Data-Driven Simulation and Optimization Approaches to Incorporate Production Variability in Sales and Operations Planning

Bruno A. Calfa*, Anshul Agarwal†, Scott J. Bury†, John M. Wassick†, Ignacio E. Grossmann*

May 18, 2015

Abstract

We propose two data-driven, optimization-based frameworks (simulation-optimization and bi-objective optimization) to account for production variability in the operations planning stage of the Sales and Operations Planning (S&OP) of an enterprise. Production variability is measured as the deviation between historical planned (target) and actual (achieved) production rates. A statistical technique, namely, quantile regression is used to model the distribution of deviation values given planned production rates. Scenarios are constructed by sampling from the distribution of deviation values and used as inputs to the proposed optimization-based frameworks. Advantages and disadvantages of the two proposed frameworks are discussed. The applicability of the proposed methodology is illustrated with a detailed analysis of the results of a motivating example and a real-world production planning problem from a chemical company.

Keywords: Sales and Operations Planning, Data-Driven Optimization, Production Variability, Multi-Objective Optimization, Quantile Regression

1 Introduction

Sales and Operations Planning (S&OP) is a business and decision-making process through which a company makes certain that tactical plans in every business area balance demand and supply for products. Therefore, S&OP links the corporate strategic plan to daily operations plans. The overall result of the S&OP process is an operating plan to allocate company resources (Grimson & Pyke, 2007).

Attempts in the literature to systematically survey case studies of the S&OP process adopt the Capability Maturity Model (CMM) (Paulk, Weber, & Chrissis, 1993). A recent

*Department of Chemical Engineering. Carnegie Mellon University. Pittsburgh, PA 15213, USA.
†The Dow Chemical Company. Midland, MI 48674, USA.

review that surveys some maturity models can be found in Thomé *et al.* (2012). Maturity models define stages for the S&OP process that include activities such as meetings, demand forecasting, integration of procurement, production, and distribution plans, and performance measurements. Figure 1 illustrates typical stages in the S&OP process. We note that uncertainty and variability affect decisions in both stages 1 (Sales Planning) and 2 (Operations Planning). Forecasting demand for products takes into account future market conditions that are not known exactly (i.e., demand uncertainty), and the operation of plants is subject to unplanned events and imperfect implementation (i.e., production variability).

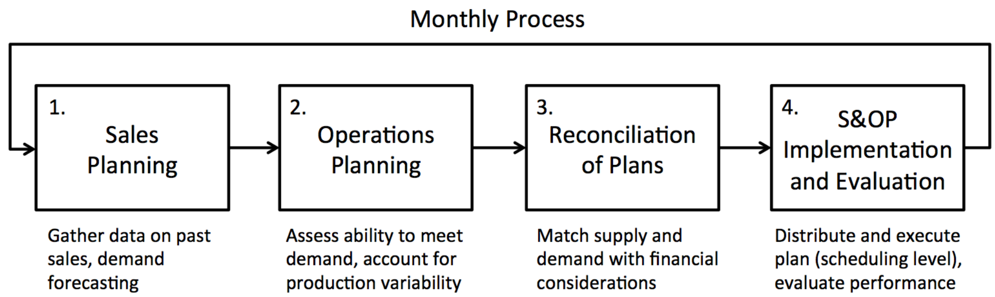

Figure 1: Typical stages in the S&OP process. Adapted from Ling & Goddard (1988).

The Operations Research and Management Science (OR&MS) and Process Systems Engineering (PSE) communities have contributed with optimization-based tactical production planning models as well as solution strategies for different industry sectors. The review papers by Mula *et al.* (2006) and Sodhi & Tang (2009), and the book by Pinedo (2009) present an extensive literature survey about models for tactical production planning and scheduling with uncertainty considerations, and classify the literature based on the production planning area and the modeling approach. Production planning and scheduling problems in the chemical, petrochemical, and pharmaceutical industries are reviewed in the works by Sung & Maravelias (2007); Verderame *et al.* (2010). The authors identify typical sources of uncertainty in different applications and how they are usually modeled in the context of optimization under uncertainty.

Some works have considered hybrid simulation and optimization to account for uncertainty in generating tactical production plans. For instance, Li, González, & Zhu (2009) studied a dedicated remanufacturing system of electronic products with stochastic batch arrival times. The system is modeled in the discrete-event simulator Arena® 10.0 by Rockwell Software, and its objective is to analyze the effect of operational changes on the profit performance of this dedicated remanufacturing system. The optimization approach was based on Genetic Algorithms. Lim, Alpan, & Penz (2014) propose a simulation-optimization approach for managing S&OP in a problem related to the automotive industry. The stochastic parameter is the demand that is assumed to be uniformly distributed. The simulation part is coded in the Java programming language, and the optimization formulation accounts for multiple criteria through ϵ-constraints (see Subsection 4.2). We note that, on the one hand, there has been a modest effort to address the effect of uncertainty in the activities pertaining to sales and procurement planning; on the other hand, there is limited work on incorporating production variability in the operations planning stage.

In this work, we focus on the operations planning stage of the S&OP process (see Figure 1). More specifically, we propose two data-driven optimization-based approaches to account for uncertainty (in this case, production variability) when generating a tactical production plan. The production variability is quantified as the deviation between historical planned and actual production rates. The statistical technique of quantile regression (Koenker, 2005) is used to model the distribution of deviation values for a given planned rate. This distribution is then sampled from in order to construct scenarios. The main contributions of this work are summarized below.

- Statistical modeling via quantile regression of historical production data to quantify production variability;
- Simulation-optimization and bi-objective optimization frameworks to account for production variability when generating a tactical production plan in the S&OP process; and
- Generated tactical production plan with tradeoff information between average profit and risk (i.e., Pareto efficient frontier).

This paper is organized as follows. Section 2 defines the problem and presents the high-level methodology. Section 3 provides a brief overview of classical and quantile regression, which are statistical techniques that can be used to model historical production variability. Section 4 describes the two proposed optimization-based solution strategies to incorporate production variability in the operations planning stage of the S&OP process. The proposed approach is illustrated by two numerical examples in Section 5: motivating example and industrial case study. In the latter, we propose modeling approaches to account for highly integrated networks. Conclusions are drawn in Section 6.

2 Problem Statement and Methodology

The approach proposed in this work is illustrated with an application related to the Chemical Process Industry (CPI). In particular, we deal with Enterprise-wide Optimization (EWO) decisions of highly-integrated chemical production sites, which produce basic chemicals and their downstream derivatives (Wassick, 2009). However, we note that the proposed methodology is general and can be applied to other types of industries.

For this problem, we consider a process network of chemical plants and focus on the effects of production variability in the operations planning stage. For simplicity, we assume that the future monthly demand is given and is deterministic over the planning horizon. Also given is the minimum/maximum installed production capacity of each plant and its production costs. Transportation and inventory holding costs, inventory capacity, and initial inventory are given. Future planned maintenance outages of production plants may also be given. Given a multi-period linear programming (LP) production planning model for the given process network (similar to the model presented in Sahinidis *et al.* (1989), excluding capacity expansion considerations), the objectives are two-fold: (1) propose a production plan incorporating historical production variability data, and (2) measure the performance of the proposed plan in the form of a tradeoff between average profit and risk. Production variability is incorporated in a two-stage the stochastic programming production planning formulation whose details are given in Subsection 4.2.

The overall strategy to generate production plans by taking into account the variability of S&OP data is shown in Figure 2. From available historical data that consist of actual and planned production rates, quantile regression models (see Section 3) are built and used to characterize the production variability, i.e., deviation between planned and actual rates. Deviation values are then sampled from these statistical models and used in an optimization-based framework to generate production plans with profit vs. risk tradeoff information (see Section 4). We propose and discuss the advantages and disadvantages of two different frameworks: (1) a simulation-optimization approach, and (2) a bi-objective optimization approach.

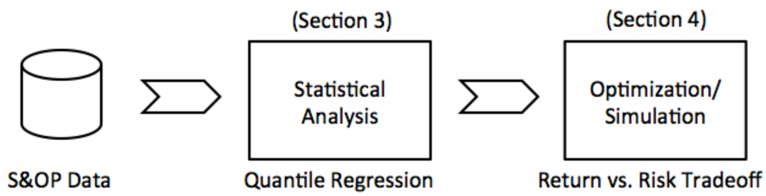

Figure 2: Overall strategy to account for historical production variability when generating production plans.

3 Modeling Production Variability with Quantile Regression

In this section, we first provide a brief overview of classical and quantile regression analysis, then we illustrate how the latter can be used to model production variability. The notation for this section is as follows. Let X and Y denote the predictor (or covariate or input) and the response (or output) random variables whose values are denoted by x and y, respectively. For example, we may define X as the planned production rate and Y as the deviation between planned and actual rates. We restrict the discussion to continuous random variables only.

In regression analysis, the regression model is generally written as,

$$Y = f(X) + \epsilon \tag{1}$$

where $f(\cdot)$ is a mathematical formula that expresses the relationship between X and Y, and ϵ is the random error term assumed to have mean zero, homoskedastic (i.e., its variance is constant over the range of X values), and uncorrelated with X (Montgomery & Runger, 2003). In linear (classical) regression, for example, $f(X) = \beta_0 + \beta_1 X$, where β_0 and β_1 are parameters to be estimated. Nonlinear parametric and nonparametric functions can also be used to model the relationship between X and Y.

We want to predict Y values for given X values. In classical regression, we write

$$\hat{Y} = \mathbb{E}[Y|X = x] \tag{2}$$

where $\hat{Y}$ denotes the predicted response variable and $\mathbb{E}[\cdot]$ is the expectation operator. In other words, $\hat{Y}$ is the *mean* of Y values conditional on $X = x$. Therefore, for a distribution of

X, the result of a classical regression analysis is a single point (the mean) of the distribution of Y.

A more general approach to regression analysis is quantile regression (Koenker, 2005). A quantile is the value that divides a data set in two subsets. The 50^{th} 100-quantile (also called 50^{th} percentile or median) separates the higher half of a data set from the lower half. In other words, there is at most 50% probability that a random variable will be less than the median. The 4-quantiles are called quartiles, the 5-quantiles are called quintiles and so on. In this paper, we will use quantile and 100-quantile interchangeably.

In quantile regression, the predicted Y values correspond to quantiles of the distribution of Y conditional on $X = x$. Mathematically,

$$\hat{Y} = Q_\alpha[Y|X = x] \tag{3}$$

where $Q_\alpha[\cdot]$ denotes the 100α-th quantile and $\alpha \in [0, 1]$ is the probability level. Similarly to classical regression, quantile regression can be performed parametrically (linear and nonlinear models such as smoothing splines) as well as nonparametrically (e.g., kernel smoothing (Li & Racine, 2007)).

Figure 3 illustrates regression analysis in the general case. In the general case of regression analysis, the objective is to model the relationship between the distribution of X and the distribution of Y, i.e., to predict $F_Y(y)$ from $F_X(x)$, where $F_Y(\cdot)$ and $F_X(\cdot)$ are the cumulative distribution functions of Y and X, respectively. Classical regression provides the mean whereas quantile regression provides *any* quantile of the distribution of Y conditional on $X = x$.

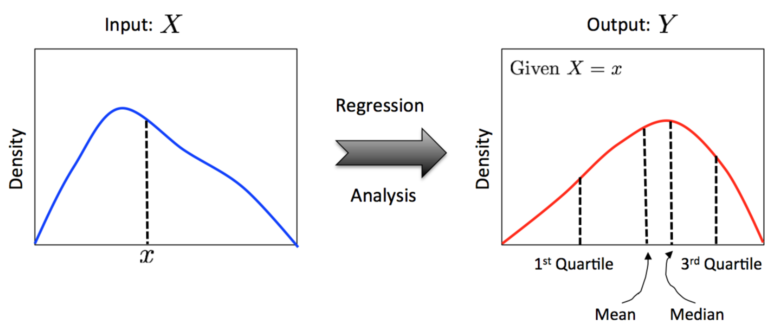

Figure 3: General case of regression analysis. Classical regression only provides mean output, whereas quantile regression provides any point of the output distribution.

The proposed approach for using quantile regression to model production variability in S&OP data is as follows. From historical planned and actual production rates, the deviation between them is calculated as $\Delta = \text{Plan} - \text{Actual}$. In this paper, we consider Δ to be "static", i.e., it is not a stochastic process that evolves in time. The regression analysis consists of regressing Δ on Plan, i.e., obtain the regression function $Q_\alpha[\Delta|\text{Plan} = \text{planned value}]$ for a given probability level α. Finally, a distribution of Δ values given a planned value can be obtained by estimating quantiles for several probability levels (e.g., $\alpha = \{0,\ 0.01,\ \ldots,\ 1\}$). Figure 4 shows an example of a Δ vs. Plan plot for a given chemical plant and the estimated quantiles conditional on two different planned values. The top plot shows that the distribution of Δ values (i.e., conditional quantiles) varies depending on the planned value.

This can also be seen from the bottom plots, where the range of Δ values is larger for the planned value of 10 w.u. (weight units, bottom left plot) than for the planned value of 1.5 w.u. (bottom right plot).

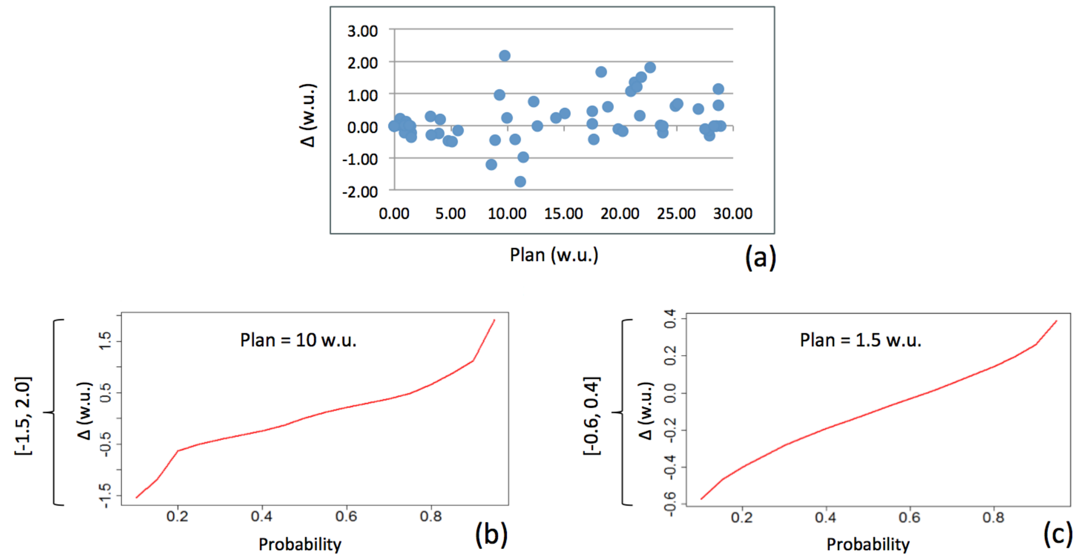

Figure 4: Example of modeling production variability with quantile regression, where Δ = Plan − Actual. Legend: (a) Δ vs. Plan plot, (b) estimated quantiles conditional on Plan = 10 w.u. (weight units), and (c) estimated quantiles conditional on Plan = 1.5 w.u..

Remark. As will be discussed in the next section, it may be difficult or impossible to employ a quantile regression model to generate samples within an optimization formulation. One possible approximation is to disregard the covariate in the quantile regression (i.e., the Plan values), and generate samples from the distribution of the Δ values alone. This is the approach taken in the numerical examples discussed later in this paper.

4 Simulation and Optimization Frameworks

The objective of the proposed approach is two-fold: (1) account for historical production variability when generating an optimal Sales and Operations Planning (S&OP) production plan, and (2) provide tradeoff information about the generated plan in the form of average profit vs. risk. We describe two optimization-based frameworks whose potential advantages and disadvantages are listed in Table 1. The frameworks are detailed in the next two sub-sections.

Table 1: Potential advantages (+) and disadvantages (−) of optimization-based frameworks. Legend: Sim-Opt = Simulation-Optimization framework, Bi-Opt = Bi-Objective Optimization framework, $\Delta|$Plan = deviation conditional on planned values in the context of quantile regression (see Section 3), DFO = Derivative-Free Optimization.

Sim-Opt	Bi-Opt
+ Easy to accommodate for arbitrary $\Delta|$Plan;	+ Simultaneous generation of plan and minimization of risk;
− Expensive simulations as number of scenarios increases;	+ Explicit handling of constraint violation;
− No explicit handling of constraint violation;	− Difficult or impossible to accommodate for arbitrary $\Delta|$Plan;
− Decrease in efficiency if high-dimensional DFO problem.	− Optimization model may be large and nonlinear.

4.1 Simulation-Optimization Framework (Sim-Opt)

The Simulation-Optimization framework (Sim-Opt) consists of alternating between a simulator and a Derivative-Free Optimization (DFO) (Conn, Scheinberg, & Vicente, 2009) solver as illustrated in Figure 5. The purpose of the DFO solver is to set the production target (i.e., generate the S&OP production plan). For a new proposed plan, Δ values are generated using quantile regression and used to form scenarios to be evaluated by the simulator. In the simulator, the production rates of the plants that are subject to variability are fixed to the respective plan proposed by the DFO minus the respective Δ value, i.e., Production Rate = Production Target − Δ. Recall that Δ = Plan − Actual; therefore, by subtracting the estimated Δ value from the production target, we estimate the actual production rate for a plant.

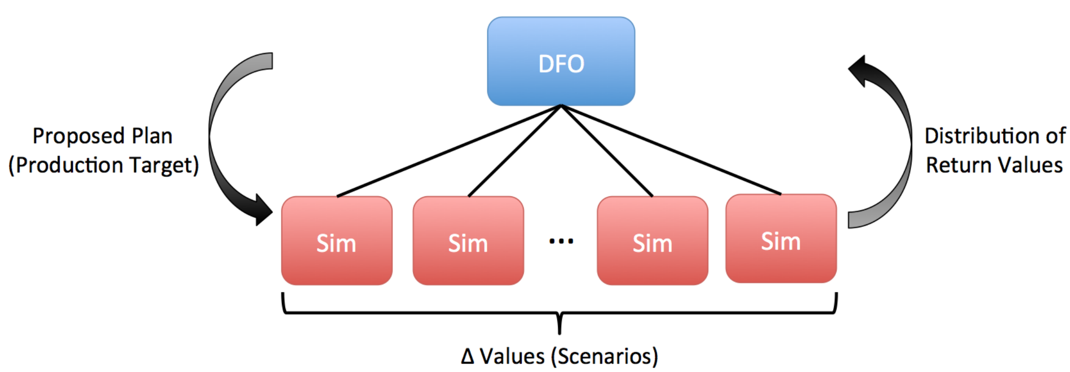

Figure 5: Schematic of the Sim-Opt framework. The DFO solver sets the production targets based on which the Δ values are estimated and form scenarios for the simulator. The simulator evaluates the current production target and returns a distribution of financial performance values (profit, cost etc.).

The simulator can be a black-box S&OP software or a production planning optimization model that acts as a simulator by fixing certain "input" variables, such as production rates. Different scenarios containing Δ values are passed to the simulator, which evaluates the impact of the proposed plan on a performance metric, such as profit or cost. Note that each scenario is independent of the other, which makes this approach amenable to parallelization.

A recent review on DFO solvers and algorithms is given by Rios & Sahinidis (2013). We note that some limitations of DFO include the decrease in efficiency for high-dimensional optimization problems and the number of necessary function evaluations (i.e., calls to the simulator, which may be computationally expensive due to the number of scenarios) in order to achieve significant progress. In the type of problem addressed in this paper, if the production rate of a plant is indexed by time periods and the planning time horizon considered has twelve time periods, then for each plant that is subject to variability, twelve decision variables are needed.

One of the objectives of this paper is to show how a tradeoff curve of average profit vs. risk (Pareto efficient frontier) of the proposed production plan can be used in the analysis of results. This tradeoff curve can be obtained by applying a bi-criterion approach to the Sim-Opt framework as explained as follows. For the case of a black-box simulator, in which no model is available, multi-objective DFO algorithms can be used to construct the Pareto efficient frontier. The literature on multi-objective DFO is generally divided into two classes: Direct Search Methods (DSM) of directional type and Evolutionary Multi-objective Optimization (EMO) algorithms. Reviews are given in Zhou *et al.* (2011); Custódio, Emmerich, & Madeira (2012). If an optimization model is used as the simulator, then standard multi-objective optimization techniques can be used, such as the ϵ-constraint method (see Subsection 4.2) (Hwang & Masud, 1979; Miettinen, 1999).

4.2 Bi-Objective Optimization Framework (Bi-Opt)

The Bi-Objective Optimization framework (Bi-Opt) simultaneously proposes a production plan and minimizes the risk of operating a such plan under production variability consideration. This is accomplished by solving an optimization problem, which requires the production planning optimization model to be fully known. The proposed optimization model is a bi-objective two-stage stochastic program (Birge & Louveaux, 2011) whose first-stage variables are the production targets and second-stage variables are the remaining variables of the model (e.g., flows and inventory). The two objectives are the average profit value to be maximized and a risk measure (e.g., financial risk) to be minimized.

A general deterministic equivalent model of the bi-objective two-stage stochastic program is as follows,

$$\begin{aligned} \min_{x_s, PT} \quad & \text{Risk}(x_s, PT) \\ \max_{x_s, PT} \quad & f_0(PT) + \sum_{s \in S} \text{p}_s f_s(x_s) \\ \text{s.t.} \quad & g(x_s) \leq 0 \qquad \forall\, s \in S \\ & PR_s = PT - \Delta_s \qquad \forall\, s \in S \end{aligned} \tag{4}$$

where the first objective minimizes risk while the second objective maximizes the expected profit. In equation (4), PT are the first-stage production target variables, while x_s is a vector of two-stage decision variables (including the production rate variables, PR_s) that is defined for each scenario $s \in S$, p_s is a constant vector of probability of scenario s, $f_0(\cdot)$, $f_s(\cdot)$, and $g(\cdot)$ are linear functions that define the multi-period LP planning model. In this paper, functions $f_0(\cdot)$ and $f_s(\cdot)$ correspond to the first- and second-stage profit terms, respectively, whereas functions $g(\cdot)$ represent linear material and inventory balances and capacities.

A major contribution of this paper is to model production variability as shown in the last set of equality constraints of equation (4). This set of constraints fixes PR_s to the production target (PT) minus the deviation value for a given scenario s (Δ_s). Recall that subtracting Δ_s from the production target (i.e., production plan) results in the estimated actual production rate. Also, note that PT is not indexed by scenarios, since it is a vector of first-stage variables. The production target is the production plan that is sought to be implemented in practice.

The bi-objective optimization problem in equation (4) can be cast as a single-objective model using standard multi-objective optimization techniques as mentioned in the previous section. The ϵ-constraint method results in two possible models,

$$\begin{aligned} \max_{x_s, PT} \quad & f_0(PT) + \sum_{s \in S} \mathrm{p}_s f_s(x_s) \\ \text{s.t.} \quad & \mathrm{Risk}(x_s, PT) \leq \epsilon \\ & g(x_s) \leq 0 \qquad \forall\, s \in S \\ & PR_s = PT - \Delta_s \qquad \forall\, s \in S \end{aligned} \tag{5}$$

or,

$$\begin{aligned} \min_{x_s, PT} \quad & \mathrm{Risk}(x_s, PT) \\ \text{s.t.} \quad & f_0(PT) + \sum_{s \in S} \mathrm{p}_s f_s(x_s) \geq \epsilon \\ & g(x_s) \leq 0 \qquad \forall\, s \in S \\ & PR_s = PT - \Delta_s \qquad \forall\, s \in S \end{aligned} \tag{6}$$

where ϵ is a threshold value that represents the maximum risk (equation (5)) or minimum average profit (equation (6)) the decision maker is willing to have. The Pareto efficient frontier can be constructed by varying the value of ϵ and resolving the optimization problem.

If the objective function to be maximized in equation (4) is a financial performance indicator (e.g., profit or cost, if minimization), then one possibility is to use a financial risk measure for the expression of $\mathrm{Risk}(\cdot, \cdot)$. Different financial risk management strategies have been proposed in the literature (see Sarykalin, Serraino, & Uryasev (2008) for a tutorial). Some of these risk measures are presented below. Note that each measure operates on the distribution of values of the financial performance indicator.

- *Variance*: It is a measure of the spread of a distribution that operates symmetrically on all values with respect to the expected value. Minimization of variance can be interpreted as the minimization of the square of the L^2-norm between the financial performance in a scenario and the average financial performance over all scenarios.

$$\mathrm{Risk}_2(x_s, PT) = \sum_{s \in S} \mathrm{p}_s \left[f_s(x_s, PT) - \bar{f}(x_s, PT) \right]^2 \tag{7}$$

where $\bar{f}(x_s, PT) = \sum_{s \in S} \mathrm{p}_s f_s(x_s, PT)$ is the expected value of the distribution of profit values.

- *Semivariance*: It is a deviation measure similar to the variance, but it operates on values above or below the expected value. It is also similar to downside risk where the threshold is the expected value of a distribution.

$$\mathrm{Risk}_{2+}(x_s, PT) = \sum_{s \in S} \mathrm{p}_s \left[f_s(x_s, PT) - \bar{f}(x_s, PT) \right]_+^2 \tag{8}$$

$$\mathrm{Risk}_{2-}(x_s, PT) = \sum_{s \in S} \mathrm{p}_s \left[\bar{f}(x_s, PT) - f_s(x_s, PT) \right]_+^2 \tag{9}$$

where $[a]_+ = \max\{0, a\}$.

- *Mean Absolute Deviation (MAD)*: The MAD (also known as the average absolute deviation about the mean) also measures the dispersion of a distribution, but in an absolute sense. Analogously to the variance, its minimization can be seen as the minimization of an L^1-norm.

$$\mathrm{Risk}_1(x_s, PT) = \sum_{s \in S} \mathrm{p}_s \left| f_s(x_s, PT) - \bar{f}(x_s, PT) \right| \tag{10}$$

- *Maximum Absolute Deviation*: It is analogous to the MAD, but its minimization is equivalent to minimizing the L^∞-norm (Cai *et al.*, 2000).

$$\mathrm{Risk}_\infty(x_s, PT) = \max_{s \in S} \sum_{s \in S} \mathrm{p}_s \left| f_s(x_s, PT) - \bar{f}(x_s, PT) \right| \tag{11}$$

- *Conditional Value-at-Risk (CVaR)*: It is also called expected shortfall and is a quantile-based risk measure similarly to Value-at-Risk (VaR), which is the quantile of a distribution for a given probability level α. The minimization of CVaR was first proposed by Rockafellar & Uryasev (2000).

$$\mathrm{Risk}_{\mathrm{CVaR}_\alpha}(x_s, PT) = \gamma - \frac{1}{1-\alpha} \sum_{s \in S} \mathrm{p}_s \left[-f_s(x_s, PT) - \gamma \right]_+ \tag{12}$$

where $\gamma \in \mathbb{R}$ (additional variable).

5 Numerical Examples

The proposed approach to deal with production variability in the operations planning stage of the Sales & Operations Planning (S&OP) process is illustrated with a motivating example and an industrial case study. The two-stage scenario tree for the stochastic models has its node values (outcomes) fixed to the sampled Δ values from the quantile regression analysis, and the probabilities of the scenarios, p_s, were calculated using the data-driven scenario generation approach described in Calfa *et al.* (2014) (see L^2 DMP formulation).

All optimization models were implemented in AIMMS 3.13 (Roelofs & Bisschop, 2013) and solved on a desktop computer with the following specifications: Dell Optiplex 990 with 4 Intel® Core™ i7-2600 CPUs at 3.40 GHz (total 8 threads), 8 GB of RAM, and running Windows 7 Enterprise. All linear programming (LP) and convex quadratically-constrained programming (QCP) models were solved with Gurobi 5.6.

The Sim-Opt approach consists of a main script in MATLAB (The MathWorks Inc., 2014) in which the DFO algorithm `fminsearchbnd`[1] (Nelder-Mead simplex search algorithm) sets the production targets PT that are fixed in the AIMMS model (the simulator). In other words, the DFO algorithm proposes a production plan, which is evaluated by the simulator (two-stage stochastic programming model). In order to perform a comparison between Sim-Opt and Bi-Opot approaches, we used the same Δ values in both, even though the Sim-Opt can accommodate the sampling of Δ values conditional on the proposed production plan.

5.1 Motivating Example

The goal of the motivating example is to demonstrate that different allocation schemes of chemicals in a process network are obtained when production variability is considered and some risk measure is adopted. Typically, only the margin (sales from revenue minus operating cost) of individual products is used as a criterion for deciding their allocation throughout the network. By also accounting for production variability, larger amounts of a feedstock chemical may be allocated to lower-margin, but potentially more reliable plants (to be defined in the next paragraph) than to higher-margin, but less reliable plants that "compete" for the same raw material. Even though this may seem counter-intuitive at first, we show through this example the trade-off between the expected or average *overall* profit and the risk of choosing an allocation scheme, i.e., less risk with allocation favoring low-margin and more reliable plants vs. high risk with allocation favoring high-margin and less reliable plants. Please recall that by reliability we mean the spread of Δ values around zero, i.e., the deviation of actual production rates from the respective planned values. A detailed analysis of the results is given to illustrate the applicability of the proposed methodology.

The process network is shown in Figure 6. Each plant produces a single product, which receives the same name as the plant that produces it. Therefore, we will use plant and product interchangeably. The main objective is to demonstrate the different allocation schemes of chemical A to the downstream plants (B–G) between deterministic and stochastic (risk neutral and averse) solutions by considering production variability of the downstream plants. Note that the order of plant's *reliability* is the reverse of the order of plant's margin (revenue from sales minus operating costs), i.e., the most reliable plant (G) is the lowest-margin plant, whereas the less reliable plant (B) is the highest-margin one. In this context, reliability is represented by the spread of the deviation between historical planned and actual production rates (Δ values) around the origin, which is along the lines of a root mean square calculation. In other words, a more reliable plant means $\Delta \approx 0$, i.e., it is more likely to actually achieve its planned values proposed in the operations planning stage of the S&OP process.

[1]The function `fminsearchbnd` extends MATLAB's built-in function `fminsearch` by considering bounds on the decision variables. The bounds used in the computational experiments were the plant minimum/maximum capacities. See http://www.mathworks.com/matlabcentral/fileexchange/8277-fminsearchbnd--fminsearchcon for implementation (retrieved on March 22, 2015).

Please note that, in this paper, we do not refer to reliability in the sense of maintainability or probability of failures of a system or component.

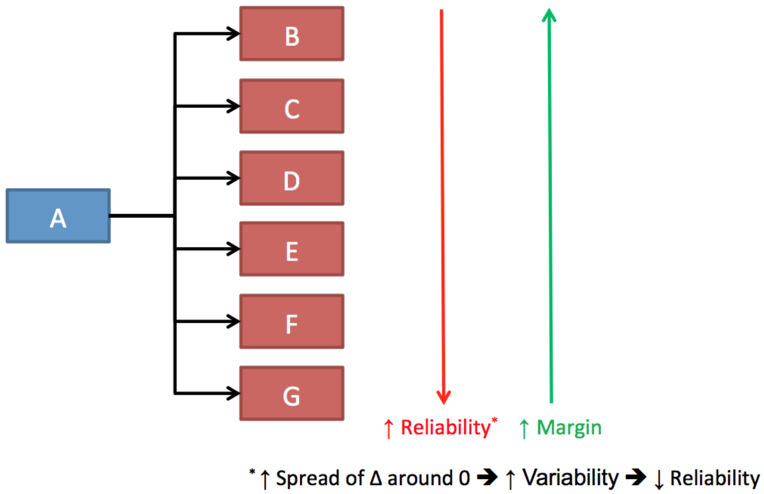

Figure 6: Process network structure of the motivating example. Plant reliability is related to the spread of the deviation between historical planned and actual production rates around the origin. Specifically in this example, the more reliable a plant is, the lower its margin is.

The multi-period, LP production planning model is similar to the one presented in Sahinidis *et al.* (1989), but excluding capacity expansion decisions. That is, the model consists of simple input-output relationships of material and inventory balances. Possible demand and plant capacity violations are captured with non-negative slack variables added to the respective constraints and penalized in the objective function. When production variability is taken into account, the deterministic equivalent model of the two-stage stochastic production planning model can be generically written as in equation (4), where the proposed S&OP production plan or target, PT, is a vector of first-stage variables. The model also has slack variables that capture unsatisfied demand and capacity violations that are penalized in the objective function. Twenty scenarios (samples from the quantile regression analysis) were considered in all stochastic models. We perform a detailed analysis of the results of four cases defined as follows:

Case 1. Deterministic

- No production variability, i.e., $\Delta_s = 0$.
- Identifier: 1. Det

Case 2. Risk Neutral Stochastic with Fixed Production Target

- Production variability is considered, i.e., $\Delta_s \neq 0$.
- *Fixed* the values of production targets, PT, to the optimal production rates obtained by solving Case 1.
- The purpose of this case is to evaluate the performance of the deterministic production plan in an uncertain environment.

– Identifier: 2. Stoch Fix

Case 3. Risk Neutral Stochastic with Variable Production Target

– Similar to Case 2, but with *variable* production targets, PT.

– Production targets (first-stage decisions) are optimally set while taking into account the historical variability in production rates.

– Identifier: 3. Stoch Var

Case 4. Risk Averse Stochastic (Bi-Opt Framework)

– Similar to Case 3, but with two objective functions, one of them measuring risk.

– Identifier: 4. Bi-Opt

We should note that Cases 1–3 give rise to LP models since the underlying planning model is linear. Case 4 gives rise to a convex QCP model, because we use the variance of the profit as the risk measure. The QCP model is convex because the variance is taken with respect to the profit, which is a linear function, and the scenario probabilities are non-negative; thus, the $\text{Risk}(x_s, PT)$ function is a sum of non-negative quadratic terms.

We first focus on the results obtained with the Bi-Opt approach (Subsection 4.2), and then comment on its differences with the Sim-Opt approach (Subsection 4.1). We begin with Cases 1–3, and then discuss Case 4, which has Case 3 as a special case.

5.1.1 Cases 1, 2, and 3

We start the analysis by comparing average overall margin and its standard deviation for the first three cases. The overall profit (or simply profit) is the difference between the revenue from the sale of all products and total costs (operating and inventory). Operating costs are proportional to production rates, and inventory costs are proportional to the amount of chemicals stored in each period. As it can be seen in Table 2, adjusting the production target (Case 3) in the face of production variability yields more profitable and less risky (smaller standard deviation) production plans. In other words, production targets obtained with the deterministic model (Case 1) yield lower average profit with larger spread (higher risk) when production variability is taken into account (Case 2). In addition, some plant capacity violations were observed in the solution of Case 2, which is clearly undesirable. The optimization results convey that the value of simultaneously proposing the production targets and accounting for production variability is $(320.60 - 272.84)\,\text{m.u.} = 47.77\,\text{m.u.}$ (monetary units) on average.

Table 2: Average profit and its standard deviation in monetary units (m.u.) for Cases 1–3 in the motivating example. The standard deviation of the profit is a measure of its spread, i.e., financial risk.

	Case		
	1. Det	2. Stoch Fix	3. Stoch Var
Average	375.15	272.84	320.60
Std Dev	–	28.30	25.82

The difference in average profitability between the solution of Case 2 and Case 3 is also explained by the average overall service level (SL) defined in equation (13). The overall SL is the complement of the fraction of total demand satisfied from sales of all products.

$$\mathbb{E}[\text{SL}] = 1 - \frac{\sum_{s \in S} \text{p}_s Sales_s}{\text{Total Demand}} \tag{13}$$

The overall SL for the deterministic solution (Case 1) is 100% (i.e., all demand is satisfied), and the average overall SL for Case 2 and Case 3 is 82.68% and 92.24%, respectively. The breakdown of unmet demand for each product is given in Table 3. From Case 2 to Case 3, the demand satisfaction of high-margin products (B, C, and D) increases relatively more than for the low-margin products (E, F, and G). In fact, the demand satisfaction of product G, which is produced by the most reliable and lowest-margin plant, actually decreased in Case 3. When analyzing the results of Case 4 later on, it will be clear that the solution of Case 3 favors more allocation of product A to the high-margin plants, since the objective function (profit) is not constrained by any risk measure (i.e., more risky condition).

Table 3: Average unmet demand in weight units (w.u.) for each product in Case 2 and Case 3 in the motivating example.

Product	**Case**	
	2. Stoch Fix	3. Stoch Var
B	64	6
C	55	10
D	41	4
E	34	10
F	18	10
G	8	34

The amounts of chemical A allocated to the six downstream plants are shown in Table 4, which complements the results shown in the previous table. As it would be expected after the analysis of unmet demand, in Case 3, relatively more amounts of A are allocated to high-margin plants than for low-margin ones. The relatively lower amount of product A allocated to plant C when compared to other plants is due to the overall lower demand for product C across time periods. The negative percentage change for plant G (lowest-margin) means that it receives less A in Case 3 than in Case 2.

Table 4: Average allocated amounts of A in weight units (w.u.) to each downstream plant in Case 2 and Case 3 in the motivating example. The percentage change column is the relative change between the two cases, i.e., (Case 3 – Case 2)/Case 3.

Plant	**Case**		
	2. Stoch Fix	3. Stoch Var	Change
B	160	204	28%
C	96	143	49%
D	131	164	25%
E	147	178	21%
F	343	356	4%
G	148	111	−25%

Let us analyze the results for Case 4. The goal is to evaluate the impact of controlling some measure of risk by including an additional constraint (ϵ-constraint) on the allocation scheme of chemical A to downstream plants. Consider two subcases of Case 4: Subcase 4.A uses an explicit financial risk measure (variance of profit) and Subcase 4.B uses the individual expected service level for one of the high-margin products.

5.1.2 Subcase 4.A: Variance of Profit

The ϵ-constraint in equation (4) takes the following form,

$$\text{Risk}(x_s, PT) = \sum_{s \in S} \text{p}_s \left[\text{Profit}_s(x_s, PT) - \overline{\text{Profit}}(x_s, PT)\right]^2 \leq \epsilon \tag{14}$$

where $\text{Profit}_s(\cdot,\cdot)$ denotes only the profit calculation of the objective function, i.e., excluding penalized slack variables, and $\overline{\text{Profit}}(x_s, PT) = \sum_{s \in S} \text{p}_s \text{Profit}_s(x_s, PT)$ is the average profit. In this subcase, ϵ is interpreted as the maximum allowed variance of profit and has units of $(\text{m.u.})^2$, where "m.u." stands for monetary units.

Note that Case 3 is a special case of Subcase 4.A in which ϵ takes a large enough value so that the constraint is not active at the solution, i.e., the financial risk is unconstrained and the model becomes risk neutral. Thus, the solution to Case 3 represents the condition of maximum variance of the profit, which is the right-most point in the Pareto efficient curve of average profit vs. variance (or standard deviation) of profit in Figure 7. In addition to the solution of Case 3, the bi-objective optimization model (convex QCP) was solved ten times for different values of ϵ, ranging from 60 to 600 with a stride of 60, and the solutions are represented by points on the Pareto efficient frontier. Note that from left to right the spread of the profit across scenarios increase, i.e., more risky solutions.

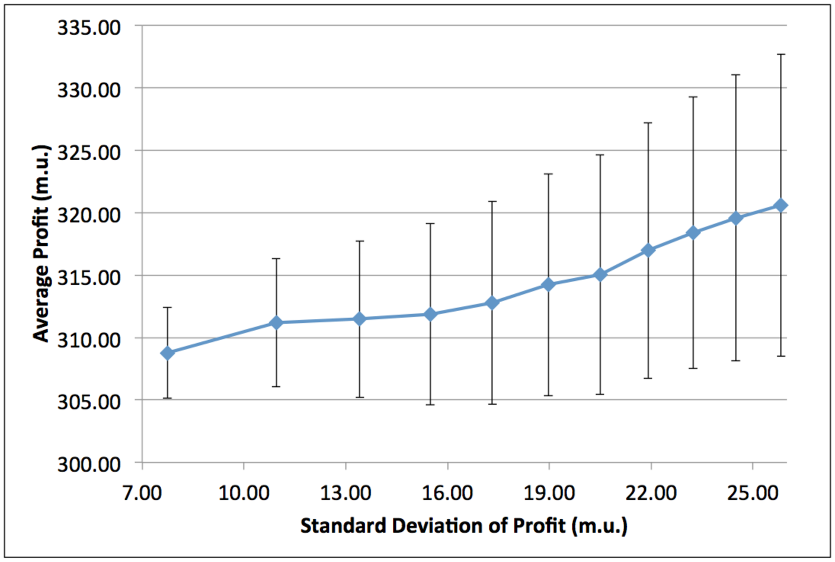

Figure 7: Pareto efficient frontier for Subcase 4.A in the motivating example. "m.u." stands for monetary units. Error bars represent 95% confidence intervals on the average values.

Figure 8 shows the same Pareto efficient frontier together with the average overall service level (SL) as defined in equation (13). Note that there is an increase in the average overall SL for the solutions from points P1 to P2, and after point P2 until point P3 the average overall SL levels off. Therefore, we classify the solutions as belonging to two regions: Region I (less risky) and Region II (more risky). After point P2, the amount of A allocated to all downstream plants remains practically the same with the exception of the least and most reliable plants, B and G, respectively. As the ϵ value increases (more risky condition), there is a shift of the amount of A allocated from G to B.

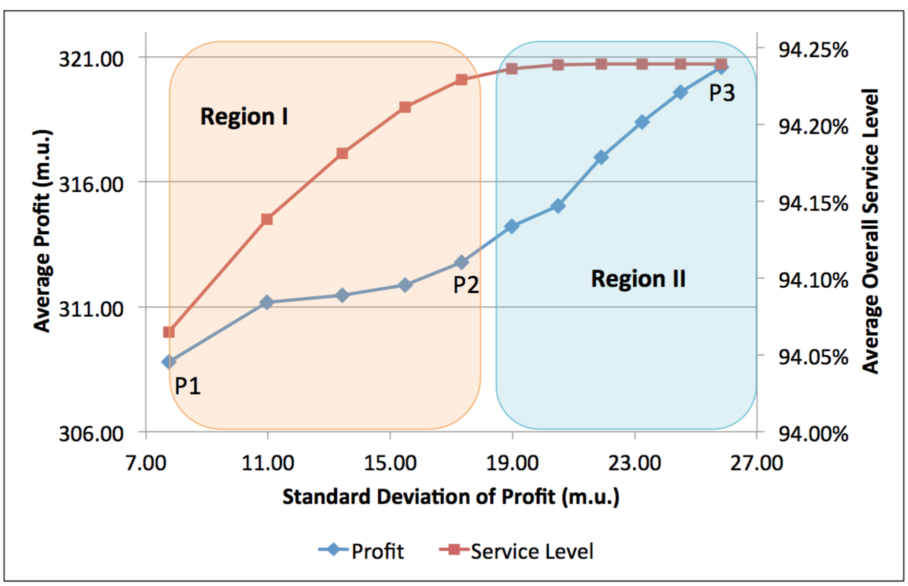

Figure 8: Pareto efficient frontier and average overall service level for Subcase 4.A in the motivating example. "m.u." stands for monetary units.

Figure 9 shows the allocation amounts of A to the downstream plants for each point in the Pareto efficient curve for the two plants in the extremes of the reliability-margin scale. The same overall trend is observed for the other plants: more A is allocated to less reliable, high-margin plants (B, C, and D) as risk (ϵ value) increases; conversely, less A is allocated to more reliable, low-margin plants (E, F, and G) from left to right in the figure.

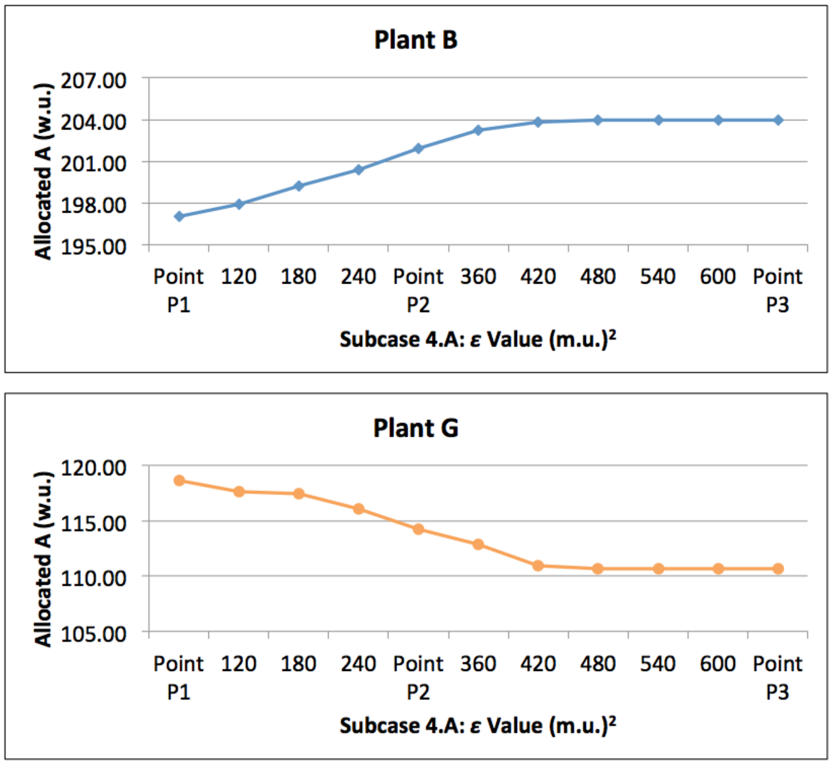

Figure 9: Allocation scheme for two downstream plants in Subcase 4.A in the motivating example. Plant B is the less reliable and highest-margin, whereas plant G is the most reliable and lowest-margin of the downstream plants. The ϵ value denotes the variance of the profit (financial risk). "w.u." and "m.u." stand for weight and monetary units, respectively.

5.1.3 Subcase 4.B: Average Service Level of Product C

The ϵ-constraint in equation (4) takes the following form,

$$\text{Risk}(x_s, PT) = 1 - \frac{\sum_{s \in S} \text{p}_s Sales_s^{\text{C}}}{\text{Total Demand}} \leq \epsilon \tag{15}$$

where $Sales_s^{\text{C}}$ indicates that only sales for product C are considered, thus the ϵ-constraint is a calculation of the complement of the average *individual* service level (SL) of product C. In this subcase, ϵ is interpreted as the maximum allowed fraction of unmet demand of product C and is dimensionless. It can be expressed as a percentage, e.g., $\epsilon = 1\%$ means that at most 1% of the demand of product C can be unmet, or equivalently, at least 99% of the demand of C must be satisfied.

The motivation behind this subcase is two-fold: (1) it uses a non-conventional form of the ϵ-constraint for the risk (i.e., not an explicit financial risk measure); (2) the average

SL of the high-margin product C in the solution of Case 3 is 93.85%, and it is desired to evaluate the impacts of enforcing a higher average SL of this valuable product on the overall profitability of the production plan.

Figure 10 shows the Pareto efficient curve of the average profit vs. average individual service level of product C. In addition to the solution of Case 3, the bi-objective optimization model (LP) was solved four times for different values of ϵ (5%, 4%, 3%, and 2%). In this motivating example, 99% (i.e., $\epsilon = 1\%$) or higher average SL of product C is not feasible. Note that as the average SL of C increases, the overall average profit decreases and its spread (represented by the error bars) increases. Even though more demand of product C is met from left to right, the overall production plan becomes less profitable on average.

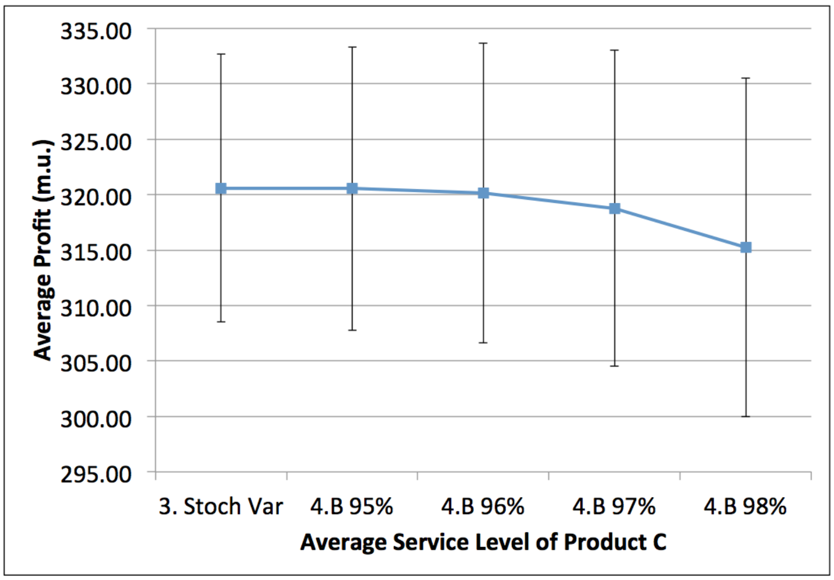

Figure 10: Pareto efficient frontier for Subcase 4.B in the motivating example. Error bars represent 95% confidence intervals on the average values. "m.u." stands for monetary units.

Figure 11 helps explain the result in the previous figure. The average overall SL decreases as the average individual SL of product C is forced to increase. In other words, by requiring higher SL for product C, there is a shift of A allocated from other plants to plant C in order to ensure its desired SL. Consequently, the individual demand satisfaction of the other products decreases from left to right in the figure.

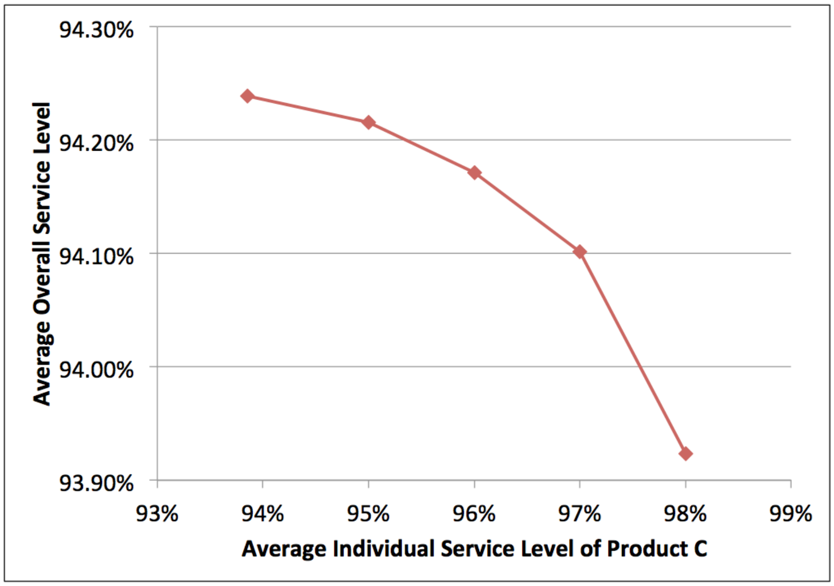

Figure 11: Effect of increasing average service level of product C on average overall service level in Subcase 4.B in the motivating example.

5.1.4 Computational Statistics

To conclude the analysis of the motivating example, Table 5 presents the computational statistics of the optimization models solved in the four cases. The time corresponds to the wall time, including loading and solving the models. The time listed under Subcases 4.A and 4.B corresponds to the average wall time for all points in the Pareto efficient frontier in each subcase. Recall that Cases 1 through 3 are LP models, Subcase 4.A contains convex QCP models, and Subcase 4.B has LP models.

Table 5: Computational statistics for optimization models in all cases and subcases in the motivating example.

	Case and Subcase				
	1. Det	2. Stoch Fix	3. Stoch Var	4.A Bi-Opt	4.B Bi-Opt
Variables	5,186	101,778	101,874	101,874	101,874
Constraints	3,602	73,482	73,482	73,483	73,483
Time [s]	2.59	4.38	5.05	5.31	5.12

5.1.5 Sim-Opt vs. Bi-Opt

We note that Cases 1 and 2 are the same for both Sim-Opt and Bi-Opt frameworks, since they correspond to the deterministic problem (no production variability) and the stochastic problem with fixed production targets (no proposed production plan), respectively. The Sim-Opt framework for Case 3 took 1,000 iterations (imposed limit), 1,160 function evaluations (i.e., calls to the simulator, which is the two-stage stochastic programming model implemented in AIMMS), and 10,745 seconds to achieve an expected profit of 294.43 m.u., while the Bi-Opt framework yielded an expected profit of 320.60 m.u. in less than 5 seconds. In addition, the solution of the Sim-Opt framework exhibited 91.36% expected overall service level, which is lower than that obtained with the Bi-Opt framework (94.24%).

The number of decision variables in this case is 72, since they correspond to the monthly production targets of 6 plants for a time horizon of one year. Moreover, every call to the simulator takes approximately 7 seconds, while the time spent by the DFO algorithm per iteration is negligible. These two factors make the Sim-Opt framework computationally more intensive than the Bi-Opt approach, in spite of having more flexibility for sampling Δ values conditional on proposed production targets. Similar observations were made for the solution of Case 4 with the Sim-Opt approach.

5.2 Industrial Case Study

The industrial test case concerns the optimal production planning of chemical sites. Each site contains several plants that are highly integrated. The plants can also transfer products between sites. The chemical sites contain more than 12 production facilities and manufacture several products. The time horizon of one year is divided into monthly time periods. The objective of the optimization model is to maximize the total profit. Due to confidentially reasons, we only discuss the modeling changes from the motivating example and the computational results.

In Case 2 of the motivating example, the production targets, PT, are fixed to the corresponding production rates obtained from the solution of the deterministic model (Case 1). The goal is to evaluate the performance of the production targets proposed by the deterministic model when production variability is taken into account. In order to perform a similar study on a highly integrated system, some modeling modifications may be necessary. If the general deterministic equivalent form of the two-stage stochastic programming model (see equation (4)) describes a highly integrated system, the equality constraint that captures the production variability,

$$PR_s = PT - \Delta_s \qquad \forall\, s \in S \tag{16}$$

for *fixed* PT, may cause infeasibility. For instance, it may not be feasible to satisfy material and inventory balances in the network by imposing such a constraint on production rates with fixed production targets. To circumvent this potential problem, we propose two modifications to the implementation of Case 2 for a highly integrated system: (1) unfix the production target decision variables, PT, and (2) penalize the deviation of the production targets set by stochastic model from the corresponding values obtained in the solution of the deterministic model, $\mathrm{PT}^{\mathrm{det}}$ (a constant parameter vector). Therefore, the general optimization problem for Case 2 is rewritten as follows:

$$\begin{aligned}
&\min_{x_s, PT} && \mathrm{Risk}(x_s, PT) \\
&\max_{x_s, PT} && f_0(PT) + \sum_{s \in S} \mathrm{p}_s f_s(x_s) - \psi \cdot ||PT - \mathrm{PT}^{\mathrm{det}}||_p \\
&\text{s.t.} && g(x_s) \le 0 && \forall\, s \in S \\
& && PR_s = PT - \Delta_s && \forall\, s \in S
\end{aligned} \tag{17}$$

where ψ is a penalty factor and $||\cdot||_p$ is an L^p-norm. In order to preserve the linearity of the production planning model used in this work, we employed the L^1-norm, which resulted in the following formulation:

$$\begin{aligned} \min_{x_s, PT} \quad & \text{Risk}(x_s, PT) \\ \max_{x_s, PT} \quad & f_0(PT) + \sum_{s \in S} \text{p}_s f_s(x_s) - \psi \cdot (PT^+ + PT^-) \\ \text{s.t.} \quad & g(x_s) \le 0 && \forall\, s \in S \\ & PR_s = PT - \Delta_s && \forall\, s \in S \\ & PT - \text{PT}^{\text{det}} = PT^+ - PT^- \\ & PT^+, PT^- \ge 0 \end{aligned} \tag{18}$$

where PT^+ and PT^- capture the positive and negative deviations.

The high degree of integration in the network also has to be considered in the statistical modeling. If two plants are directly connected (e.g., plant A feeds plant B), then it may not be realistic (and likely lead to infeasibilities) to consider them independent from the point of view of production rates. In other words, deviations from plan in an upstream plant may affect production in a downstream plant, and vice versa.

Different approaches can be used to account for this dependence between plants for which production variability is taken into account. If only Δ values are used to characterize production variability (i.e., the covariate Plan values are disregarded), then we propose the following approaches (illustrated in Figure 12) for any two connected plants A and B:

- Assume a *parametric* classical regression model for the Δ values between the upstream and downstream plants, and from the generated samples of one of the plants, calculate the corresponding (expected) sample of the other plant. For example, if a linear regression model such as $\Delta^{\text{Plant B}} = \beta_0 + \beta_1 \cdot \Delta^{\text{Plant A}}$ is considered, then (1) estimate the regression function (in this case, the model parameters β_0 and β_1) by regressing Δ values of Plant B on Δ values of Plant A, then (2) generate samples of Δ values for Plant A, and finally (3) calculate the corresponding Δ values for Plant B using the linear model.

- Similar to the previous approach, but instead assume a *nonparametric* classical regression model for the Δ values between the upstream and downstream plants. In other words, the regression model is generally written as $\Delta^{\text{Plant B}} = g(\Delta^{\text{Plant A}})$, where $g(\cdot)$ is the nonparametric regression function (e.g., kernel regression (Li & Racine, 2007)). Follow the same three-step procedure discussed in the previous approach.

- Estimate the joint distribution of Δ values for both plants A and B, i.e., obtain $\hat{F}(\Delta^{\text{Plant A}}, \Delta^{\text{Plant B}})$, and then sample from this estimated multivariate distribution.

- Perform bootstrap resampling (Efron, 1979), i.e., sample with replacement data points from the original data set. As in the previous approach, the samples constitute the outcomes (i.e., scenarios) in the stochastic programming framework.

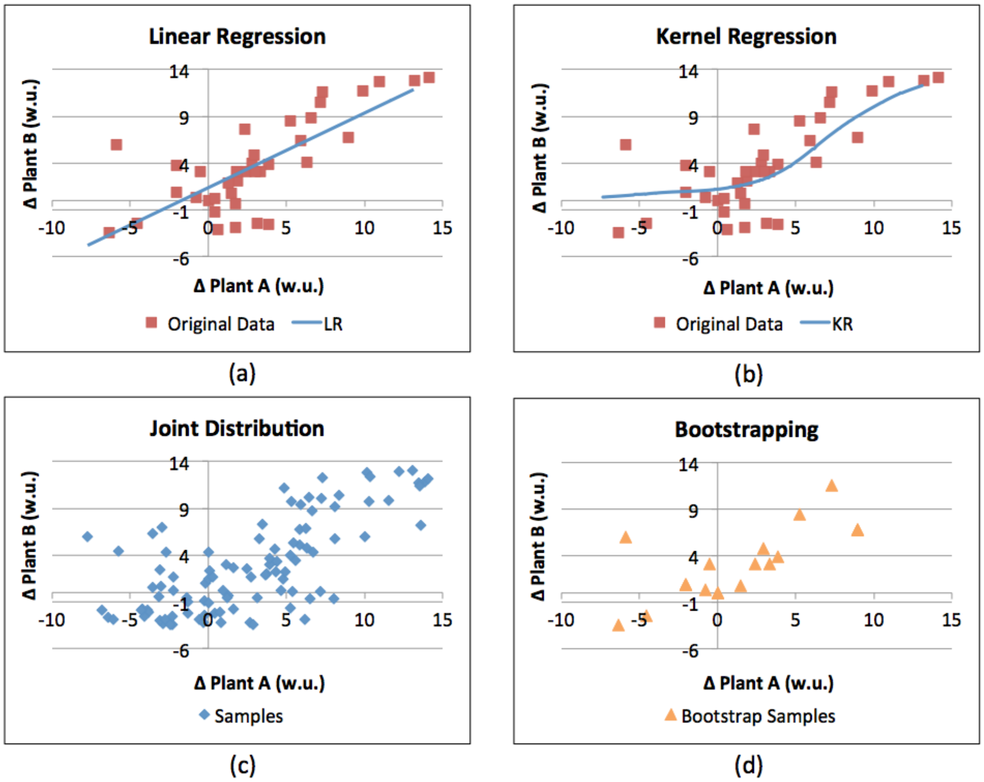

Figure 12: Illustration of the three proposed approaches to account for dependence of the Δ values for directly connected plants in the network. The relationships between the Δ values of the upstream and downstream plants are captured by: (a) linear regression (LR), (b) kernel regression (KR) using a second-order Gaussian kernel, (c) estimated joint distribution, and (d) bootstrap resampling. "w.u." stands for weight units.

When the Plan values are taken into account in modeling production variability (i.e., the quantile regression approach discussed in Section 3), then the problem essentially becomes modeling the relationship between *conditional* distributions, $\Delta^{\text{Plant B}}|\text{Plan}^{\text{Plant B}}$ given $\Delta^{\text{Plant A}}|\text{Plan}^{\text{Plant A}}$. A general approach would be to estimate the joint conditional distribution of all random variables involved, $\hat{F}(\Delta^{\text{Plant A}}, \Delta^{\text{Plant B}}|\text{Plan}^{\text{Plant A}}, \text{Plan}^{\text{Plant B}})$, and then sample from this estimated multivariate distribution. This general approach may pose computational and algorithmic challenges. An *approximate* approach to generate samples of $\Delta^{\text{Plant B}}|\text{Plan}^{\text{Plant B}}$ given $\Delta^{\text{Plant A}}|\text{Plan}^{\text{Plant A}}$ is described in the following steps:

1. Obtain the estimated distribution of $\Delta^{\text{Plant A}}|\text{Plan}^{\text{Plant A}}$, $\hat{F}_A(\Delta^{\text{Plant A}}|\text{Plan}^{\text{Plant A}})$, and sample from this conditional distribution. Let the samples be denoted by S_A.

2. Repeat the previous step replacing "A" with "B".

3. Obtain the estimated distribution $\hat{F}_S(S_B|S_A)$, and sample from this conditional distribution. These final "approximate" samples relate $\Delta^{\text{Plant B}}|\text{Plan}^{\text{Plant B}}$ given $\Delta^{\text{Plant A}}|\text{Plan}^{\text{Plant A}}$.

Table 6 presents the computational statistics of the optimization models solved in the four cases. We report both solution times (CPU times for the solver) and wall times (includes loading and solving the models), which become significant as the problem instance increases. Similar to the Motivating Example, Cases 1–3 and Subcase 4.B are LP models, while Subcase 4.A contains a convex QCP model.

Table 6: Computational statistics for optimization models in all cases and subcases in the industrial case study.

	Case and Subcase				
	1. Det	2. Stoch Fix	3. Stoch Var	4.A Bi-Opt	4.B Bi-Opt
Variables	124,116	2,447,057	2,446,913	2,446,914	2,446,913
Constraints	88,790	1,777,313	1,777,241	1,777,243	1,777,242
Solution Time [s]	0.45	9.17	10.22	15.61	17.43
Wall Time [s]	15.26	50.17	53.77	60.20	62.76

The Sim-Opt framework for Case 3 was run for 100 iterations (imposed limit), 181 function evaluations (i.e., calls to the two-stage stochastic programming model implemented in AIMMS), and 10,130 seconds to achieve an expected profit of $1,293.45\,\mathrm{m.u.}$, while the Bi-Opt framework yielded an expected profit of $1,299.81\,\mathrm{m.u.}$ in less than 54 seconds of wall time. Moreover, the solution of the Sim-Opt framework yielded 97.92% expected overall service level, which is lower than that obtained with the Bi-Opt framework (98.60%). Similar to the motivating example, the number of decision variables for the DFO algorithm is 72 (monthly production targets for 6 plants and time horizon of one year). Each call to the AIMMS model takes approximately 80 seconds, which contributes significantly to the overall solution time of the Sim-Opt framework. Similar observations were made for the solution of Case 4.

6 Conclusions

In this paper, we have addressed uncertainty in the operations planning stage of the Sales & Operations Planning (S&OP) process of a manufacturing company. The uncertainty is attributed to production variability, which is caused by unplanned events that can result in actual production rates lower or higher than their planned values. In order to model production variability, we defined Δ as the deviation between historical planned and actual production rates, and used quantile regression to predict quantiles of the distribution of Δ conditional on planned production rates. The predicted quantiles form samples or scenarios in a two-stage bi-objective stochastic programming production planning model, whose first-stage variables are the production plans or targets that are sought to be implemented in practice. One objective represents the financial performance of the production plan (e.g., profit), whereas the other is a risk measure (explicitly financial or not).

The applicability of the proposed approach was illustrated with a motivating example and an industrial case study. The motivating example consisted of a small process network with a feedstock plant that serves six plants. The downstream plants are sorted in reverse order of reliability and margin, i.e., the most reliable plant is also the lowest-margin one, and the less reliable plant is the highest-margin one. The motivation behind this was to show that depending on the maximum desired level of risk, the optimization model decides to allocate more of the feedstock chemical to more or less reliable plants. In other words, the optimization model considers not only the individual margin of the plants, but also their reliability in order to obtain a solution with maximum *expected* profit. For the industrial case study, we proposed modeling approaches to address the connectivity in the network, which may create dependence in production variability profiles.

Acknowledgment

The authors gratefully acknowledge financial support from The Dow Chemical Company.

References

Birge, J. R., and Louveaux, F. 2011. *Introduction to Stochastic Programming.* Springer Science+Business Media, LLC., second edition. New York, NY. USA.

Cai, X.; Teo, K.-L.; Yang, X.; and Zhou, X. Y. 2000. Portfolio Optimization Under a Minimax Rule. *Management Science.* 46(7):957–972.

Calfa, B. A.; Agarwal, A.; Grossmann, I. E.; and Wassick, J. M. 2014. Data-Driven Multi-Stage Scenario Tree Generation via Statistical Property and Distribution Matching. *Computers & Chemical Engineering.* 68(1):7–23.

Conn, A. R.; Scheinberg, K.; and Vicente, L. N. 2009. *Introduction to Derivative-Free Optimization.* Society for Industrial and Applied Mathematics (SIAM) and the Mathematical Programming Society (MPS). Philadelphia, PA. USA.

Custódio, A.; Emmerich, M.; and Madeira, J. F. A. 2012. Recent Developments in Derivative-Free Multiobjective Optimisation. *Computational Technology Reviews.* 5(1):1–30.

Efron, B. 1979. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics* 7(1):1–26.

Grimson, J. A., and Pyke, D. F. 2007. Sales and Operations Planning: An Exploratory Study and Framework. *The International Journal of Logistics Management.* 18(3):322–346.

Hwang, C.-L., and Masud, A. S. M. 1979. *Multiple Objective Decision Making, Methods and Applications: A State-of-the-Art Survey*, volume 164 of *Lecture Notes in Economics and Mathematical Systems.* Springer-Verlag. New York, NY. USA.

Koenker, R. 2005. *Quantile Regression.* Cambridge University Press. New York, NY. USA.

Li, Q., and Racine, J. S. 2007. *Nonparametric Econometrics: Theory and Practice.* Themes in Modern Econometrics. Princeton University Press. New Jersey, NJ. USA.

Li, J.; González, M.; and Zhu, Y. 2009. A Hybrid Simulation Optimization Method for Production Planning of Dedicated Remanufacturing. *International Journal of Production Economics.* 117(2):286–301.

Lim, L. L.; Alpan, G.; and Penz, B. 2014. A Simulation-Optimization Approach for Managing the Sales and Operations Planning in the Automotive Industry. Technical Report 212, Laboratoire G-SCOP. ISSN: 1298-020X.

Ling, R. C., and Goddard, W. E. 1988. *Orchestrating Success: Improve Control of the Business with Sales & Operations Planning.* John Willey & Sons, Inc. New York, NY. USA.

Miettinen, K. 1999. *Nonlinear Multiobjective Optimization*, volume 12 of *International Series in Operations Research & Management Science.* Kluwer Academic Publishers. Boston, MA. USA.

Montgomery, D. C., and Runger, G. C. 2003. *Applied Statistics and Probability for Engineers.* John Wiley & Sons, Inc., third edition. New York, NY. USA.

Mula, J.; Poler, R.; García-Sabater, J. P.; and Lario, F. C. 2006. Models for Production Planning under Uncertainty: A Review. *The International Journal of Production Economics.* 103(1):271–285.

Paulk, M. C.; Weber, C. V.; and Chrissis, M. B. 1993. *The Capability Maturity Model for Software, Version 1.1.* Software Engineering Institute. Carnegie Mellon University. Pittsburgh, PA. USA.

Pinedo, M. L. 2009. *Planning and Scheduling in Manufacturing and Services.* Springer Science+Business Media, LLC., second edition. New York, NY. USA.

Rios, L. M., and Sahinidis, N. V. 2013. Derivative-Free Optimization: A Review of Algorithms and Comparison of Software Implementations. *Journal of Global Optimization.* 56(3):1247–1293.

Rockafellar, R. T., and Uryasev, S. 2000. Optimization of Conditional Value-at-Risk. *Journal of Risk.* 2:21–41.

Roelofs, M., and Bisschop, J. 2013. *Advanced Interactive Multidimensional Modeling System (AIMMS).* http://www.aimms.com/.

Sahinidis, N. V.; Grossmann, I. E.; Fornari, R. E.; and Chathrathi, M. 1989. Optimization Model for Long Range Planning in the Chemical Industry. *Computers & Chemical Engineering.* 13(9):1049–1063.

Sarykalin, S.; Serraino, G.; and Uryasev, S. 2008. Value-at-Risk vs. Conditional Value-at-Risk in Risk Management and Optimization. In Chen, Z.-L., and Raghavan, S., eds., *Tutorials in Operations Research.* INFORMS. 270–294.

Sodhi, M. S., and Tang, C. S. 2009. Modeling Supply-Chain Planning under Demand Uncertainty Using Stochastic Programming: A Survey Motivated by Asset-Liability Management. *International Journal of Production Economics.* 121(2):728–738.

Sung, C., and Maravelias, C. T. 2007. Production Planning in Process Systems Engineering. In Pistikopoulos, E. N.; Georgiadis, M. C.; Dua, V.; and Papageorgiou, L. G., eds., *Process Systems Engineering: Supply Chain Optimization*, volume 4. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA. chapter 9.

The MathWorks Inc. 2014. *MATLAB® R2014a*. The MathWorks Inc., Natick, Massachusetts, United States. http://www.mathworks.com/.

Thomé, A. M. T.; Scavarda, L. F.; Fernandez, N. S.; and Scavarda, A. J. 2012. Sales and Operations Planning: A Research Synthesis. *International Journal of Production Economics.* 138(1):1–13.

Verderame, P. M.; Elia, J. A.; Li, J.; and Floudas, C. A. 2010. Planning and Scheduling under Uncertainty: A Review Across Multiple Sectors. *Industrial & Engineering Chemistry Research.* 49(9):3993–4017.

Wassick, J. M. 2009. Enterprise-wide Optimization in an Integrated Chemical Complex. *Computers & Chemical Engineering.* 33(12):1950–1963.

Zhou, A.; Qu, B.-Y.; Li, H.; Zhao, S.-H.; Suganthan, P. N.; and Zhang, Q. 2011. Multiobjective Evolutionary Algorithms: A Survey of the State of the Art. *Swarm and Evolutionary Computation.* 1(1):32–49.