Proceedings of the 2017 Winter Simulation Conference
W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, eds.

LOAD DEPENDENT LEAD TIME MODELLING: A ROBUST OPTIMIZATION APPROACH

Erinç Albey

Department of Industrial Engineering
Özyeğin University
İstanbul, 34794, TURKEY

İhsan Yanıkoğlu

Department of Industrial Engineering
Özyeğin University
İstanbul, 34794, TURKEY

Reha Uzsoy

Edward P. Fitts Department of Industrial and Systems Engineering
North Carolina State University
Raleigh, NC 27695-7906, USA

ABSTRACT

Although production planning models using nonlinear CFs have shown promising results for semiconductor wafer fabrication facilities, the lack of an effective methodology for estimating the CFs is a significant obstacle to their implementation. Current practice focuses on developing point estimates using least-squares regression approaches. This paper compares the performance of a production planning model using a multi-dimensional CF and its robust counterpart under several experimental settings. As expected, as the level of uncertainty is increased, the resulting production plan deviates from the optimal solution of the deterministic model. On the other hand, production plans found using the robust counterpart are less vulnerable to parameter estimation errors.

1 INTRODUCTION

The production release planning problem (PRP) seeks to match the output of a production system to its demand in an optimal or near-optimal way by determining the quantity of each product released into the system in any time period. This requires PRP models to anticipate the impact of their release decisions on the output and costs of the production system, which is difficult because the production system operates in continuous time while planning is carried out periodically. Queueing models , simulation models and industrial observation all show that the throughput of a production system is a complex, nonlinear function of the workload state of the system, including product mix (Karmarkar 1989). The workload in any planning period, in turn, is determined by the release plan developed by the production planning system. This mutual dependency between throughput and the planned release quantities is referred to as the *planning circularity* (Asmundsson, Rardin, Turkseven, and Uzsoy 2009).

The planning circularity is of particular interest to semiconductor manufacturers since the high capital cost of manufacturing facilities requires that they run at high utilization levels to be economically viable. At high utilization, even small fluctuations in utilization can cause significant variation in cycle time, adversely affecting on-time delivery and requiring higher safety stocks to ensure high customer service. The cyclic nature of demand for many semiconductor devices requires that these firms either maintain high output volumes by aggressively soliciting additional business

978-1-5386-3428-8/17/$31.00 ©2017 IEEE

at less rewarding prices, or reduce utilization to meet demand and try to compete on reduced cycle times. A production planning system that can accurately account for the cycle time impacts of resource utilization would provide a significant advantage in the latter case, although not necessarily in the former. Foundries, on the other hand, must maintain high on-time delivery performance with a diverse and ever-changing product mix, which again suggests that a planning model capable of accounting for queueing behavior may yield meaningful advantage. Another important issue is the effective planning of safety stocks throughout the semiconductor supply chain. Decades of research on inventory theory (Axsäter 2010, Zipkin 2000) has shown that the amount of safety stock needed to maintain a specified level of customer service, however defined, is determined to a large extent by the probability distribution of the lead time demand, the amount of demand arriving over the time required to replenish the inventory. The need to produce additional material to replenish safety stocks results in increased resource utilization which, in turn, increases the cycle time of the production facility and alters the lead time demand distribution. Hence the development of production planning models that comprehend the relation between resource utilization and cycle time holds out the possibility of an integrated approach to planning releases to maintain safety stocks in the face of uncertain supply and demand.

Two principal approaches have been proposed to address the planning circularity. The first of these decomposes the PRP problem into two subproblems: an optimization problem that determines the optimal releases for a given set of cycle time estimates, and an estimation problem that uses a simulation or queueing model to estimate the cycle times implied by a given release schedule. The planning procedure iterates between these two subproblems until convergence is achieved (Albey and Bilge 2011, Bang and Kim 2010, Hung and Leachman 1996). However, these procedures are time-consuming (Irdem, Kacar, and Uzsoy 2010) and their convergence behaviour is unclear (Kacar, Irdem, and Uzsoy 2012). The second class of approaches use nonlinear clearing functions (CFs), which formulate a mathematical relationship between the expected throughput of a machine in a planning period and its workload in that period (Albey and Uzsoy 2015, Albey, Bilge, and Uzsoy 2014, Asmundsson, Rardin, Turkseven, and Uzsoy 2009, Kang, Albey, Hwang, and Uzsoy 2014, Karmarkar 1989, Missbauer 2002, Srinivasan, Carey, Morton, et al. 1988). Most of these use a single aggregate state variable that does not distinguish different products. Extensive testing has shown that these models perform well under a wide range of conditions (Asmundsson, Rardin, Turkseven, and Uzsoy 2009, Kacar, Irdem, and Uzsoy 2012). However, when product mix affects the ability of the resource to produce output, as in the presence of setup times, this approach performs poorly, motivating the development of multi-dimensional CFs (MDCFs) for multi-product systems (Albey, Bilge, and Uzsoy 2014). (Albey, Bilge, and Uzsoy 2017) introduce several new product based MDCFs that show promising performance although their computational burden increases with instance size.

However, practical use of CFs has been hampered by the lack of effective methods for estimating them. While steady-state queueing models can be used to develop analytic expressions for CFs (Asmundsson, Rardin, Turkseven, and Uzsoy 2009, Karmarkar 1989, Missbauer 2002), the finite length of planning periods calls their validity into question. It has been shown that the functional form of the CF must take into account the transient behavior of the underlying queueing system, but this results in intractable optimization models (Missbauer 2009). The prevalent approach in the literature is to use least-squares regression to fit a postulated functional form to empirical data. However, the performance of this approach has been found to be lacking. This approach is subject to three types of errors (Stinstra and Den Hertog 2008, also see): 1) *simulation error*, due to discrepancies between reality and the computer model (e.g., distributional assumptions used in the simulation experiment), 2) *metamodelling error*, due to the difference between the "true" relation of the inputs (or factors) to the outputs (or response) and the relation imposed by the metamodel (e.g., estimating a CF using a linear function of workload when the actual relation is nonlinear), and,

3) *implementation error* which arises when the optimal solution for the computed model cannot be implemented with the same precision in a physical setting.

In this paper we postulate a functional form for the CF whose parameters are random variables with support in a specified uncertainty set, and incorporate it as a constraint in the PRP model, thus formulating the production planning problem as an optimization model with stochastic data. We adopt the robust optimization (RO) approach (Ben-Tal, El Ghaoui, and Nemirovski 2009, Ben-Tal and Nemirovski 2008, Bertsimas, Brown, and Caramanis 2011, Gorissen, Yanıkoğlu, and den Hertog 2015) which yields tractable optimization models that do not require specific distributional assumptions on the uncertain parameters. The model seeks release schedules that are immunized against the errors above, which will combine to result in errors in the estimated values of the CF parameters.

The following section reviews previous related work on optimization under uncertainty and the estimation of CFs. Section 3 briefly reviews the CF idea and presents the deterministic MDCF based PRP model. Section 4 reformulates the MDCF in a RO framework assuming an ellipsoidal uncertainty set assumption. Section 5 demonstrates how level of uncertainty affects the performance of robust counterpart, and tests how deterministic model and the robust counterpart perform when the underlying uncertainty set assumption is changed. Section 6 concludes the work and presents possible future research directions.

2 PREVIOUS RELATED WORK

Clearing functions (CFs) for production planning models were suggested by several authors in the late 1980s. One of the first forms is a linear CF under which a production resource converts a specified fraction of the work available to it at the start of a period into output during the period (Graves 1986). Concave CFs using a single state variable were introduced later (Karmarkar 1989, Srinivasan, Carey, Morton, et al. 1988). (Missbauer 2002) suggests a different concave saturating CF based on modeling the production resource as a steady-state $M/G/1$ queue. These functional forms are based on steady-state queueing models, with the exception of that of (Srinivasan, Carey, Morton, et al. 1988). (Missbauer 2009) uses transient analysis of the $M/M/1$ queue to show that the form of the CF of a transient queue is substantially different from that obtained in steady state. However, explicit representation of transient behavior results in nonlinear integer optimization models that are computationally challenging. (Kang, Albey, Hwang, and Uzsoy 2014) derive multiproduct CFs from steady-state queueing models to capture the impacts of lot sizing decisions, which yield promising results.

The most common approach to estimating CFs to postulate a functional form and estimate its parameters from data using least- squares regression. This approach is followed by (Asmundsson, Rardin, Turkseven, and Uzsoy 2009), who fit the functional form of (Srinivasan, Carey, Morton, et al. 1988) to data using least-squares regression, and then piecewise linearize the resulting concave function in their Allocated Clearing Function (ACF) formulation. However, they find that this approach yields poor fits, and use a percentile fit to obtain better results. (Missbauer 2011) argues that in order to represent the transient behavior of the resource the CF must incorporate additional independent variables. (Haeussler and Missbauer 2014) experiment with additional independent variables and functional forms using linear regression using data from simulation models and empirical data from a manufacturer of optical media. Similarly, (Albey, Bilge, and Uzsoy 2014) postulate several multi-dimensional CFs that are effective in the presence of setup times. (Kacar and Uzsoy 2014) use simulation data from a scaled-down semiconductor wafer fabrication facility and perform a similar analysis. These studies find that additional independent variables improve the fits of the CFs as measured by the adjusted R^2 value. In these studies the CFs are fit to data obtained without using a planning model, which creates concerns as the behavior of a queue with controlled releases is likely to be quite different from that of an open queueing system. Gopalswamy and Uzsoy (2017) propose an Iterative Refinement approach that obtains improved results at high utilization levels. (Kacar and Uzsoy 2015) use simulation optimization to obtain CF fits that optimize the performance of the production system, as opposed to the fit of the CF to the data, again obtaining substantially improved results. A significant fraction of the literature on CFs (Asmundsson, Rardin, and Uzsoy 2006, Kacar, Irdem, and Uzsoy 2012, Kacar, Monch, and Uzsoy 2013, Kacar, Mönch, and Uzsoy 2016) has focused on the wafer fabrication environment.

While improved statistical models for fitting CFs to empirical data remain an active area of research, this paper treats the parameters of the CF as random variables and embeds them in a stochastic formulation of the production planning problem. There are several alternative methods to address data uncertainty in optimization, including *stochastic programming* (SP), *simulation based optimization* (SO) and *robust optimization* (RO). SP (Prékopa 1995, Birge and Louveaux 2011, Ruszczyński and Shapiro 2003) assumes that the probability distribution (or its moments) of any uncertain data is known and the deterministic counterpart of the stochastic problem is tractable. In contrast to SP, RO does not require distributional assumptions, but instead assumes that the support of the uncertain data (its *uncertainty se*) is known, yielding a distribution-free tractable approach (Ben-Tal, El Ghaoui, and Nemirovski 2009, Ben-Tal and Nemirovski 2008, Bertsimas, Brown, and Caramanis 2011, Gorissen, Yanıkoğlu, and den Hertog 2015). Although RO has been successfully applied in a number of fields (Lobo 2000, Fredriksson, Forsgren, and Hårdemark 2011), applications in manufacturing are relatively scarce. Some examples are (Hood, Bermon, and Barahona 2003), (Barahona, Bermon, Gunluk, and Hood 2005), and (Chien and Zheng 2012), which focus on capacity planning or expansion in semiconductor manufacturing under demand uncertainty. (Aouam and Uzsoy 2015) compare SP and RO for a simple production system subject to stochastic demand and find that RO yields an effective approach although parameter values must be selected with care. (Bertsimas and Thiele 2006) apply RO to inventory management problems, while applications to production planning problems include (Lasserre and Mercé 1990), (Leung and Wu 2004), and (Verderame and Floudas 2009).

3 MULTI-DIMENSIONAL CF BASED PRODUCTION PLANNING MODEL

The single-variable CF of (Karmarkar 1989) shown in (1) has been shown empirically to perform well for both a single product and the aggregate output of a multiple product system in the absence of setups.

$$TH = \frac{C \times WIP}{M + WIP}. \tag{1}$$

C in (1) represents the maximum possible output in a period; M is a user-specified parameter controlling the curvature of the CF; WIP denotes the expected work-in-process (WIP) level or workload aggregated over products during the planning period; and TH the expected throughput in the period. In the multi-product case, given the machine capacity dedicated to products $j \neq i$, the remaining capacity is consumed by product i in a manner dependent on only the average WIP level of product i. Therefore, in determining the throughput of a given product, the capacity consumed by other products should be explicitly accounted for in the functional form. (Albey, Bilge, and Uzsoy 2014) propose several product-based disaggregated MDCFs for a single machine based on this logic with the form shown in (2).

$$TH_i \leq f(\vec{WIP}, \vec{TH}) \quad \forall i \in I. \tag{2}$$

TH_i in (2) denotes the expected throughput of product i, $\vec{WIP}$ the vector of average WIP levels of all products during the planning period, and $\vec{TH}$ the vector of throughput values of products $j \neq i$. WIP-based MDCFs generalize (1), and take the form

$$TH_i = \frac{\sum_j a_{ij} WIP_j^{avg}}{M_i + \sum_j b_{ij} WIP_j^{avg}} \quad \forall i \in I, \tag{3}$$

where C represents the maximum possible output and M_i controls the curvature of the CF for product i. WIP_i^{avg} denotes the average WIP of product i during the planning period; and TH_i the expected throughput of product i. The a_{ij}, b_{ij} in (3) are parameters estimated by curve fitting techniques (Albey, Bilge, and Uzsoy 2014).

(Albey, Bilge, and Uzsoy 2014) try several families of functions (ratios of polynomials, logarithmic transformations, several Taylor series and sigmoidals) and found that form in (3) performs well compared to others.

There are several motivations for MDCFs of this type for semiconductor manufacturing. The presence of multiple products with different processing and failure characteristics results in a more variable effective service time distribution, making it difficult for a single-variable CF to accurately estimate system output under different operating conditions. The output of a given tool group may depend on batching policies, sequencing decisions, and require setups and test wafers, rendering the output dependent on not only the aggregate workload but the mix of products available to the tool group during a planning period. These factors result in a situation where the output of a tool group can vary significantly over time due to a range of factors which are difficult to isolate. This paper acknowledges this situation by treating the parameters of the estimated MDCFs as random variables over a specified uncertainty set in a single-stage multi-product system. We define the following notation:

Indices:

i, j	$\equiv$ product index, $i, j \in I$, where I is the set of products
t	$\equiv$ period indices, $t \in T$, where T is the set of planning periods

Parameters:

ϵ_i	$\equiv$ unit processing time of product i
c_t	$\equiv$ available processing time of machine in period t
d_{it}	$\equiv$ demand of product i at period t
ϕ_i	$\equiv$ unit production cost of product i
ω_i	$\equiv$ unit WIP holding cost of product i
ρ_i	$\equiv$ unit release cost of product i
π_i	$\equiv$ unit finished goods inventory holding cost of product i
β_i	$\equiv$ unit backorder cost of product i

Decision Variables:

BO_{it} $\equiv$ amount of product i backordered at the end of period t

R_{it} $\equiv$ amount of product i released in period t

TH_{it} $\equiv$ amount of product i that is produced in period t

WIP_{it} $\equiv$ amount of work in process (WIP) of product i at the end of period t

FGI_{it} $\equiv$ amount of finished goods inventory (FGI) of product i at the end of period t

The MDCF based PRP model (MDCF-PRP) can now be stated as follows:

$$\min \sum_{i \in I} \sum_{t \in T} (\phi_i TH_{it} + \pi_i FGI_{it} + \omega_i WIP_{it} + \rho_i R_{it} + \beta_i BO_{it})$$

$$s.t.$$

$$FGI_{it-1} + TH_{it} + BO_{it} - BO_{it-1} - FGI_{it} = d_{it} \qquad \forall i \in I, \forall t \in T \qquad (4)$$

$$WIP_{it} = WIP_{it-1} - TH_{it} + R_{it} \qquad \forall i \in I, \forall t \in T \qquad (5)$$

$$WIP_{it}^{avg} = 0.5\epsilon_i(WIP_{it-1} + R_{it} + WIP_{it}) \qquad \forall i \in I, \forall t \in T \qquad (6)$$

$$\epsilon_i TH_{it} \leq \frac{\sum_j a_{ij} WIP_{jt}^{avg}}{M_i + \sum_j b_{ij} WIP_{jt}^{avg}} \qquad \forall i \in I, \forall t \in T \qquad (7)$$

$$\sum_i \epsilon_i TH_{it} \leq c_t \qquad \forall t \in T \qquad (8)$$

$$R_{it}, TH_{it}, FGI_{it}, WIP_{it}, BO_{it} \geq 0 \qquad \forall i \in I, \forall t \in T.$$

The MDCF-PRP model minimizes the sum of production, release, backorder, finished goods inventory and WIP holding costs over all periods. Constraints (4) and (5) are the balance constraints for finished goods inventory and WIP, respectively. Since the model tracks WIP levels at period boundaries but the MDCF uses the time-average WIP over the period, the time-average WIP is estimated by (6). The MDCF in constraint (7) limits the throughput of each product in each period. Constraint (8) ensures that the processing time required by all products completed in a period does not exceed the nominal capacity of the machine. $WIP_{i,0}$ and $FGI_{i,0}$ represent the initial WIP and FGI of product i at the beginning of the planning horizon.

4 ROBUST REFORMULATION OF MDCF BASED PRODUCTION PLANNING MODEL

The MDCF in (7) can be reformulated as

$$\sum_j a_{ij} WIP_{jt}^{avg} - \epsilon_i TH_{it}(M_i + \sum_j b_{ij} WIP_{jt}^{avg}) \geq 0 \quad \forall i \in I, \forall t \in T. \qquad (9)$$

We assume that the regressors $\mathbf{a}_i \in \mathbb{R}^n$ and $\mathbf{b}_i \in \mathbb{R}^n$ are uncertain in that they reside in uncertainty sets whose center points are given by the nominal data ($\mathbf{a}_i^0 \in \mathbb{R}^n$, $\mathbf{b}_i^0 \in \mathbb{R}^n$) obtained by regression. Formally, $\mathbf{a}_i := \mathbf{a}_i^0 + \mathbf{Q}_i\boldsymbol{\zeta}$ and $\mathbf{b}_i := \mathbf{b}_i^0 + \mathbf{W}_i\boldsymbol{\xi}$ where $\mathbf{Q}_i \in \mathbb{R}^{n\times n}$(:=diag($\mathbf{q}_i$)) and $\mathbf{W}_i \in \mathbb{R}^{n\times n}$(:=diag($\mathbf{q}_i$)), and the support of the primitive uncertainty parameters $\boldsymbol{\zeta} \in \mathbb{R}^n$ and $\boldsymbol{\xi} \in \mathbb{R}^n$ are assumed to reside in a unit box, i.e, $\mathcal{Z} = \{\boldsymbol{\zeta} : ||\boldsymbol{\zeta}||_\infty \leq 1\}$ and $\mathcal{X} = \{\boldsymbol{\xi} : ||\boldsymbol{\xi}||_\infty \leq 1\}$, where the infinity norm is given by

$||\boldsymbol{\xi}||_\infty := \sum_i |\xi_i|$. Correlation of uncertainties among different product types is assumed to be zero, and the uncertain vectors $\mathbf{a}_i$ and $\mathbf{b}_i$ are assumed to be independent, i.e., $\boldsymbol{\zeta}$ and $\boldsymbol{\xi}$ are independent from each other. More precisely, the off-diagonal elements of $\mathbf{W}_i$ and $\mathbf{Q}_i$ are zero, and $\mathbf{a}_i$ and $\mathbf{b}_i$ are defined in different uncertainty sets $\mathcal{Z}$ and $\mathcal{X}$ (which shall be extended later).

Consequently, the intractable semi-infinite representation of the robust constraint is

$$\begin{aligned}\sum_j a^0_{ij} WIP^{avg}_{jt} - \epsilon_i TH_{it}(M_i + \sum_j b^0_{ij} WIP^{avg}_{jt}) + \sum_j q_{ij}\zeta_j WIP^{avg}_{jt} \\ -\epsilon_i TH_{it} \sum_j w_{ij}\xi_j WIP^{avg}_{jt} \geq 0 \quad \forall \zeta \in \mathcal{Z}, \forall \xi \in \mathcal{X}, \forall i \in I, \forall t \in T.\end{aligned} \tag{10}$$

To remove the universal ($\forall$) quantifiers in (10), we minimize the left-hand side of the constraint over $\boldsymbol{\xi} \in \mathcal{X}$ and $\boldsymbol{\zeta} \in \mathcal{Z}$ for fixed values of WIP^{avg}_{jt} and TH_{it} and later take the dual and remove the "max" term. The 'tractable' robust counterpart (RC) of (10) thus obtained is:

$$\begin{aligned}\sum_j a^0_{ij} WIP^{avg}_{jt} - \epsilon_i TH_{it}(M_i + \sum_j b^0_{ij} WIP^{avg}_{jt}) - \sum_j \left| q_{ij} WIP^{avg}_{jt} \right| \\ -\epsilon_i TH_{it} \sum_j \left| w_{ij} WIP^{avg}_{jt} \right| \geq 0 \quad \forall i \in I, \forall t \in T.\end{aligned} \tag{11}$$

Under box uncertainty, the RC reformulation has the same complexity as (9), even though (11) requires additional variables to linearize the absolute values. Details on deriving the RC are given in (Gorissen, Yanıkoğlu, and den Hertog 2015, §2).

Alternatively, we can extend the independence assumption of $\mathbf{a}_i$ and $\mathbf{b}_i$ by defining them in the same ellipsoidal uncertainty set:

$$\hat{\mathcal{Z}} = \{\hat{\boldsymbol{\zeta}} \in \mathbb{R}^{2n} : ||\hat{\boldsymbol{\zeta}}||_2 \leq 1\},$$

where $\hat{\boldsymbol{\zeta}} := [\boldsymbol{\zeta}; \boldsymbol{\xi}]$ (";" operator concatenates two column vectors in the one with increased dimension) and the Euclidean norm is given by $||\hat{\boldsymbol{\zeta}}||_2 := \sqrt{\sum_i \zeta_i^2}$. Using the same notation, the uncertainty can be modeled as

$$[\mathbf{a}_i; \mathbf{b}_i] = [\mathbf{a}^0_i; \mathbf{b}^0_i] + [\mathbf{Q}_i, \mathbf{W}_i][\boldsymbol{\zeta}; \boldsymbol{\xi}]$$

and the 'tractable' RC becomes

$$\begin{aligned}\sum_j a^0_{ij} WIP^{avg}_{jt} - \epsilon_i TH_{it}(M_i + \sum_j b^0_{ij} WIP^{avg}_{jt}) \\ -\sqrt{\sum_j \left(q_{ij} WIP^{avg}_{jt}\right)^2 + (\epsilon_i TH_{it})^2 \sum_j \left(w_{ij} WIP^{avg}_{jt}\right)^2} \geq 0 \quad \forall i \in I, \forall t \in T,\end{aligned} \tag{12}$$

which can be efficiently solved by second-order cone programming (SOCP) solvers when TH_{it} is fixed.

Replacing (7) with either (11) or (12) yields different robust counterparts of the MDCF based production release planning model. In the next section, we demonstrate how uncertainty levels affect the different components of the PRP cost. In a second experiment, we compare the "vulnerability" (robustness) of the robust counterpart to that of deterministic model by sampling different values for estimated function parameters and comparing the frequency of infeasibility encountered in the models.

5 COMPUTATIONAL EXPERIMENTS

We briefly review the approach of (Albey, Bilge, and Uzsoy 2014) used to fit the MDCFs, and then present the computational experiments evaluating the performance of the robust formulation. We consider a single machine producing four products over a planning horizon of $T = 20$ planning periods. The total workload (i.e. demand) in each period is created to yield an average utilization of 95%, and then disaggregated to individual products by using randomly generated product mix vectors. Individual products, whose processing times follow lognormal distributions, are released to the shop floor after sequencing in a periodic pattern (Albey, Bilge, and Uzsoy 2014). Table 1 presents unit backorder, WIP holding, FGI holding, release, and production costs; the mean μ of the corresponding normal distribution and the coefficient of variation of the processing times for each product.

Simulations of the production system are performed to observe the output of the system under different product mix and workload levels. Using this data, parameters for the MDCFs are estimated using curve fitting techniques. A full factorial experimental design is run for multiple levels of each of these factors. The product mixes are chosen such that ratio of each product in the mix takes values of 0, 1, 2, 3 or 4, and the actual mix ratios are obtained by normalizing this combination vector. Hence a mix combination of $\{1,2,4,1\}$, whose mix ratios sum to 8, corresponds to a product mix of $\{1/8,2/8,4/8,1/8\} = \{0.125,0.250,0.500,0.125\}$. Since each product has five different levels for the mix factors this yields a total of $5^4 = 625$ combinations. However, some combinations such as $\{1,1,1,1\}$ and $\{2,2,2,2\}$ correspond to the same (in this case, equal) allocation of demand among products. Therefore to avoid unnecessary weighting of combinations in the experimental design, the 96 duplicate mixes are removed. The final number of mix combinations is therefore $625-96 = 529$. For each product mix, the system is simulated for 1000 periods, for a total of 529,000 periods of simulation. The data collection phase is followed by a fitting procedure to estimate the parameters of the MDCF. Once the MDCF has been obtained, the planning models are solved using the KNITRO solver through the GAMS interface. We consider two different robust optimization models, Robust-v1 and Robust-v2, which are obtained by replacing MDCF constraint (7) in deterministic model by (11) and (12), respectively.

Our first experiment, whose results are presented in Table 2, compares the objective function value of the deterministic and robust models to validate the behavior of the latter. As the robust model becomes more conservative, it releases less work and backorders more than the deterministic model, yielding higher total costs. This is expected since the robust solution is immunized against uncertainty, i.e., its feasible space is tighter than that of the deterministic model.

A direct comparison of the objective function values of the deterministic and robust models may be misleading because they use different definitions of feasibility. In particular, a deterministic solution may be infeasible for some (or all) realizations of the uncertain parameter that differ from the nominal data. To examine this issue we run a Monte Carlo simulation experiment by sampling uncertain data from two box uncertainty sets whose box limits are set 10% and 20% away from the point estimates of the uncertain parameters. Experiments are run for 100 random pairs (q_{ij}, w_{ij}) for each uncertainty scenario and the feasibility of (7) for the deterministic and the robust solution are compared. In the second experiment, we give both models a recovery option by allowing outsourcing at an additional cost when initial plans are not met. The unit outsourcing cost is set to 150% of unit backorder cost.

Results are presented in Table 3. The third column shows the percentage of infeasible instances before introducing the recovery option, while the fourth column shows the percent of MDCF constraints that are infeasible for the sampled uncertainty realizations (out of all constraints over all periods). The fifth column lists the expected outsourcing cost with the recovery option. A significant fraction of the deterministic solutions become infeasible in the presence of uncertainty in the MDCF parameters. The fraction of periods with an infeasibility in the MDCF constraint increases with the degree of uncertainty, as would be expected. The robust model reduces the number of infeasibilities by an order of magnitude, and the number of infeasibilities does not increase with the level of uncertainty.

Table 1: Cost and demand related parameter values.

Product	Backorder	WIP holding	FGI holding	Release	Production	μ	c.v.
1	15	1	1.5	5	2	100	0.82
2	25	1	1.5	5	2	150	0.54
3	35	1	1.5	5	2	200	0.41
4	50	1	1.5	5	2	300	0.27

Table 2: Comparison of deterministic and robust models at uncertainty levels 10% and 20%.

	Deterministic	Uncertainty level = 10%		Uncertainty level = 20%	
		Robust-v1	Robust-v2	Robust-v1	Robust-v2
Release Cost	11.636	11.126	10.865	11.182	10.336
FGI Holding Cost	628	594	450	598	712
WIP Holding Cost	2.039	1.191	4.302	1.195	695
BO Cost	118.341	166.026	156.093	166.857	161.993
Production Cost	4.601	4.402	4.058	4.424	4.111
Total Cost	137.244	183.341	175.769	184.256	177.846

Table 3: Results of Monte Carlo experiment

Uncertainty	Model	Infeasibility	MDCF constraint infeasibility	Outsourcing cost
10%	Deterministic	49%	30%	1890
10%	Robust-v2	10%	4%	121
20%	Deterministic	52%	33%	4409
20%	Robust-v2	10%	4%	161

6 CONCLUSIONS

Although nonlinear CFs have proven to be useful metamodels representing the behavior of production resources, a satisfactory framework for estimating them for practical production systems remains elusive. Current practice focuses on developing point estimates using conventional least-squares regression approaches, which have not yet yielded satisfactory solutions. Some of the reasons for these difficulties have been discussed by (Gopalswamy and Uzsoy 2017). This paper takes an alternative approach which seeks to explicitly account for the uncertainties in the MDCF parameters induced by estimation errors using robust optimization. We compare the performance of a production planning model using a multi-dimensional CF and its robust counterpart under several experimental settings. As expected, as the level of uncertainty in the function parameters in the robust counterpart is increased, the resulting production plan deviates from the optimal solution of the deterministic model. However, the production plans found by the robust counterpart are proven to be less vulnerable to estimation errors, which are a natural consequence of fitting processes.

The application of these methods to support production planning models for semiconductor manufacturing facilities poses a number of challenges, but also presents some unique opportunities. The prevalence of real-time shop-floor data from manufacturing execution systems in wafer fabs provides potentially very large data sets to which these approaches can be applied. However, the sources of uncertainty affecting output from the various tools groups in wafer fabs need to be better characterized to determine the best representation of uncertainty sets to use with this approach. The presence of timing constraints between certain steps in the wafer fabrication process may affect the performance of such models by introducing strong dependencies between the output of different tool groups, which it is not immediately clear how to model. Finally, an important issue remains the size of the resulting formulations and their computational complexity. Even single-variable CFs require the definition of decision variables and constraints fore ach operation of each product performed on each tool group, which leads to a significant increase in model size over the linear programming models with workload-independent lead times (Kacar, Mönch, and Uzsoy 2016).

A number of interesting directions for future research are apparent. The first of these is to compare the realized performance of the production system under the different planning models by simulating the execution of the proposed release schedules. Based on our experience with using two different models with different sets of assumptions for the same underlying problem, in this case dynamic lot sizing (Kang, Albey, Hwang, and Uzsoy 2014), it is quite possible that although the objective function value of the deterministic model may be better than those of the robust models, the robust model may result in significantly better cost performance. Another interesting direction would be to formulate the problem of fitting the MDCF to the data as a robust optimization problem, instead of using the conventional minimization of the sum of squared errors. Finally, application of this approach to fitting piecewise linear single-variable CFs that can be directly implemented in the formulation of (Asmundsson, Rardin, Turkseven, and Uzsoy 2009) as well as other functional forms would be of great interest.

REFERENCES

Albey, E., and Ü. Bilge. 2011. "A hierarchical approach to FMS planning and control with simulation-based capacity anticipation". *International Journal of Production Research* 49 (11): 3319–3342.

Albey, E., Ü. Bilge, and R. Uzsoy. 2014. "An exploratory study of disaggregated clearing functions for production systems with multiple products". *International Journal of Production Research* 52 (18): 5301–5322.

Albey, E., Ü. Bilge, and R. Uzsoy. 2017. "Multi-dimensional clearing functions for aggregate capacity modelling in multi-stage production systems". *International Journal of Production Research* 55 (14): 4164–4179.

Albey, E., and R. Uzsoy. 2015. "Lead time modeling in production planning". In *Winter Simulation Conference (WSC), 2015*, 1996–2007. IEEE.

Aouam, T., and R. Uzsoy. 2015. "Zero-order production planning models with stochastic demand and workload-dependent lead times". *International Journal of Production Research* 53 (6): 1661–1679.

Asmundsson, J., R. L. Rardin, C. H. Turkseven, and R. Uzsoy. 2009. "Production planning with resources subject to congestion". *Naval Research Logistics (NRL)* 56 (2): 142–157.

Asmundsson, J., R. L. Rardin, and R. Uzsoy. 2006. "Tractable nonlinear production planning models for semiconductor wafer fabrication facilities". *IEEE Transactions on Semiconductor Manufacturing* 19 (1): 95–111.

Axsäter, S. 2010. "A capacity constrained production-inventory system with stochastic demand and production times". *International Journal of Production Research* 48 (20): 6203–6209.

Bang, J.-Y., and Y.-D. Kim. 2010. "Hierarchical production planning for semiconductor wafer fabrication based on linear programming and discrete-event simulation". *IEEE Transactions on Automation Science and Engineering* 7 (2): 326–336.

Barahona, F., S. Bermon, O. Gunluk, and S. Hood. 2005. "Robust capacity planning in semiconductor manufacturing". *Naval Research Logistics* 52 (5): 459–468.

Ben-Tal, A., L. El Ghaoui, and A. Nemirovski. 2009. *Robust Optimization*. Princeton Series in Applied Mathematics. Princeton University Press.

Ben-Tal, A., and A. Nemirovski. 2008. "Selected topics in robust convex optimization". *Mathematical Programming* 112 (1): 125–158.

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

Bertsimas, D., and A. Thiele. 2006. "A robust optimization approach to inventory theory". *Operations Research* 54 (1): 150–168.

Birge, J. R., and F. V. Louveaux. 2011. *Introduction to Stochastic Programming*. Springer.

Chien, C.-F., and J.-N. Zheng. 2012. "Mini–max regret strategy for robust capacity expansion decisions in semiconductor manufacturing". *Journal of Intelligent Manufacturing* 23 (6): 2151–2159.

Fredriksson, A., A. Forsgren, and B. Hårdemark. 2011. "Minimax optimization for handling range and setup uncertainties in proton therapy". *Medical Physics* 38 (3): 1672–1684.

Gopalswamy, K., and R. Uzsoy. 2017. "An Iterative Refinement Approach to Fitting Clearing Functions to Data from Simulation Models of Production Systems". In *Winter Simulation Conference (WSC)*. IEEE.

Gorissen, B. L., İ. Yanıkoğlu, and D. den Hertog. 2015. "A practical guide to robust optimization". *Omega* 53:124–137.

Graves, S. C. 1986. "A tactical planning model for a job shop". *Operations Research* 34 (4): 522–533.

Haeussler, S., and H. Missbauer. 2014. "Empirical validation of meta-models of work centres in order release planning". *International Journal of Production Economics* 149:102–116.

Hood, S. J., S. Bermon, and F. Barahona. 2003. "Capacity planning under demand uncertainty for semiconductor manufacturing". *IEEE Transactions on Semiconductor Manufacturing* 16 (2): 273–280.

Hung, Y.-F., and R. C. Leachman. 1996. "A production planning methodology for semiconductor manufacturing based on iterative simulation and linear programming calculations". *IEEE Transactions on Semiconductor manufacturing* 9 (2): 257–269.

Irdem, D. F., N. B. Kacar, and R. Uzsoy. 2010. "An exploratory analysis of two iterative linear programming–simulation approaches for production planning". *IEEE Transactions on Semiconductor Manufacturing* 23 (3): 442–455.

Kacar, N. B., D. F. Irdem, and R. Uzsoy. 2012. "An experimental comparison of production planning using clearing functions and iterative linear programming-simulation algorithms". *IEEE Transactions on Semiconductor Manufacturing* 25 (1): 104–117.

Kacar, N. B., L. Monch, and R. Uzsoy. 2013. "Planning wafer starts using nonlinear clearing functions: A large-scale experiment". *IEEE Transactions on Semiconductor Manufacturing* 26 (4): 602–612.

Kacar, N. B., L. Mönch, and R. Uzsoy. 2016. "Modeling cycle times in production planning models for wafer fabrication". *IEEE Transactions on Semiconductor Manufacturing* 29 (2): 153–167.

Kacar, N. B., and R. Uzsoy. 2014. "A comparison of multiple linear regression approaches for fitting clearing functions to empirical data". *International Journal of Production Research* 52 (11): 3164–3184.

Kacar, N. B., and R. Uzsoy. 2015. "Estimating clearing functions for production resources using simulation optimization". *IEEE Transactions on Automation Science and Engineering* 12 (2): 539–552.

Kang, Y., E. Albey, S. Hwang, and R. Uzsoy. 2014. "The impact of lot-sizing in multiple product environments with congestion". *Journal of Manufacturing Systems* 33 (3): 436–444.

Karmarkar, U. S. 1989. "Capacity loading and release planning with work-in-progress (WIP) and leadtimes". *Journal of Manufacturing and Operations Management* 2:105–123.

Lasserre, J., and C. Mercé. 1990. "Robust hierarchical production planning under uncertainty". *Annals of Operations Research* 26 (1): 73–87.

Leung, S. C., and Y. Wu. 2004. "A robust optimization model for stochastic aggregate production planning". *Production Planning & Control* 15 (5): 502–514.

Lobo, M. S. 2000. *Robust and convex optimization with applications in finance*. Ph. D. thesis, stanford university.

Missbauer, H. 2002. "Aggregate order release planning for time-varying demand". *International Journal of Production Research* 40 (3): 699–718.

Missbauer, H. 2009. "Models of the transient behaviour of production units to optimize the aggregate material flow". *International Journal of Production Economics* 118 (2): 387–397.

Missbauer, H. 2011. "Order release planning with clearing functions: a queueing-theoretical analysis of the clearing function concept". *International Journal of Production Economics* 131 (1): 399–406.

Prékopa, A. 1995. *Stochastic Programming*. Kluwer Academic Publishers.

Ruszczyński, A. P., and A. Shapiro. 2003. *Stochastic Programming (handbooks in operations research and management science)*. Elsevier.

Srinivasan, A., M. Carey, T. E. Morton et al. 1988. "Resource pricing and aggregate scheduling in manufacturing systems". Technical report, Carnegie Mellon University, Tepper School of Business.

Stinstra, E., and D. Den Hertog. 2008. "Robust optimization using computer experiments". *European Journal of Operational Research* 191 (3): 816–837.

Verderame, P. M., and C. A. Floudas. 2009. "Operational planning of large-scale industrial batch plants under demand due date and amount uncertainty. I. Robust optimization framework". *Industrial & Engineering Chemistry Research* 48 (15): 7214–7231.

Zipkin, P. H. 2000. *Foundations of inventory management*.

AUTHOR BIOGRAPHIES

ERİNÇ ALBEY is an assistant professor in the Industrial Engineering Department at Özyeğin University. He received his B.Sc., M.Sc. and Ph.D. degrees in Industrial Engineering from Boğaziçi University, Istanbul, Turkey. He worked as a researcher at Bogazici University Flexible Automation and Intelligent Manufacturing Laboratory and as a postdoctoral research associate in the Edward P. Fitts Department of Industrial and Systems Engineering, North Carolina State University. His e-mail address is erinc.albey@ozyegin.edu.tr.

İHSAN YANIKOĞLU is an assistant professor in the Industrial Engineering Department at Özyeğin University. He received his B.Sc. and M.Sc. degrees in Industrial Engineering from Bilkent University, and his M.Phil. and Ph.D. degrees in Operations Research and Econometrics from Tilburg University. His research focuses on developing practicable robust optimization methodologies and their applications. His email address is ihsan.yanikoglu@ozyegin.edu.tr.

REHA UZSOY is Clifton A. Anderson Distinguished Professor in the Edward P. Fitts Department of Industrial and Systems Engineering at North Carolina State University. He holds B.Sc. degrees in Industrial Engineering and Mathematics and an MS in Industrial Engineering from Boğaziçi University, Istanbul, Turkey, and received his Ph.D. in Industrial and Systems Engineering in 1990 from the University of Florida. His teaching and research interests are in production planning, scheduling, and supply chain management. His email address is ruzsoy@ncsu.edu.