming.
We contribute to existing literature in several ways. First, we develop a two-stage model that provides recourse and reference plans based on aggregating products into optimal families. Second, we measure the value of recourse in rolling-horizon planning with real-world data. In particular, we show that recourse is especially beneficial when capacity is limited. Finally, we compare the use of traditional nervousness mitigation strategies based on frozen decisions and our novel approach based on product aggregation. We show that freezing decisions on the raw-material level does not limit planning flexibility while providing significant stability improvements. On the other hand, the aggregation-based strategy can improve planning flexibility, communicability and stability, thus outperforming the traditional strategy of freezing production decisions.

2.3. Computational challenges

The evaluation of stochastic models is challenging due to several factors including long computation times, the need for complex simulation settings, and the strong dependence of results on the assumptions used in simulations.

Barrier 7 (Tractability). *Stochastic models often exhibit a trade-off between accuracy and long computation times.*

Stochastic programming approaches, and especially multi-stage formulations, lead to notoriously long computation times. Significant attention has been given to designing scenario trees with optimal size. In particular several methods have been developed to reduce the size of scenario trees while retaining their accuracy (Dupačová et al., 2003; Heitsch and Römisch, 2003). Other approaches to improve computation times include decomposition techniques such as progressive hedging (Watson and Woodruff, 2011) and stochastic dual dynamic programming (Shapiro, 2011). Yet, solving times depend not only on the scenario tree but also on the recourse structure. The trade-off between computation times and flexibility offered by recourse also needs to be analysed.

Barrier 8 (Evaluation). *The performance of stochastic models should be evaluated accurately despite limited available data.*

A reliable assessment of expected performance is essential to foster the adoption of new models. This reliability can be achieved by simulating the model in a setting close to its practical use. Simulations can be implemented in a rolling-horizon fashion to respect the planning structure and performed in an out-of-sample fashion to accurately evaluate the uncertainty model. To the best of the authors' knowledge, out-of-sample evaluations have not been applied in production planning to evaluate stochastic models based on real-world data. Out-of-sample evaluations have been more commonly applied to inventory management and in particular to newsvendor problems (Beutel and Minner, 2012; Bertsimas et al., 2018b; Huber et al., 2019; Oroojlooyjadid et al., 2020). When data is scarce and is used for both model calibration and evaluation, carefully designing the simulation experiments is crucial.

Contributions. We study the trade-off between model accuracy and tractability by varying the scenario size as well as the recourse structure. In both cases, we show that efficient trade-offs can be found. To tune and evaluate the models, we propose the first out-of-sample rolling-horizon evaluation of stochastic production planning models with real-world data. We highlight the importance of out-of-sample evaluation by measuring the bias of in-sample evaluations.

3. Real-world case study

In this section, we introduce the industry problem and show the relevance of the barriers identified above. While barriers may be common to many production planning problems, we believe that solution approaches are inherently problem specific. We discuss the form of the barriers in the case study and provide an overview of the strategies to overcome them.

3.1. Problem setting

Our industry partner is a world-leading agrochemical company managing a global supply chain with a large product portfolio, long production lead times and complex planning problems. We focus on the production of a restricted product portfolio of pesticides that embodies the planning challenges of the firm. Since the use of pesticides follows the crops' growth cycle, demand patterns exhibit strong seasonality, and accurately forecasting demand is limited by unpredictable parameters such as weather conditions.

The production of synthetic pesticides contains two main steps: the active ingredient synthesis, in which the molecules forming the base of the finished products are synthesised, and the formulation step, in which one or several active ingredients are combined and diluted. The active ingredient synthesis is the most complex process with important capital investment, long lead times and low flexibility. At this level, production is conducted in long campaigns that realise over several months to a year. Short-term changes to campaigns are limited by cleaning operations that can last up to several weeks. As the most value-adding process, the active ingredient synthesis highlights the inherent challenge of agrochemical supply-chain management: production has low flexibility and long lead times whereas demand is dynamic and hard to predict even in the short future. Production and supply planners are thus looking for advanced strategies to manage demand uncertainty and to ensure efficient operations throughout the supply chain.

Because of the complexity of the global network, the active ingredient synthesis and formulation are planned sequentially. Formulation planners derive the intended production over the planning horizon and deduce the active ingredient requirements that are communicated to upstream planners. The aim of our industry collaboration is to improve the formulation planning step to derive plans that satisfy the uncertain demand while ensuring that stable raw-material orders are provided to upstream planners. In effect, this improved formulation planning would act as a dampening step, reducing the uncertainty of the demand forecast as it propagates through the supply chain.

3.2. Overcoming the barriers

From the identification of the uncertain processes to the model development and evaluation, this industry problem encompasses the barriers of stochastic programming described in Section 2. We discuss the specific forms taken by the barriers in this industry case and present an overview of our strategies to overcome them.

Uncertainty. Historical forecasts and past demands are available for the last four years. Because of the seasonality of demand, this data set corresponds to only few observation of the entire demand process. Defining the uncertain processes from this limited data set is challenging since demand is dynamic and forecasts are inaccurate. To overcome Barrier 2 (Uncertainty definition), we derive seasonal models of uncertainty. Both demand-driven and forecast-driven are analysed based on the uncertainty of demand and forecasts respectively. The two approaches provide different samples for the empirical demand distributions, which can be either used directly or to estimate probability distributions. To overcome Barrier 3 (Uncertainty models), we implement both approaches, estimating normal and uniform distributions from the empirical samples. Scenarios trees are created and integrated in two-stage stochastic models.

Planning processes. The supply chain and production processes of the industry case are complex. In particular, the active ingredients have long production lead times and are especially sensitive to planning nervousness. Yet, flexibility is essential to ensure that demand can be met despite poor forecasts accuracy. We overcome the barriers linked to planning flexibility, stability and communicability in several way. To overcome Barrier 5 (Communicability), we ensure that a reference plan is always available on both the production plan and raw-material levels. Raw-material orders and inventory are explicitly modelled. Long-term visibility is essential for raw-material planning. However, a detailed production plan is only required by downstream planners to determine the schedule of formulation campaigns. Hence, we can aggregate communications on the production plan level by defining product families. The definition of the families is a key part of our approach. To ensure that aggregated plans provide the information necessary to derive production schedules, families are defined through a multi-objective optimisation models with custom rewards and constraints that reflect production processes. Product families allow to overcome Barrier 4 (Flexibility representation) by introducing production recourse. First-stage capacity reserves are placed on the family level, which can be used flexibly by products within in the family through recourse decisions. We observe that plan changes within product families tend to compensate in rolling horizon so that aggregating decisions on the family level also improves planning stability. Thus, we compare the nervousness mitigation techniques of freezing and aggregating decisions to overcome Barrier 6 (Plan stability). The different planning strategies are integrated in a mixed-integer linear problems that optimally determines the share of first-stage and recourse production decisions.

Numerical study. The models are evaluated through rolling-horizon simulations and extensive sensitivity analyses are performed. We overcome Barrier 7 (Tractability) by studying the effect of the size of the scenario trees and the number of product families that both increase the number of recourse variables. In both cases, efficient trade-offs can be found between solution quality and model complexity. The simulation setting is crucial to overcome Barrier 8 (Evaluation). The out-of-samples rolling-horizon simulation framework proposed is especially powerful to make generalisable conclusions from the limited amount of data and avoid in-sample bias. To finalise the model evaluation, meaningful benchmarks are defined from historical company data, providing an accurate assessment of expected improvements compared to current practice. Barrier 1 (Data scarcity) is the most fundamental and challenging barrier to overcome. It underlies all strategies applied in this paper, and is only overcome at the end of the numerical study, once we finally identify the best model configuration and prove its benefits experimentally.

4. Modelling the uncertain process

In this section, we discuss the definition and representation of the uncertain process. We present two uncertainty models based on demand uncertainty and forecast error respectively. Scenarios are obtained for both models and used to construct two-stage scenario trees.

4.1. Seasonal demand uncertainty

A data set is available covering Y seasons of S periods each. Historical demand of the portfolio of K products has been observed where $d_k^{s,y}$ is the demand for product k observed in period s of season y. To reflect seasonality, the first uncertainty model assumes that demand follows a stationary distribution in each period of the season and that demand periods within the season may be correlated. The planner can either use the empirical distribution $\mathcal{D}_s = \left\{d_k^{s,y}, y \in \{1, \ldots, Y\}\right\}$ derived from past observations of demand in period s of the season, or estimate a probability distribution to derive additional scenarios. This uncertainty model is based solely on past demand data and ignores forecasts available in each review period. It is a static approach that does not benefit from forecast updates obtained in rolling horizon.

4.2. Seasonal forecast error

In rolling horizon, an updated forecast is obtained in each review cycle covering a planning horizon of T periods. Let $f_{k,t}^{s,y}$ be the forecast for product k in period t of the planning horizon as seen in review period s of season y. We introduce an additional time index to distinguish the different versions of forecast relating to the same demand period.

To model uncertainty in a forecast-driven fashion and reflect the seasonality of both the demand and forecast processes, we introduce the concept of seasonal forecast error. The forecast error associated to planning period s of season y is defined by by $\mathbf{e^{s,y}} = (e^{s,y}_{k,t}) \in \mathbb{R}^{K\times T}$ where

$$e^{s,y}_{k,t} = d^{s+t-1,y}_{k} - f^{s,y}_{k,t}, \quad \forall k \in \mathcal{K}, t \in \mathcal{T}.$$

The novelty of this model is to assume that each planning period s has its own forecast error distribution, which is stationary across seasons. For each review period s in the season, the set of forecast error $\mathcal{E}_s = \{\mathbf{e^{s,y}}, y \in \{1, \dots, Y\}\}$ is the empirical distribution of the (unknown) multivariate random forecast error distribution. The forecast error can be measured a posteriori for all periods for which the actual demand has been observed. The empirical distribution can be used to estimate the parameters of an assumed distribution and sampled to create additional forecast error scenarios. Since it is not straightforward to decide a priori which uncertainty model provides the best results, we compare their performance numerically through out-of-sample simulations in Section 6.2.

4.3. Two-stage scenario tree

Let y_o be the current season for which we want to derive a production plan using Y past seasons. Demand-driven samples obtained in Section 4.1 can be used directly to form a scenario tree.

Forecast error samples can also be used to generate a scenario tree by correcting the currently available forecast with forecast error samples. Let $N - 1$ be the number of forecast error samples equal to Y if one uses the empirical distribution. A two-stage scenario tree can be constructed as a fan containing N equiprobable sample paths. The first path is set to the deterministic demand forecast $f^{1,s,y_o}_{k,t} = f^{s,y_o}_{k,t}$ for all products over the planning horizon. The remaining scenarios can be determined as

$$f^{,s,y_o}_{k,t,n+1} = f^{s,y_o}_{k,t} + e^{s}_{k,t,n}, \quad \forall k \in \mathcal{K}, t \in \mathcal{T}, n \in \{1, \dots, N-1\}.$$

to correct the deterministic forecast with the seasonal forecast error samples of the same review period. Scenarios with negative demand are corrected to take the value zero.

4.4. Summary

We have shown strategies to overcome Barrier 2 (Uncertainty definition) and Barrier 3 (Uncertainty model) with limited available data. Seasonal uncertainty models based on forecast and demand data have been presented and integrated in two-stage scenarios trees. The use of scenarios based on the empirical distribution and estimated distribution have been considered.

5. Stochastic planning model with flexibility, stability and communicability

The uncertainty model and scenario tree presented in the previous section are now integrated in a stochastic planning model. We present different recourse structures and strategies to ensure communicability and flexibility of reference plans for both production and raw-material decisions. The notation for this section is summarised in Table A.1 in the appendix.

5.1. Stochastic model without recourse

The planner manages a portfolio of K products made from A raw materials and needs to determine a production plan over a horizon of T periods. Consider a general product structure in which raw material can be used for several products and each product can require multiple raw materials. The bill of material is given by $\mathbf{U} = (u_{k,a}) \in \mathbb{R}^{K \times A}$ where $u_{k,a}$ is the amount of raw material a required to produce one unit of product k. The planner is responsible for several production sites that serve a regional market. Each site contains parallel lines with different capacity κ_l and product portfolio. The set of production lines at site w is denoted by $\mathcal{L}_w$. The set of products that can be formulated on line l is given by $\mathcal{K}_l = \{k \in \mathcal{K} \mid \rho_{k,l} = 1\}$. Raw-material inventory is kept in a single warehouse and shared over the production sites whereas finished goods are held at the production sites. At the end of each review period, the company incurs a per-unit holding costs ν_a for raw-material a and $\mu_{k,w}$ for product k in site w. The problem setting is illustrated in Figure 1 for two production sites and five lines.

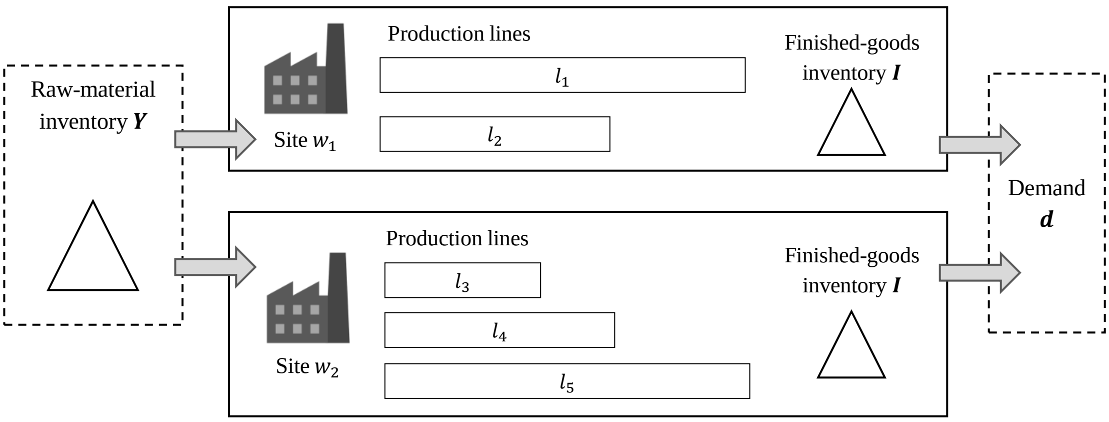

Figure 1: Supply, production and inventory system for $W = 2$ sites and $L = 5$ lines.

In each review period, the planner uses a scenario tree $f \in \mathbb{R}^{(K \times T \times N)}$ to determine a production plan over the horizon and to communicate raw-material orders to the upstream level. The planner's goal is to satisfy the uncertain demand while minimising the inventory costs of raw materials and finished goods. Unmet demand is considered a lost sale and penalised with per-unit cost γ_k. The planning model is formulated as a two-stage stochastic model. Production decisions and raw-material orders are set as first-stage variable in order to provide a reference plan over the horizon. The inventory, sales and lost-sales decisions are set as recourse variables. The model is presented in Problem (1) where the season and review period indices are dropped for clarity.

$$\min \quad \sum_{t=1}^{T}\left(\frac{1}{N}\sum_{k=1}^{K}\sum_{w=1}^{W}\mu_{k,w}\sum_{n=1}^{N} i_{k,w,t,n} + \sum_{a=1}^{A}\nu_a \cdot y_{a,t} + \frac{1}{N}\sum_{k=1}^{K}\gamma_k\sum_{n=1}^{N}\cdot b_{k,t,n}\right) \tag{1a}$$

$$\text{s.t.} \quad i_{k,w,t,n} = i_{k,w,t-1,n} + \sum_{l\in\mathcal{L}_w} q_{k,l,t} - s_{k,w,t,n}, \qquad \forall k,w,t,n \tag{1b}$$

$$f_{k,t,n} = b_{k,t,n} + \sum_{w=1}^{W} s_{k,w,t,n}, \qquad \forall k,t,n \tag{1c}$$

$$\sum_{k=1}^{K} q_{k,l,t} \le \kappa_l, \qquad \forall l,t \tag{1d}$$

$$y_{a,t} = y_{a,t-1} + z_{a,t} - \sum_{k=1}^{K}\sum_{l=1}^{L}\beta_{k,a}\cdot q_{k,l,t}, \qquad \forall a,t \tag{1e}$$

$$q_{k,l,t}, y_{a,t}, z_{a,t} \ge 0, \qquad \forall k,w,l,a,t \tag{1f}$$

$$i_{k,w,t,n}, b_{k,t,n}, s_{k,t,n} \ge 0, \qquad \forall k,l,w,t,n \tag{1g}$$

The objective function in (1a) minimises the expected costs of inventory and lost sales over the different scenarios where the lost-sales penalty cost γ_k adjusts the conservativeness of the solution. Constraint (1b) describes the inventory balance at the production sites. Constraint (1c) ensures that demand is satisfied from sales or accounted as a lost sale in each scenario path. Constraint (1d) limits the production on each line to its capacity in each period. Constraint (1e) describes the raw-material inventory balance. Constraints (1f) and (1g) specify the domain of the first-stage and recourse decisions variables respectively.

To improve planning stability, production and raw-material decisions can be frozen over the short-term horizon, prohibiting changes from decisions made in the previous review period. The frozen horizon can be implemented through the additional constraints

$$z_{a,t} = z^0_{a,t}, \qquad \forall a, t \le \tau_a \tag{2a}$$

$$q_{k,l,t} = q^0_{k,l,t}, \quad \forall k,l,t \le \tau_k \tag{2b}$$

where $z^0_{a,t}$ and $q^0_{k,l,t}$ are raw-material orders and production values determined in the previous review period. The length of the frozen horizon for production decisions τ_k and raw-material orders τ_a is chosen by the planner.

The stochastic model presented in (1) overcomes Barrier 5 (Communicability) by providing a reference plan on both the production and raw-material levels. Barrier 6 (Plan stability) can be overcome by freezing decisions on either the raw-material, the production levels, or both. However, the resulting model provides low flexibility since there is no recourse production and short-term decisions are frozen.

5.2. *Improving flexibility through production recourse*

By allowing recourse, decisions can be adapted to each scenario leading to less conservative here-and-now decisions. However, recourse variables limit planning communicability since the planner does not determine a unique reference plan but a tree of decisions. We introduce a stochastic model with recourse that provides high flexibility and communicability. The model is based on product families built through a data-driven optimisation model with custom rewards and constraints that reflect the product structure. The families are integrated in the planning model that reserves capacity on the family level through first-stage decisions. Thus, a reference plan is obtained on the family level. Recourse production decisions that consume the reserved capacity are implemented for products within families. In the numerical study, we show that aggregating production decisions over products improves planning stability since production changes tend to compensate within product families.

5.2.1. *Product families: a multi-objective problem*

The product-to-family assignment problem is a multi-criteria decision problem. Together with our industrial partner, we identify properties that the final assignment should exhibit: (1) the family assignment should cover many products, (2) a large share of the demand should be covered in product families, and (3) products with high uncertainty should be prioritised in the allocation. Each property is formulated as a normalised reward function, so that the rewards can be weighted easily to reflect planners' preferences. The product families should also respect the operational constraints and provide high visibility to site planners and schedulers. The product families are built in a data-driven fashion by using the historical data set of Y seasons.

Custom reward functions. Let $x_{k,f}$ be the binary variable equal to 1 if product k is assigned to family f. The first reward function is given by

$$\psi_1(X) = \frac{1}{K}\sum_{f=1}^{F}\sum_{k=1}^{K} x_{k,f}$$

and simply counts the number of products assigned to families. The second reward quantifies the share of demand covered by assigned products. It is given by

$$\psi_2(X) = \frac{1}{\sum_{k=1}^{K} td_k}\sum_{f=1}^{F}\sum_{k=1}^{K} td_k \cdot x_{k,f}$$

where $td_k = \sum_{y=1}^{Y}\sum_{s=1}^{S} d_k^{s,y}$ is the total demand of product k over the data set. The third reward prioritises products with high uncertainty. It is expressed by

$$\psi_3(X) = \frac{1}{\sum_{i=1}^{K} fe_i}\sum_{f=1}^{F}\sum_{k=1}^{K} fe_k \cdot x_{k,f}$$

where fe_k represents the difficulty to forecast product k. In this paper, we measure the uncertainty of a product using the weighted mean absolute percent error (wMAPE). This measure is normalised and allows to compare the forecast error of products with different demand share. The wMAPE forecast error in review period s of season y is given by

$$fe_{k,t}^{s,y} = \frac{\sum_{t=1}^{T} \omega_t \mid d_k^{s+t-1,y} - f_{k,t}^{s,y} \mid}{\sum_{t=1}^{T} \omega_t \cdot d_k^{s+t-1,y}} \quad \forall k, t.$$

where the weighting factor ω_t emphasises forecast error over the short-term horizon. The average product forecast error is then calculated as $fe_k = \frac{1}{T}\sum_{t=1}^{T} fe_{k,t}$.

Model formulation. The optimisation model is formulated in Problem (3).

$$\max_{x} \quad \sum_{i=1}^{3} w_i \cdot \psi_i(x) \tag{3a}$$

$$\text{s.t.} \quad \sum_{f=1}^{F} x_{k,f} \leq 1, \quad \forall k \tag{3b}$$

$$x_{k_1,f} \cdot x_{k_2,f} \leq \rho_{k_1,l} \cdot \rho_{k_2,l}, \quad \forall k_1, k_2, l, f \tag{3c}$$

$$x_{k_1,f} \cdot x_{k_2,f} \leq 1 - m_{k_1,a_1} \cdot m_{k_2,a_2} \cdot (1 - m_{k_1,a_2}) \cdot (1 - m_{k_2,a_1}) \tag{3d}$$

$$- m_{k_1,a_2} \cdot m_{k_2,a_1} \cdot (1 - m_{k_1,a_1}) \cdot (1 - m_{k_2,a_2}), \quad \forall k_1, k_2, a_1, a_2, l, f,$$

$$x_{k,f} \in \{0; 1\}, \quad \forall k, f. \tag{3e}$$

The objective function in (3a) maximises the weighted sum of rewards. Constraint (3b) ensures that a product is assigned to at most one family. Constraint (3c) specifies that all products within a family must be produced on the same set of production lines. Constraint (3d) states that there is always a sequence of products feasible without cleaning operation within a product family. Although cleaning operations are outside the scope of tactical planning, we ensure that the reference plan on the family level provides high visibility for the site schedulers. In the agrochemical industry, cleaning operations are conducted each time a raw material is removed when switching equipment from one product to the next. Let $\mathbf{m} = (m_{k,a})$ be the raw-material usage matrix where $m_{k,a}$ is equal to 1 if product k requires raw material a and 0 otherwise. Constraint (3d) holds for any number of products and raw materials. Although the above formulation is non-linear, the product of binary variables in Constraints (3c) and (3d) can be linearised by adding auxiliary variables $z_{k1,k2,f}$ and the following constraints:

$$z_{k_1,k_2,f} \leq x_{k_1,f}, \quad \forall k_1, k_2, f \tag{4}$$

$$z_{k_1,k_2,f} \leq x_{k_2,f}, \quad \forall k_1, k_2, f \tag{5}$$

$$z_{k_1,k_2,f} \geq x_{k_1,f} + x_{k_2,f} - 1, \quad \forall k_1, k_2, f. \tag{6}$$

5.2.2. Two-stage stochastic model with production recourse

The product families are integrated in a stochastic model that allows recourse production decisions. The extended stochastic model with family reserves and production recourse is formulated in Problem (7).

$$\min \quad \sum_{t=1}^{T}\sum_{k=1}^{K}\sum_{w=1}^{W}\mu_{k,w}\sum_{n=1}^{N}\frac{1}{N}i_{k,w,t,n} + \sum_{a=1}^{A}\nu_a \cdot \sum_{n=1}^{N}\frac{1}{N}y_{a,t,n} + \sum_{k=1}^{K}\gamma_k\sum_{n=1}^{N}\frac{1}{N}\cdot b_{k,t,n} \tag{7a}$$

$$\text{s.t.} \quad i_{k,w,t,n} = i_{k,w,t-1,n} + \sum_{l\in\mathcal{L}_w}\left(q_{k,l,t} + r_{k,l,t,n}\right) - s_{k,w,t,n}, \qquad \forall k, w, t, n \tag{7b}$$

$$f_{k,t,n} = b_{k,t,n} + \sum_{w=1}^{W} s_{k,w,t,n}, \qquad \forall k, t, n \tag{7c}$$

$$\sum_{k=1}^{K} q_{k,l,t} + \sum_{f=1}^{F} h_{f,l,t} \leq \kappa_l, \qquad \forall l, t \tag{7d}$$

$$\sum_{k\in\mathcal{K}_f} r_{k,l,t,n} \leq h_{f,l,t}, \qquad \forall f, l, t, n \tag{7e}$$

$$r_{k,l,t,n} \leq \sum_{f=1}^{F} x_{k,f}\cdot h_{f,l,t}, \qquad \forall k, l, t, n \tag{7f}$$

$$q_{k,l,t} + r_{k,l,t,n} \leq \kappa_l \cdot \rho_{k,l}, \qquad \forall k, l, t, n \tag{7g}$$

$$y_{a,t,n} = y_{a,t-1,n} + z_{a,t} - \sum_{k=1}^{K}\sum_{l=1}^{L}\beta_{k,a}\cdot(q_{k,l,t} + r_{k,l,t,n}), \qquad \forall a,n,t \qquad (7h)$$

$$u_{k,t,n}\cdot\sum_{l=1}^{L}\kappa_l\cdot\rho_{k,l} \geq \sum_{l=1}^{L} r_{k,l,t,n}, \qquad \forall k,t,n \qquad (7i)$$

$$\sum_{n=1}^{N} u_{k,t,n} = N-1, \qquad \forall k,t \qquad (7j)$$

$$\sum_{k\in\mathcal{K}_f}\sum_{l=1}^{L} r_{k,l,t,n} \geq \sum_{l=1}^{L} h_{f,l,t} - v_{f,t,n}\cdot\sum_{l=1}^{L}\kappa_l, \qquad \forall f,t,n \qquad (7k)$$

$$\sum_{n=1}^{N} v_{f,t,n} = N-1, \qquad \forall f,t \qquad (7l)$$

$$r_{k,l,t,n}, h_{f,l,1} = 0, \qquad \forall k,l,f,n \qquad (7m)$$

$$z_{a,t} = z^0_{a,t}, \qquad \forall a, t\leq\tau_a \qquad (7n)$$

$$q_{k,l,t} = q^0_{k,l,t}, \qquad \forall k,l,t\leq\tau_k \qquad (7o)$$

$$h_{f,l,t} = h^0_{f,l,t}, \qquad \forall f,l,t\leq\tau_k \qquad (7p)$$

$$q_{k,l,t}, h_{f,l,t}, z_{a,t}, \geq 0, \qquad \forall k,l,a,f,t \qquad (7q)$$

$$i_{k,w,t,n}, b_{k,t,n}, s_{k,t,n}, y_{a,t,n}, r_{k,l,t,n} \geq 0, \qquad \forall k,l,w,a,t,n \qquad (7r)$$

$$u_{k,t,n}, v_{f,t,n} \in \{0;1\}, \qquad \forall f,k,t,n. \qquad (7s)$$

The objective function in (7a) minimises the expected costs of finished-goods inventory, raw-material inventory and lost sales. Constraint (7b) describes the inventory balance of finished goods at each production site. Constraint (7c) tracks the demand satisfaction from the sites. Constraint (7d) ensures that production and capacity reserves on each line do not exceed available capacity. Constraint (7e) states that recourse production within a family is restricted by its capacity reserve in each scenario. Constraint (7f) ensures that there is no recourse production for unassigned products. Constraint (7g) specifies that production on a line is restricted to its feasible portfolio. Constraint (7h) describes the raw material balance. Constraints (7i) and (7j) ensure that the minimum recourse production over all scenarios is zero for each product and time period. Constraints (7k) and (7l) force the maximum recourse production over all scenario to be equal to the capacity buffer reserved for each product in each period. These two sets of constraints ensure that the capacity buffer reserved for each family accounts exactly for the volatile part of demand. Constraint (7m) states that there is no recourse variable in the first period. Constraints(7n) implements a frozen horizon on the raw-material orders. Constraints (7o) and (7p) implement a frozen horizon on first-stage production decisions and capacity reserves respectively. Constraints (7q) and (7r) express the domain of the continuous first-stage and recourse variables respectively. Constraint (7s) defines the auxiliary binary variables to identify the minimum and maximum production recourse over the scenarios.

5.3. Summary

The stochastic model with recourse determines the optimal first-stage raw-material orders, production and capacity reserves that allows flexible second-stage production decisions, overcoming Barrier 4 (Flexibility representation). Recourse production can only be used for products within families if enough resources have been reserved. Although it uses scenarios, the model overcomes Barrier 5 (Communicability) by providing an aggregated reference plan defined so that a detailed production schedule can still be derived by downstream planners. Barrier 6 (Stability) is overcome by aggregating first-stage decisions on the family level. The model also explicitly distinguishes between parts of the plan likely to be conducted (first-stage production) and parts of the plan potentially subject to changes (recourse production).

6. Numerical study

The numerical study applies the strategies developed in the previous sections to the real-world case study and details the final steps to overcome barriers of stochastic programming in practice. Due to the large number of parameters and performance indicators, it is difficult to investigate their interactions in a full factorial experiment. Instead, we analyse sequentially the strategies related to the uncertainty process from Section 4 and the planning structure from Section 5. First, we present the simulation setting, the performance metrics and the problem parameters. Second, we compare the performance of demand-driven and forecast-driven uncertainty models as well as the use of empirical or estimated distributions. Third, we evaluate the stochastic models with varying nervousness mitigation strategies. Finally, we compare our model to the current practice of our industrial partner.

Simulations are implemented in the Julia programming language (Bezanson et al., 2017) and are run on an Intel(R) Core(TM) i7-4810MQ processor at 2.80Ghz using 16GB of RAM. The optimisation problems are formulated using JuMP (Dunning et al., 2017) and solved with Gurobi 9.0. The relative MIP gap is set to 0.1% for all instances of the stochastic model with production recourse.

6.1. Simulation setting

All simulations from model parameterisation to final evaluation are conducted in an out-of-sample rolling-horizon fashion. In each review period, the following steps are taken: (1) a production plan is calculated over the planning horizon using the available forecast or scenario tree, (2) the production quantity of the first period is added to the on-hand inventory, (3) the actual demand is observed, (4) sales are subtracted from the inventory and lost sales are observed if demand is higher than the on-hand inventory, and (5) the new inventory position is determined.

We gather a data set containing the planning history of our industrial partner over $Y = 4$ seasons of $S = 12$ months each. Rolling-horizon simulations are run for each season independently using the other $(Y - 1)$ seasons to construct the uncertainty model. In each review period, the forecast and demand are taken from the historical data set of our industrial partner. This simulation setting allows to carry Y independent out-of-sample simulations.

The initial inventory in the first period of the season is set to the historical inventory of the company. Each simulation is started $\tau = \max(\tau_a, \tau_k)$ periods earlier than the first period of the season and the demand and forecast are set to zero during this warm-up phase. The corresponding review periods are ignored for the model evaluation. To neglect the interactions between consecutive seasons, we replace demand and forecast values by zero for all periods later than the last period of the current simulation season.

6.1.1. Key performance indicators

The models are evaluated using four key performance indicators: the service level, the inventory costs, the planning nervousness and the nervousness of the raw-material orders.

Key trade-off: service level and inventory. The service level is measured as the proportion of satisfied demand over the season given by

$$sl = \frac{\sum_{s=1}^{S} \sum_{k=1}^{K} (d_{k,s} - b_{k,s}^{(r)})}{\sum_{s=1}^{S} \sum_{k=1}^{K} d_{k,s}}$$

where $b^{(r)}$ are the realised lost sales of product k in simulation period s respectively. The inventory costs are measured as the sum of finished-goods and raw-material inventory costs over the season as

$$ic = \sum_{s=1}^{S} \sum_{k=1}^{K} \sum_{w=1}^{W} \mu_{k,w} \cdot i_{k,w,s}^{(r)} + \sum_{s=1}^{S} \sum_{a=1}^{A} \nu_a \cdot y_{a,s}^{(r)}$$

where $i_{k,w,s}^{(r)}$ and $y_{a,s}^{(r)}$ are the realised finished-goods inventory and raw-material inventory observed at the end of period s. All inventory costs reported are normalised by dividing them by the company's average historical inventory costs.

Planning stability. There is a large body of literature discussing how to measure nervousness. Quantity-oriented and setup-oriented nervousness measures have been distinguished, which are particularly relevant in lot-sizing contexts (Tunc et al., 2013). Early measures focus on setup-oriented nervousness such as Carlson et al. (1979) who account only for the nervousness induced by adding a new setup in the plan. Sridharan et al. (1988) have proposed a quantity-oriented measure that assigns weights to periods in the horizon in order to emphasise short-term stability. While the majority of existing measures are absolute, relative nervousness measures are more interpretable. Jensen (1993) proposes a normalised nervousness measures that relate nervousness to the maximum nervousness possible, which can be determined from the available capacity. However, this measure has several practical shortcomings: it cannot be applied if the maximum nervousness is unbounded, and it might give a false sense of stability if capacity is large. We propose a novel quantity-oriented nervousness measure that is relative to the plan itself. This measure provides high interpretability and allows to compare several planning steps. For instance, we compare production planning nervousness, raw-material nervousness and forecast nervousness in Section 6.4.

Planning stability is measured independently on both the production and raw material levels. Planning nervousness is measured as the average sum of absolute changes between production volumes aggregated on the product family level. It is based on the observation that nervousness within a family is negligible compared to nervousness between families. Planning nervousness is measured as

$$nsf = \frac{1}{S-1}\sum_{s=2}^{S}\frac{\sum_{t=1}^{T-1}\sum_{f=1}^{F} \mid \sum_{l=1}^{L}(HK^{(s)}_{f,l,t} + \sum_{k\in\mathcal{K}_f} Q^{(s)}_{k,l,t} - HK^{(s-1)}_{f,l,t+1} - \sum_{k\in\mathcal{K}_f} Q^{(s-1)}_{k,l,t+1}) \mid + \sum_{k\in\mathcal{K}\setminus\mathcal{KF}} \mid \sum_{l=1}^{L} Q^{(s)}_{k,l,t} - Q^{(s-1)}_{k,l,t+1} \mid}{\max\left(\sum_{t=1}^{T-1}\sum_{l=1}^{L}\sum_{k=1}^{K} Q^{(s)}_{k,l,t} + \sum_{f=1}^{F} HK^{(s)}_{f,l,t}, \sum_{t=1}^{T-1}\sum_{l=1}^{L}\sum_{k=1}^{K} Q^{(s-1)}_{k,l,t+1} + \sum_{f=1}^{F} HK^{(s-1)}_{f,l,t+1}\right)} \tag{8}$$

where $\mathcal{K} \setminus \mathcal{KF}$ is the set of products not assigned to any family. Raw-material orders nervousness is given by

$$nsa = \frac{1}{S-1}\sum_{s=2}^{S}\frac{\sum_{t=1}^{T-1}\sum_{a=1}^{A} \mid Z^{(s)}_{a,t} - Z^{(s-1)}_{a,t+1} \mid}{\max\left(\sum_{t=1}^{T-1}\sum_{a=1}^{A} Z^{(s)}_{a,t}, \sum_{t=1}^{T-1}\sum_{a=1}^{A} Z^{(s-1)}_{a,t+1}\right)}. \tag{9}$$

The nervousness measures are quantity oriented. They account for plan changes due to both volume and timing. Nervousness is calculated relative to the reference plan. This normalisation does not guarantee that nervousness is always between 0 and 1 but increases interpretability and allows the comparison of different models.

6.1.2. Problem parameters

The product portfolio contains $K = 55$ products formulated from $A = 13$ raw materials. There are $W = 2$ production sites with 2 production lines at the first site and 3 lines at the second site. The demand is seasonal with periodicity $S = 12$ periods and the planning horizon is set to $T = 12$ periods. The raw-material frozen horizon is set to $\tau_a = 2$ periods. There is no frozen horizon for production decisions. The line capacities, bill of materials, each line's product portfolio and the inventory costs have been collected together with our industrial partner. The lost-sales penalty cost is set proportional to the product inventory cost as $\gamma_k = \lambda \cdot \max_{w \in \mathcal{W}}(\mu_{k,w})$.

6.2. Evaluation of uncertainty models

6.2.1. Pareto fronts

The stochastic models are implemented in out-of-sample rolling-horizon simulations to measure the value of different uncertainty models and decide on the optimal configuration. We compare the performance of the empirical distribution and estimated distributions as well as the use of forecast-driven and demand-driven models. The value of augmenting the scenario tree with scenarios sampled from assumed distribution is also measured.

Normal and uniform distributions are fitted to the empirical samples. A normal distribution is estimated from the empirical mean and variance of the forecast error independently for each product and time period. The bounds of the uniform distribution are taken as 80% and 120% of the minimum and maximum empirical forecast error for each product and time period. We refrain from estimating covariance parameters since the number of samples $(Y - 1)$ is significantly smaller than the number of parameters to estimate $(K^2 \times T^2)$. Demand scenarios are then sampled using Descriptive Sampling (Saliby, 1990). To identify the value of stochastic programming, we also show the Pareto front of deterministic models with exogenous safety stock calculations. The deterministic optimisation model presented in Appendix B is implemented with additional exogenous safety stocks determined by $ss_{k,t} = z \cdot \sigma_{k,t}$ where σ is the standard deviation of demand (DD) or forecast error (DF) and z is a conservativeness parameter set by the planner.

Since the planning problem has four objectives, several trade-offs exist between the performance indicators presented in Section 6.1.1. We focus on the most important trade-off between realised service level and inventory costs. While the planner is interested in achieving high demand satisfaction, it comes at the price of more conservative decisions yielding higher inventory costs. We determine the Pareto front of stochastic forecast-driven (SF) and demand-driven (SD) models using empirical, normal and uniform distributions. We study the effect of varying the number of scenarios by setting $N \in \{4, 8, 16, 32\}$ for the normal and uniform distributions while the number of empirical scenarios remains equal to $N = 4$. For each scenario tree, a sensitivity analysis of the lost-sales penalty cost factor is performed with $\lambda \in \{1, 2, 5, 10, 15, 20, 25, 30, 50, 200\}$. Similarly, the conservativeness of the deterministic models is adjusted through parameter $z \in \{0, 0.2, 0.4, \ldots, 1.8, 2\}$.

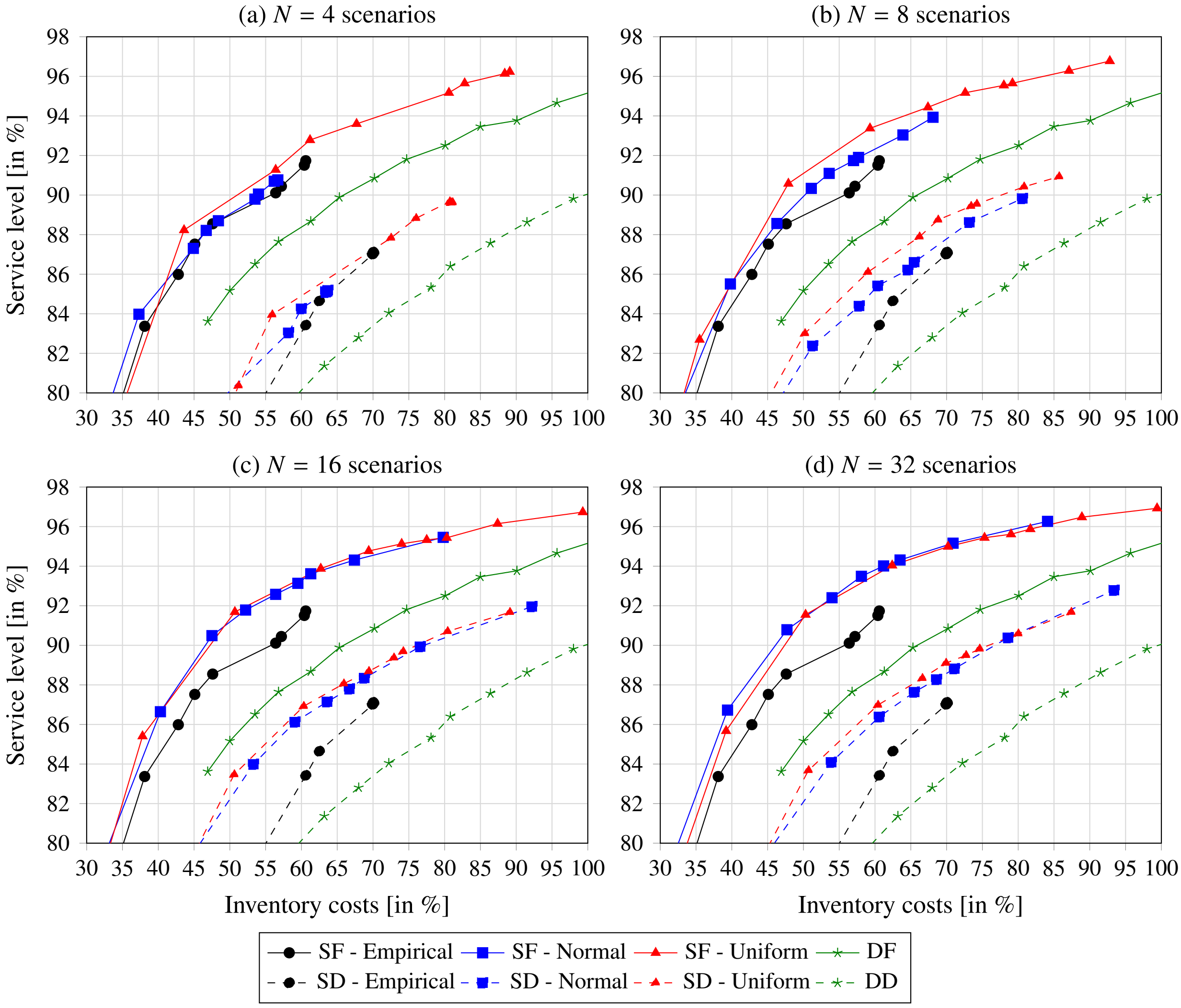

Figure 2: Pareto front between service level and inventory costs for different model configurations.

The Pareto fronts are shown in Figure 2 where each mark corresponds to the average performance over the Y seasons for a given lost-sale penalty cost or deterministic safety factor. Several conclusions can be drawn from the simulation results: (a) forecast-driven models outperform demand-driven models, (b) stochastic models dominate deterministic models with exogenous safety stock calculations, (c) the value of sampling additional scenarios decreases quickly so that only few scenarios are necessary to achieve good performance, and (d) the uniform distribution dominates other distributions for small scenario trees but appears equivalent to the normal distribution for larger tree sizes.

This analysis highlights the importance of the uncertainty modelling step identified in Barrier 2 (Uncertainty definition) when applying stochastic programming from data. Notably, the deterministic forecast-driven model outperforms all stochastic demand-driven models, confirming our intuition that defining uncertainty correctly may be more important that applying advanced stochastic techniques. Still, using a stochastic model instead of a deterministic model provides significant benefits. For small sample sizes, the uniform distribution provides the best results, which may be explained by the fact that it contains more extreme scenarios that allows it to reach high service levels. For large scenario trees, which provides a more accurate evaluation of the distribution quality, the normal and uniform probability distributions yield similar performance. The results suggest that overcoming Barrier 2 (Uncertainty definition) is even more important than Barrier 3 (Uncertainty model), even though the latter has received much more attention in the literature.

6.2.2. Out-of-sample regret

To highlight the importance of performing out-of-sample simulations, we compare the results of in-sample and out-of-samples simulations. We investigate the relative out-of-sample regret, which is defined as the difference between the average performance obtained with in-sample and out-of-sample simulations divided by the in-sample performance.

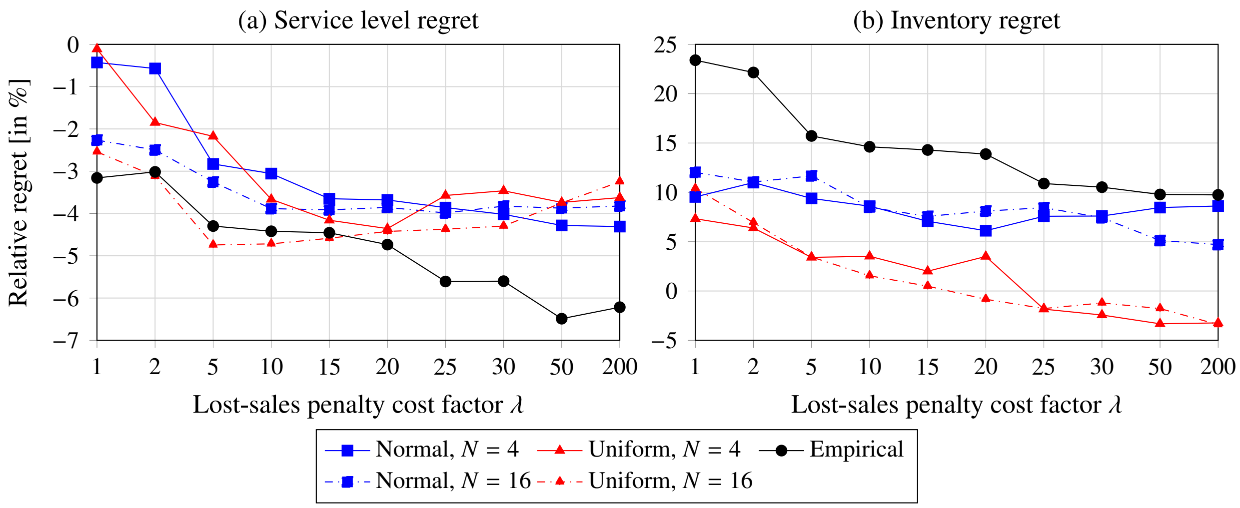

Figure 3: Out-of-sample regret of realised (a) service level and (b) inventory cost.

The relative regret of service level and inventory costs is shown on Figure 3 as a function of the lost-sales penalty factor. The figure shows that all service level regrets are negative. In-sample simulations have an optimistic bias, which is consistent over all uncertainty models. Similarly, the inventory regret shows that out-of-samples inventory costs are overall higher than their in-sample estimates. Interestingly, the empirical distribution shows the highest regret on both the service level and inventory costs. Increasing the size of the scenario trees does not reduce the out-of-sample regret of estimated distributions. On the contrary, it leads to overall higher service level regret.

6.2.3. Summary

Determining the Pareto fronts of the models with different uncertainty process definition and representation allows to overcome Barrier 2 (Uncertainty definition) and Barrier 3 (Uncertainty model). The out-of-sample evaluations are a key component for overcoming Barrier 1 (Data scarcity) and Barrier 8 (Evaluation). They provide an accurate and unbiased estimate of model performance. They also highlight the ability to generalise from past observations by estimating probability distributions and sampling from them. We overcome Barrier 7 (Tractability) by observing that a small scenario tree is enough to provide good out-of-sample performance. In the remainder of the numerical study, we use the forecast-driven stochastic model with $N = 8$ scenarios sampled from the uniform distribution. The lost-sales penalty cost factor is set to $\lambda = 15$, which ensures a satisfying trade-off between between service level and inventory costs.

6.3. Stochastic programming, recourse and planning stability

In this part, we investigate the trade-off between planning flexibility, stability and communicability. First, we illustrate the reference plan obtained with product families. Then, we evaluate the value of recourse and compare the effect of freezing and aggregating production decisions to mitigate nervousness.

6.3.1. Planning communicability

The stochastic model with production recourse presented in Section 5.2.2 determines a reference plan as a combination of capacity reserves and first-stage production decisions. An example is shown in Figure 4 for $F = 4$ product families, where production is shown relative to available capacity in each period. The figure shows the first-stage decisions aggregated over all products as well as the capacity reserves for the four families. It illustrates the variation in volume and timing between the different families over the planning horizon. The capacity reserves can be understood as the volatile part of the plan since they are used differently in each recourse scenario by the products in the family. Hence, a reference plan can be communicated while allowing flexible product-specific decisions.

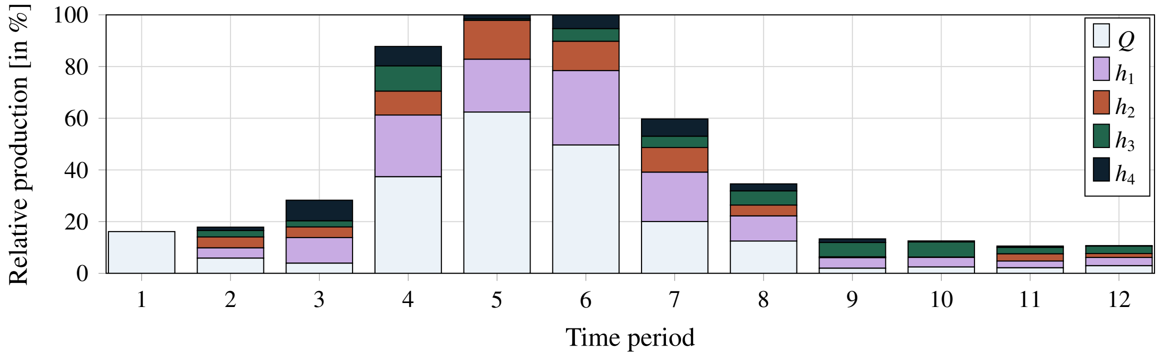

Figure 4: Capacity reserves and first-stage production decisions relative to available capacity.

6.3.2. Raw-material stability

We analyse the impact of freezing raw-material ordering decisions by varying the frozen horizon length τ_a within the set $\{0, 1, 2, 3, 4, 5, 6\}$. The results are shown in Figure 5, which shows that freezing raw-material ordering decisions is an effective strategy to improve raw-material stability although it leads to increased inventory costs.

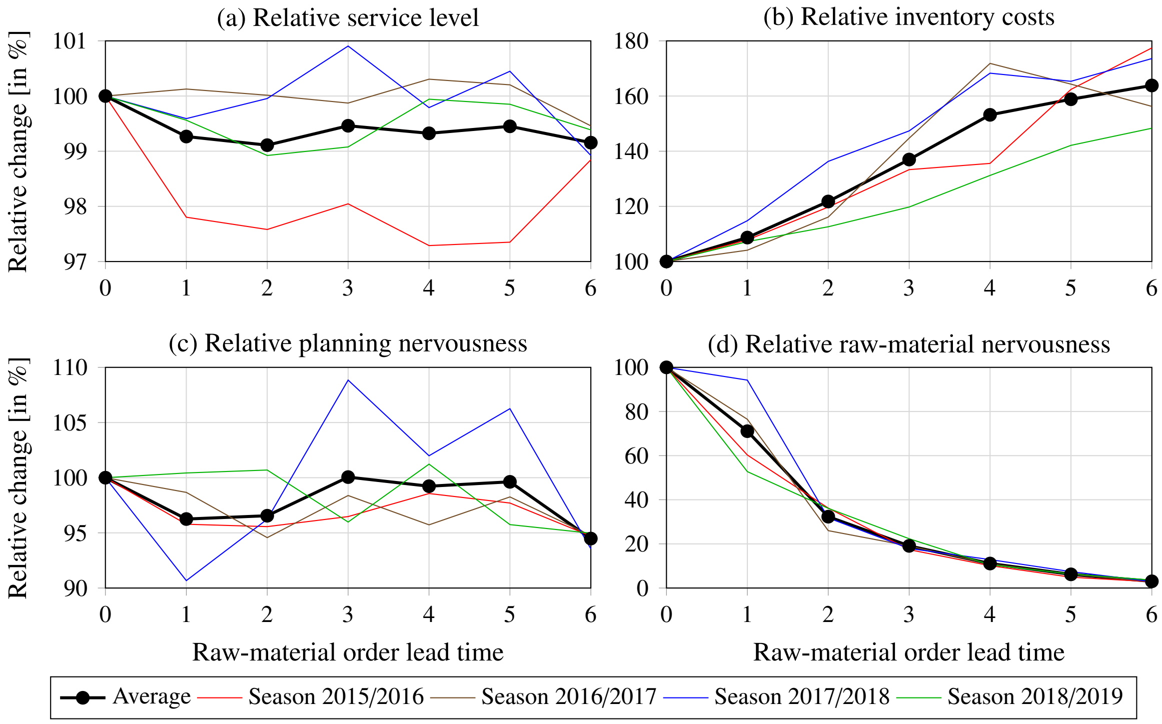

Figure 5: Sensitivity analysis of raw-material ordering lead time.

The stochastic model maintains high service level by increasing safety inventory, suggesting that the scenario tree accurately captures the raw-material uncertainty over the prediction horizon. Interestingly, there is no distinguishable effect on planning nervousness. Freezing raw-material orders does not reduce planning flexibility if enough safety inventory is available on the raw-material level. Hence, we fix the raw-material lead time to $\tau_a = 2$ to decrease raw-material nervousness with acceptable inventory costs increase.

6.3.3. Plan stability

In Section 5, two methods are presented to mitigate planning nervousness. The first strategy freezes production decisions over the short-term horizon while the second aggregates decisions over optimally defined families. We evaluate and compare their performance in a sensitivity analysis. In Figure 6, we show a side-by-side comparison of the effect of increasing the length of the frozen horizon and increasing the number of product families.

As for raw materials, implementing a frozen horizon on the production level gives significant reduction in planning nervousness. However, it leads to small decrease in average service level and comes at the cost of increased inventory costs. Freezing the production horizon also has a stabilising effect on raw-material orders since production flexibility is strongly reduced. On the other hand, the stochastic model with recourse provides high stability, high demand satisfaction and low costs. Since the product-to-family assignment model prioritises the assignment of products with high demand and large forecast errors, few product families are sufficient to observe large improvements in planning stability. As the number of families increases, planning nervousness decreases with diminishing marginal returns. For $F = 4$ families, the model can reduce inventory costs by 6% while the average service level is decreased by only 1% and planning nervousness is reduced by 40%. On the contrary, freezing production decisions overly restricts the flexibility of the model, which may lead to unacceptable cost increase.

6.3.4. Value of recourse under varying capacity utilisation

In the agrochemical industry, capacity is expensive and capacity planning is an important long-term problem. To demonstrate the robustness of our approach in diverse settings, we analyse performance under varying capacity. Available capacity is reduced in 5% increments from 100% to 40%. The average performance are shown in Figure 7 for the stochastic model without recourse corresponding to $F = 0$ family, the stochastic model with recourse and $F = 4$ families, as well as the stochastic model without family and frozen horizon $\tau_k = 1$.

Figure 7 shows that inventory increases as capacity is further reduced, leading to higher costs but also

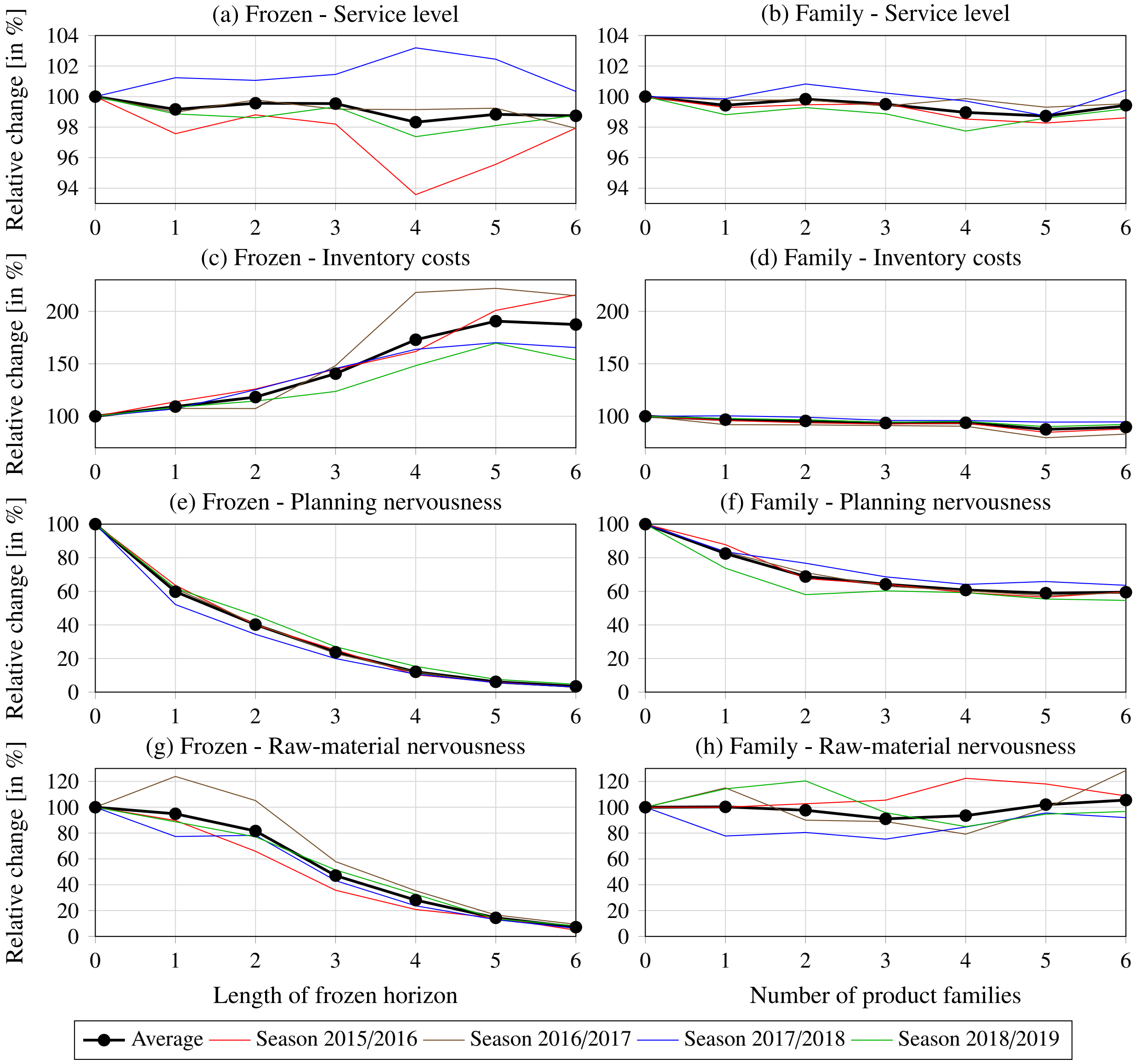

Figure 6: Comparison of nervousness mitigation strategies on planning level.

higher service level. When capacity is severely limited, a steep decline in service level is observed. The simulations highlight that the value of recourse is robust over a wide range of capacity settings. Remarkably, the stochastic model with recourse provides highest service level and lower costs when capacity is highly utilised, which corresponds to capacity reduction of 65% and lower. On the contrary, the stochastic model with frozen production decisions yields the highest inventory costs and lowest service level over all instances. Hence, production flexibility is essential to manage short-term uncertainty when capacity is limited.

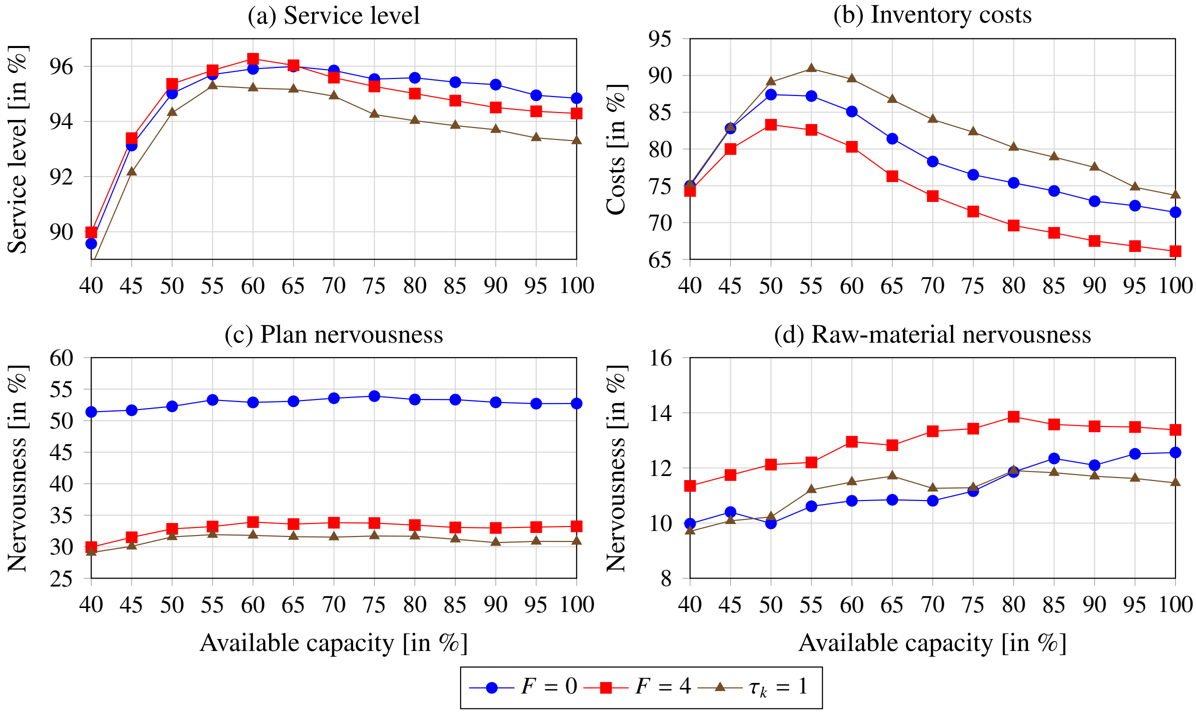

Figure 7: Performance of stochastic models under varying capacity.

It is interesting to observe that realised service level is overall higher when capacity is limited. With little available capacity, production starts earlier and uses less accurate forecasts. Hence, additional safety stock are placed, leading to both higher inventory and service level. Yet, we note that the obtained solutions is a dominated solution on the Pareto front shown in Figure 2. The Pareto analysis performed in Section 6.2 should be performed with the new capacity to decide on the optimal trade-off between service level and inventory costs and identify the lost-sales penalty cost factor that achieves the target service level.

6.3.5. Summary

The barriers linked to the planning processes have been overcome thanks to several strategies. Barrier 5 (Communicability) is solved by explicitly integrating raw-material orders and determining a reference plan on the family level. Barrier 4 (Flexibility representation) is overcome through recourse decisions, which proves especially relevant when capacity is highly utilised. Barrier 6 (Plan stability) is resolved by implementing a frozen horizon on raw-material orders and aggregating decisions over families. We show that there is not necessarily a trade-off between planning stability and flexibility. The proposed approach based on product aggregation is especially successful since it overcomes the above barriers jointly.

6.4. Comparison with industry benchmarks

To conclude the numerical study, we compare our approach to the current practices of our industrial partner. The stochastic model with $N = 8$ scenarios sampled from a uniform distribution and $F = 4$ families is compared to two benchmarks based on the historical data of our industry partner.

6.4.1. Benchmark definition

The *forecast* benchmark assesses the quality of the demand forecast. The service level of the benchmark is measured through a rolling simulation in which the on-hand inventory is set equal to the demand forecast, thus evaluating the forecast accuracy of the first period in the horizon. Planning and raw-material nervousness are determined by applying Equation (8) and Equation (9) respectively using the demand forecasts and forecasts translated into raw materials using the bill-of-material. The forecast benchmark does not lead to inventory costs, which are not reported.

The *company* benchmark represents the practice of our industrial partner. Currently, a combination of deterministic automated MRP software and expert knowledge is used to derive a production plan in each review period. The benchmark is based on the history of production plans and inventory levels. The service level of the company benchmark is measured by comparing the sum of historical on-hand inventory and the production plan implemented in rolling horizon to the demand. The inventory costs are determined from the historical inventory of raw materials and finished goods. Planning nervousness is measured using Equation (8). Raw-material orders are obtained by converting the finished-goods production plan on the raw-material level and by accounting for on-hand raw-material inventory. Raw-material nervousness is then deduced using Equation (9).

6.4.2. Simulation results

The out-of-sample rolling-horizon simulation results are presented in Figure 8 for all seasons. The average results are given in Table 1. The forecast accuracy is poor since the forecast benchmark yields lowest service level in all seasons. Interestingly, the deterministic model provides higher service level although it uses the same demand forecasts. This can be explained by the fact that the deterministic model carries inventory from one period to the next if it produces more than the actual demand. This highlights a bias in the forecasting process: demand tends to be forecast earlier than it actually realises, which leads to inventory build up that is used in later periods. Overall the company benchmark achieves a high service level and outperforms the forecast and deterministic models, highlighting the value of planner expertise.

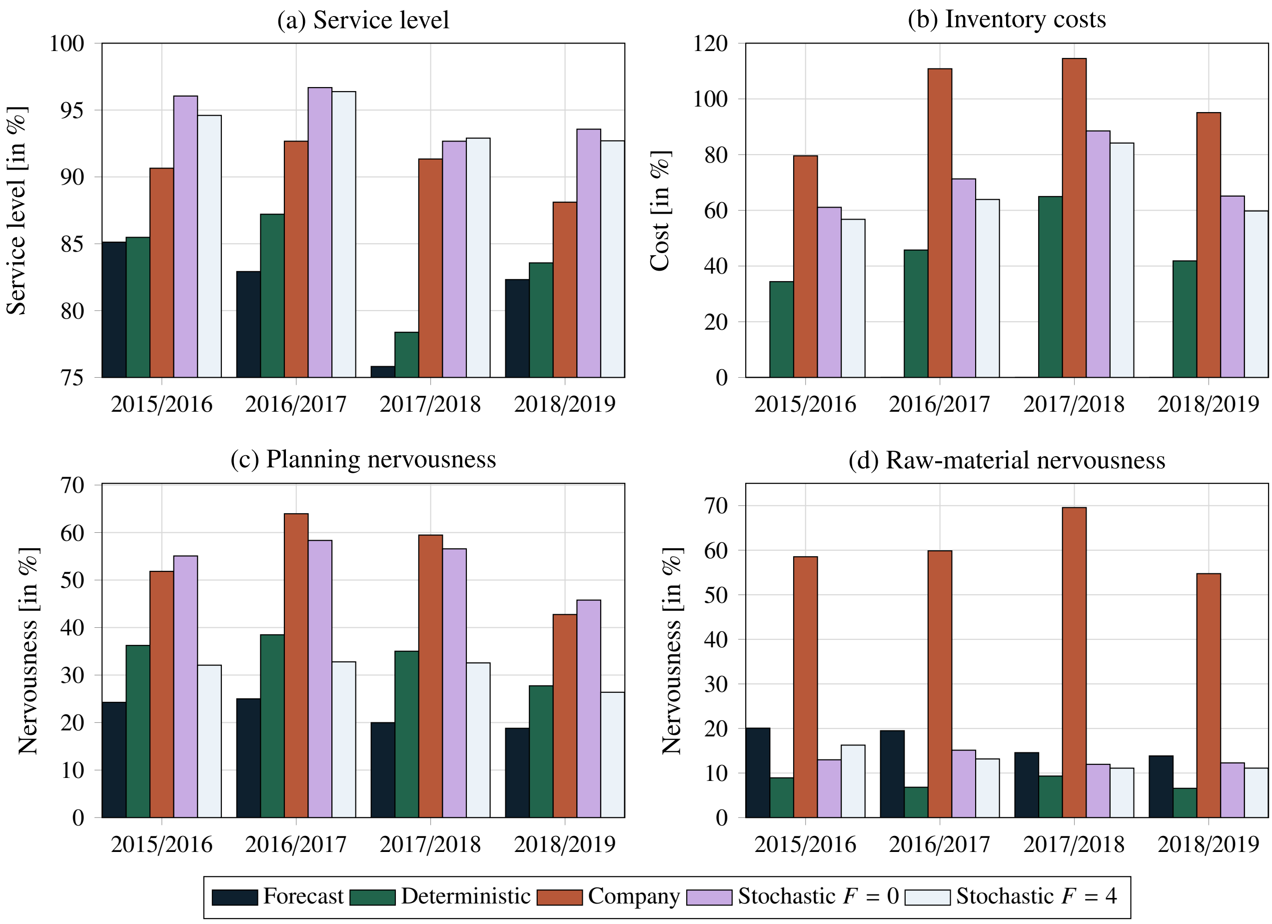

Figure 8: Simulation results over four seasons: (a) service level, (b) inventory, (c) planning nervousness, and (d) raw-material nervousness.

The stochastic model with $F = 4$ families achieves high service level consistently over the four seasons. It reduces inventory costs by more than 33% compared to the company benchmark, which suggests an efficient placement of safety stocks. It also yields substantial improvements in stability as planning nervousness is reduced by 40% thanks to the aggregation of planning decisions on the family level. Raw-material nervousness is reduced by almost 80% on average, which results in lower nervousness than the forecast benchmark. Thus, the planning step acts as a dampening step. Short-term demand variability is effectively mitigated, which provides a robust ordering signal to upstream raw-material planners.

The simulation setting and benchmark definition allows us to overcome Barrier 8 (Evaluation). The results show that the stochastic model with production recourse improves all performance indicators compared to the company historical practice: customer satisfaction is increased, inventory costs are reduced and planning is more stable on both the finished-goods and raw-material orders levels.

Table 1: Average value of KPIs over all seasons and value relative to company.

KPI	Company	Forecast	Deterministic	Stochastic $F = 0$	Stochastic $F = 4$
Service level [in %]	90.69	81.54	83.66	94.74	94.14
Relative [in %]	100	89.9	92.2	104.5	103.8
Relative inv. costs [in %]	100	-	46.7	71.5	66.2
Relative fin.-goods inv. costs [in %]	100	-	33.8	80.3	76.2
Relative raw-mat. inv. costs [in %]	100	-	62.2	61	54.2
Planning nervousness [in %]	54.5	22	34.4	53.9	31
Relative [in %]	100	40.4	63.1	99	56.8
Raw-mat. nervousness [in %]	60.7	17	7.9	13.1	12.9
Relative [in %]	100	28	13	21.6	21.3

7. Conclusion

This paper aims to foster the use of stochastic programming in master production scheduling. First, we identify barriers that challenge the application of stochastic programming and relate them to a real-world case study in the agrochemical industry. Then, we discuss how to model the uncertainty from limited available data and construct scenario trees to represent future demand. The scenario trees are integrated into stochastic planning model that reflect the planning processes. The trade-off between planning communicability, stability and flexibility are integrated in a two-stage stochastic model that determine the optimal recourse production volumes. Finally, we demonstrate our framework on the case study, determine the best model configuration through sensitivity analyses and compare its result to industry practice.

The results of this paper extend beyond the scope of the case study considered. The barriers identified are common to a wide array of manufacturing environment. We hope to stimulate the discussion on their relevance and encourage the development of solutions suitable to varied production settings. The simulation study allows us to emphasise the importance of the definition and representation of the uncertain processes. Whereas existing literature overwhelmingly assumes that uncertainty models are available, we show that the carefully modelling uncertainty is critical. Indeed, a simple deterministic model with the right uncertainty model outperforms advanced stochastic models with inaccurate uncertainty definition.

We discern several directions for future research. The analysis of nervousness mitigation by aggregating production decisions could be extended to a multi-level supply chain. The application of advanced models to characterise the forecast revision process such as the Martingale Model of Forecast Evolution of Heath and Jackson (1994) could be investigated to further exploit available data and derive robust uncertainty models.

Acknowledgements

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - 277991500/GRK2201.

References

Atadeniz, S.N., Sridharan, S.V., 2020. Effectiveness of nervousness reduction policies when capacity is constrained. International Journal of Production Research 58, 4121–4137.

Ban, G.Y., 2020. Confidence intervals for data-driven inventory policies with demand censoring. Operations Research 68, 309–326.

Ben-Tal, A., Den Hertog, D., De Waegenaere, A., Melenberg, B., Rennen, G., 2013. Robust solutions of optimization problems affected by uncertain probabilities. Management Science 59, 341–357.

Bertsimas, D., Gupta, V., Kallus, N., 2018a. Data-driven robust optimization. Mathematical Programming 167, 235–292.

Bertsimas, D., Gupta, V., Kallus, N., 2018b. Robust sample average approximation. Mathematical Programming 171, 217282.

Beutel, A.L., Minner, S., 2012. Safety stock planning under causal demand forecasting. International Journal of Production Economics 140, 637–645.

Bezanson, J., Edelman, A., Karpinski, S., Shah, V.B., 2017. Julia: A fresh approach to numerical computing. SIAM Review 59, 65–98.

Blackburn, J.D., Kropp, D.H., Millen, R.A., 1986. A comparison of strategies to dampen nervousness in MRP systems. Management Science 32, 413–429.

Bollapragada, R., Kuppusamy, S., Rao, U.S., 2015. Component procurement and end product assembly in an uncertain supply and demand environment. International Journal of Production Research 53, 969–982.

Brandimarte, P., 2006. Multi-item capacitated lot-sizing with demand uncertainty. International Journal of Production Research 44, 2997–3022.

Carlson, R.C., Jucker, J.V., Kropp, D.H., 1979. Less nervous MRP systems: A dynamic economic lot-sizing approach. Management Science 25, 754–761.

Chopra, S., Meindl, P., 2013. Supply chain management: Strategy, planning & operation. 5th ed., Pearson Prentice Hall Inc.

Cunha, A.L., Santos, M.O., Morabito, R., Barbosa-Póvoa, A., 2018. An integrated approach for production lot sizing and raw material purchasing. European Journal of Operational Research 269, 923–938.

Dunning, I., Huchette, J., Lubin, M., 2017. JuMP: A modeling language for mathematical optimization. SIAM Review 59, 295–320.

Dupačová, J., Gröwe-Kuska, N., Römisch, W., 2003. Scenario reduction in stochastic programming. Mathematical programming 95, 493–511.

Escudero, L.F., Kamesam, P.V., King, A.J., Wets, R.J., 1993. Production planning via scenario modelling. Annals of Operations Research 43, 309–335.

Gruson, M., Cordeau, J.F., Jans, R., 2021. Benders decomposition for a stochastic three-level lot sizing and replenishment problem with a distribution structure. European Journal of Operational Research 291, 206–217.

Heath, D.C., Jackson, P.L., 1994. Modeling the evolution of demand forecasts with application to safety stock analysis in production/distribution systems. IIE Transactions 26, 17–30.

Heitsch, H., Römisch, W., 2003. Scenario reduction algorithms in stochastic programming. Computational optimization and applications 24, 187–206.

Heitsch, H., Römisch, W., 2009. Scenario tree modeling for multistage stochastic programs. Mathematical Programming 118, 371–406.

Herrera, C., Belmokhtar-Berraf, S., Thomas, A., Parada, V., 2016. A reactive decision-making approach to reduce instability in a master production schedule. International Journal of Production Research 54, 2394–2404.

Ho, C.J., 1989. Evaluating the impact of operating environments on MRP system nervousness. International Journal of Production Research 27, 1115–1135.

Homem-de-Mello, T., Bayraksan, G., 2014. Monte carlo sampling-based methods for stochastic optimization. Surveys in Operations Research and Management Science 19, 56–85.

Hu, Z., Hu, G., 2018. A multi-stage stochastic programming for lot-sizing and scheduling under demand uncertainty. Computers & Industrial Engineering 119, 157–166.

Huber, J., Müller, S., Fleischmann, M., Stuckenschmidt, H., 2019. A data-driven newsvendor problem: From data to decision. European Journal of Operational Research 278, 904–915.

Jensen, T., 1993. Measuring and improving planning stability of reorder-point lot-sizing policies. International Journal of Production Economics 30, 167–178.

Kanyalkar, A.P., Adil, G.K., 2010. A robust optimisation model for aggregate and detailed planning of a multi-site procurement-production-distribution system. International Journal of Production Research 48, 635–656.

Kazemi Zanjani, M., Nourelfath, M., Ait-Kadi, D., 2010. 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, 4701–4723.

King, A.J., Wallace, S.W., 2012. Modeling with stochastic programming. Springer Science & Business Media, New York.

Klabjan, D., Simchi-Levi, D., Song, M., 2013. Robust stochastic lot-sizing by means of histograms. Production and Operations Management 22, 691–710.

Kleywegt, A.J., Shapiro, A., Homem-de-Mello, T., 2002. The sample average approximation method for stochastic discrete optimization. SIAM Journal on Optimization 12, 479–502.

Li, Q., Disney, S.M., 2017. Revisiting rescheduling: MRP nervousness and the bullwhip effect. International Journal of Production Research 55, 1992–2012.

Lin, P.C., Uzsoy, R., 2016. Chance-constrained formulations in rolling horizon production planning: an experimental study. International Journal of Production Research 54, 3927–3942.

Meistering, M., Stadtler, H., 2017. Stabilized-cycle strategy for capacitated lot sizing with multiple products: fill-rate constraints in rolling schedules. Production and Operations Management 26, 2247–2265.

Mišić, V.V., Perakis, G., 2020. Data analytics in operations management: A review. Manufacturing & Service Operations Management 22, 158–169.

Oroojlooyjadid, A., Snyder, L.V., Takáč, M., 2020. Applying deep learning to the newsvendor problem. IISE Transactions 52, 444–463.

Prak, D., Teunter, R., 2019. A general method for addressing forecasting uncertainty in inventory models. International Journal of Forecasting 35, 224–238.

Prak, D., Teunter, R., Syntetos, A., 2017. On the calculation of safety stocks when demand is forecasted. European Journal of Operational Research 256, 454–461.

Saliby, E., 1990. Descriptive sampling: a better approach to Monte Carlo simulation. Journal of the Operational Research Society 41, 1133–1142.

Shapiro, A., 2011. Analysis of stochastic dual dynamic programming method. European Journal of Operational Research 209, 63–72.

Silver, E.A., Pyke, D.F., Thomas, D.J., 2016. Inventory and Production Management in Supply Chains. 4th ed., CRC Press, Boca Raton.

Sridharan, S.V., Berry, W.L., Udayabhanu, V., 1988. Measuring master production schedule stability under rolling planning horizons. Decision Sciences 19, 147–166.

Stephan, H.A., Gschwind, T., Minner, S., 2010. Manufacturing capacity planning and the value of multi-stage stochastic programming under markovian demand. Flexible Services and Manufacturing Journal 22, 143–162.

Tavaghof-Gigloo, D., Minner, S., 2020. Planning approaches for stochastic capacitated lot-sizing with service level constraints. International Journal of Production Research , 1–21.

Thevenin, S., Adulyasak, Y., Cordeau, J.F., 2021. Material requirements planning under demand uncertainty using stochastic optimization. Production and Operations Management 30, 475–493.

Trapero, J.R., Cardós, M., Kourentzes, N., 2019. Empirical safety stock estimation based on kernel and garch models. Omega 84, 199–211.

Tunc, H., Kilic, O.A., Tarim, S.A., Eksioglu, B., 2013. A simple approach for assessing the cost of system nervousness. International Journal of Production Economics 141, 619–625.

Watson, J.P., Woodruff, D.L., 2011. Progressive hedging innovations for a class of stochastic mixed-integer resource allocation problems. Computational Management Science 8, 355–370.

Wiesemann, W., Kuhn, D., Sim, M., 2014. Distributionally robust convex optimization. Operations Research 62, 1358–1376.

Zhao, X., Xie, J., Jiang, Q., 2001. Lot-sizing rule and freezing the master production schedule under capacity constraint and deterministic demand. Production and Operations Management 10, 45–67.

Appendix A. Notation of stochastic models.

Table A.1: Notation of stochastic models

Sets	
$\mathcal{A}$	Set of raw materials $\{1, \ldots, A\}$
$\mathcal{L}$	Set of production lines $\{1, \ldots, L\}$
$\mathcal{K}$	Set of products $\{1, \ldots, K\}$
$\mathcal{K}_f$	Set of products within family f
$\mathcal{T}$	Set of time periods $\{1, \ldots, T\}$
$\mathcal{W}$	Set of production sites $\{1, \ldots, W\}$
$\mathcal{L}_w$	Set of lines $\{1, \ldots, L_w\}$ at site w
Parameters	
$f_{k,t}$	Demand forecast for product k in period t
κ_l	Capacity of line l
γ_k	Lost-sale penalty cost of product k
$\beta_{k,a}$	Consumption of raw material a per unit of product k
$\mu_{k,w}$	Inventory holding cost of product k at site w
ν_a	Inventory holding cost of raw material a
$\rho_{k,l}$	Equal to 1 if product k can be produced on line l, 0 otherwise
$x_{k,f}$	Product-to-family assignment, equal to 1 if product k is assigned to family f
Decision variables	
$q_{k,l,t}$	Production volume of product k on line l in period t
$z_{a,t}$	Order of raw material a for period t
$b_{k,t,n}$	Lost sales of product k at the end of period t in scenario n
$s_{k,w,t,n}$	Sales of product k assigned to site w in period t in scenario n
$i_{k,w,t,n}$	Inventory of product k on hand at site w at the end of period t in scenario n
$r_{k,l,t,n}$	Recourse production of product k on line l in period t in scenario n
$Y_{a,t,n}$	Inventory of raw material a at the end of period t in scenario n
$u_{k,t,n}$	Auxiliary variable to track minimum recourse production over all scenarios n for product k in period t
$v_{f,t,n}$	Auxiliary variable to track maximum recourse production over all scenarios n for family f in period t

Appendix B. Deterministic model

The deterministic production planning and raw-material ordering problem can be formulated in Problem (B.1) as a deterministic optimisation problem.

$$\min \quad \sum_{t=1}^{T}\left(\sum_{k=1}^{K}\sum_{w=1}^{W}\mu_{k,w} i_{k,w,t} + \sum_{a=1}^{A}\nu_a \cdot y_{a,t} + \sum_{k=1}^{K}\gamma_k \cdot b_{k,t}\right) \tag{B.1a}$$

$$\text{s.t.} \quad i_{k,w,t} = i_{k,w,t-1} + \sum_{l\in\mathcal{L}_w} q_{k,l,t} - s_{k,w,t}, \quad \forall k,w,t \tag{B.1b}$$

$$f_{k,t} + ss_{k,t} = b_{k,t} + \sum_{w=1}^{W} s_{k,w,t}, \quad \forall k,t \tag{B.1c}$$

$$\sum_{k=1}^{K} q_{k,l,t} \leq \kappa_l, \quad \forall l,t \tag{B.1d}$$

$$y_{a,t} = y_{a,t-1} + z_{a,t} - \sum_{k=1}^{K}\sum_{l=1}^{L}\beta_{k,a} \cdot q_{k,l,t}, \quad \forall a,t \tag{B.1e}$$

$$q_{k,l,t}, i_{k,w,t}, b_{k,t}, s_{k,t}, y_{a,t}, z_{a,t} \geq 0, \quad \forall k,w,l,a,t \tag{B.1f}$$

The objective function in (B.1a) minimises the inventory costs of finished goods and raw materials as well as the lost-sales costs over the planning horizon. The lost-sale costs relax the problem to allow feasible solutions when capacity or raw materials are insufficient to satisfy demand. As such, the lost-sales penalty cost is typically set to a high value ($\lambda = 1000$ in our numerical study). Constraint (B.1b) describes the inventory balance of finished goods in each site. Constraint (B.1c) ensures that demand is either satisfied by each site's production or counted as lost sales. Constraint (B.1d) limits the production of each line to its capacity in each period. Constraint (B.1e) describes the raw-material inventory balance. Constraint (B.1f) specifies the domain of the continuous decision variables.