Proceedings of the 2021 Winter Simulation Conference
S. Kim, B. Feng, K. Smith, S. Masoud, Z. Zheng, C. Szabo, and M. Loper, eds.

AN EVALUATION OF STRATEGIES FOR JOB MIX SELECTION IN JOB SHOP PRODUCTION ENVIRONMENTS - CASE : A PHOTOLITHOGRAPHY WORKSTATION

Amir Ghasemi
Cathal Heavey

CONFIRM Research Centre
School of Engineering
University of Limerick
Limerick, V94 T9PX, IRELAND

ABSTRACT

In this research the impact of job mix selection in each production shift in a job shop production environment is examined. This is a critical question within photolithography workstations in semiconductor manufacturing systems. For this purpose, a recently developed Simulation Optimization (SO) method named Evolutionary Learning Based Simulation Optimization (ELBSO) is implemented to solve a set of designed Stochastic Job Shop Scheduling problems captured from a real semiconductor manufacturing data set. Experiment results indicate that the best performance in each shift occurs when machines are flexible in terms of processing different job operations, and the selected jobs for a certain shift have as equal as possible due dates.

1 INTRODUCTION

The semiconductor industry is characterized by increasing complexity of manufacturing processes due to elementary operations where finer geometries are realized on chips. This increases the number of constraints and with an increase in automation (mainly as a result of the increase in disc area), there is an increase in infrastructure and equipment costs. This, with a higher level of technology renewal rates places an ever-increasing pressure on the cost of wafers due to worldwide competition, and high customer requirements in terms of quality (Mönch et al. 2018). Within production, a photolithography toolset, requires high capital investments, is one of the most crucial processes in semiconductor manufacturing. It is mostly regarded as a bottleneck process due to the layered nature of wafer fabrication, especially for the case of Application Specific Integrated Circuit (ASIC) fabrication environments with high product mix portfolios and low volumes. In an ASIC fab, a diverse range of recipes exist due the range of products produced. Since the circuits are made up of layers, with every wafer passing through the photolithography area, which maybe up to 40 times, this typically makes photolithography a bottleneck resource. With the performance of a system determined by the bottleneck resource, optimal capacity allocation and job scheduling of a photolithography work area ensures improvement in the performance of the whole fab (Ghasemi et al. 2020).

It is worth mentioning that, similar to many production environments, the photolithography workstation can be viewed as a job shop production environment. Indeed, job shop scheduling problems are one of the most challenging problems in most production industries (Zhang et al. 2019). Within the photolithography workstation, there are two main operational constraints. First, certain machines within the photolithography tool will be qualified for different recipes (machine process capability constraints). Secondly, for some critical layers, certain machines within the toolset will be required to be used to ensure the quality of the integrated circuit (machine dedication constraints) (Ghasemi et al. 2020). Considering these two constraints semiconductor production planners decide the mix of wafers going through the photolithography workstation in each production shift. Accordingly, they follow certain strategies to select the production mix for each

978-1-6654-3311-2/21/$31.00 ©2021 IEEE

shift, which could be based on due date, processing times, or any other influencing factor of orders. The question here is about how to define the best strategy for improving the utilization and delivery levels of the photolithography workstation?

Scheduling of photolithography are large stochastic and complex problems. Explosive growth in computing power and also recent advances in Simulation Optimization (SO) techniques provides a possible means of addressing these class of problems. In Ghasemi et al. (2021) an Evolutionary Learning Based Simulation Optimization (ELBSO) method was presented to solve a stochastic job shop scheduling problem. The article demonstrated the superiority of ELBSO when compared to three different algorithms in terms of the quality of solutions and the reduction in computation time. SO methods generally expend a large amount of computation time to find a solution to the given problem, due to the necessity for the required number of simulation replications. In ELBSO, which uses Ordinal Optimization (OO), this is negated where a metamodel is used within the first phase. This greatly reduces the amount of time used to carry out simulation replication, allowing this computational time to be applied to the optimization of the addressed problem. The metamodel used within ELBSO is Genetic Programming (GP), which was found to provide very accurate solutions within the ELBSO algorithm.

The main purpose of this article is to examine the best strategy for wafer mix (job mix) considering their due dates and routes through photolithography machines to minimize total weighted earliness and tardiness costs. Moreover, we apply the ELBSO to data extracted from a real photolithography toolset, which, as stated earlier, can be classified as a Stochastic Job Shop Scheduling Problem (SJSSP). In Section 2, we will briefly review past work on optimization SJSSPs, then in Section 3, we will describe the relevance of this research to the photolithography toolset. Section 4 defines the ELBSO implementation procedure. Experiments and results are provided in Section 5. Finally, Section 7 concludes the paper and looks at possible follow-on studies.

2 OVERVIEW OF SO OF SJSSPS

Over recent years, a large body of research has been published on Job Shop Scheduling Problems (JSSP), which are one of the basic models used in manufacturing. JSSP is one of the famous mathematical optimization problems that has been proved to be NP-hard (Horng et al. 2012). Although stochasticity has been known as a crucial part of most industrial operations in the scheduling literature, there is little attention to solving SJSSPs, while Deterministic Job Shop Scheduling Problems (DJSSPs) have been widely researched (Winands et al. 2011).

As stated above, recent advances in SO research and the explosive growth in computing power have made it possible to optimize complex manufacturing system problems. In fact, Discrete Event Simulation Models (DESMs) are important tools used as a predictor of performance, allowing examination of the likely behavior of a proposed manufacturing system under experimental conditions (Trigueiro de Sousa Junior et al. 2019). While DESMs do not directly provide explanations for the observed system behavior, it is essentially a trial and error methodology, and although attention to experimental design techniques enhances its value, it does not provide a method of optimization. On the other hand, optimization techniques are key tools to improve decisions within almost all systems.

Integrating simulation models with optimization methods could establish promising Decision Support Tools (DSTs) benefiting from the advantages of both tools. Thus, SO techniques have been known as one of the most promising techniques to tackle large and stochastic real production problems such as SJSSPs. Researchers applied SOs to SJSSPs to allow an optimizer(s) to seek better solutions when integrated with DESMs (Figueira and Almada-Lobo 2014).

However, the objective calculated using simulation replications has a high computation cost. Therefore, most researchers have applied three main techniques to replace the high number of simulation replications. One approach used by several researchers, such as Horng et al. (2012) and Yang et al. (2014), is to use a small number of simulation replications executed during the search phase. For instance, although it is reported by Horng et al. (2012) that 10^5 simulation replications are enough to ensure the accuracy of the objective values, they performed 368 simulation replications to calculate the objective values of the proposed SJSSP (in the exploration phase, i.e., phase 1 of OO). A second approach is where researchers such as Shen and Zhu (2016) and Jamili (2019), convert SJSSP mathematical models to deterministic ones where a level of robustness and/or confidence is achieved when optimized. Typically, evolutionary methods are proposed to solve the deterministic models. A third approach is where metamodels are used. This was reported in Horng et al. (2012) where an Radial Basis Functions (RBF) metamodel was used in phase 1 of OO for the optimization of a hotel booking limit problem. The third approach is the most advantageous one as other methods ignore a series of stochastic scenarios while they do not provide information on the ignored scenarios. However, due to the highly complex nature of SJSSPs, there is no metamodeling method applied to SJSSPs within the literature.

To metamodel DESMs, researchers have presented a variety of methods and concepts, such as RBF (Hussain et al. 2002), Kriging (KG) metamodeling (Kleijnen 2009), while Artificial Neural Network (ANN) has been known as the predominant approach to metamodel DESMs (Dunke and Nickel 2020). Can and Heavey (2012) compared both GP and ANN in metamodeling DESMs. For the industrial case studies considered, the results showed that GP outperforms ANN in metamodeling DESMs. Surprisingly, there is no research in the literature implementing GP-based metamodels to SJSSP DESMs.

This paper contributes to the literature by applying the SO method presented in Ghasemi et al. (2021) to an SJSSP using case data set from a semiconductor manufacturing photolithography workstation, which aims to investigate the impact of job mix and due date selection for each production shift on the overall workstation performance.

2.1 Problem Description

In this section, the mathematical model of SJSSP is detailed. Table 1 summarises the notation used in this section. According to the scheduling environments notation provided by Pinedo (2016), Graham et al. (1979), we study a $J_{NM}|Range(p_i)|(\alpha \times T) + (\beta \times E)$ problem. That defines a job shop problem (J) with NM machines under condition of random processing times, which are between a specified range, and with the goal of optimizing the total weighted expected tardiness (T) and earliness (E).

According to Shen and Zhu (2016), in SJSSP the goal of scheduling all operations of N jobs on NM machines is to minimize the expected value of total costs. Therefore, the objective function for SJSSP is as follows:

$$MinF = Min\left(\sum_{j=1}^{N}((Max\{0, C_{j,NM} - d_j\} \times CT_j) + (Max\{0, d_j - C_{j,NM}\} \times CE_j))\right) \tag{1}$$

Shen and Zhu (2016) also provided the following sets of constraints for SJSSP: Sequence constraints: a job on a machine can start processing after completing its previous processing procedure,

$$S_{jw} \geq S_{j,w-1} + t'_{wj}, j \in \{1,...,N\}; w \in \{1,...,NM\} \tag{2}$$

Resource constraints: a job on a machine can start processing after the completion of the previous job,

$$O_{ml} \geq O_{m,l-1} + t'_{v_{m,l-1},w}, \text{ where } \; b_{v_{m,l-1},w} = m, w \in \{1,...,NM\}; l \in \{1,...,N\} \tag{3}$$

Time constraints: each job can be available at time zero,

$$S_{jw} \geq 0, j \in \{1,...,N\}; w \in \{1,...,NM\} \tag{4}$$

Table 1: Notations table.

Indices and Sets	
$j =$	Jobs index, $j \in \{1,...,\lvert N\rvert\}$.
$i,i' =$	operations ids, $i,i' \in \{1,...,\lvert NO\rvert\}$.
$w =$	operations indices, $w \in \{1,...,\lvert NM\rvert\}$.
$m =$	Machine index, $m \in \{1,...,\lvert NM\rvert\}$.
$l =$	positions on machines index, $l \in \{1,...,\lvert N\rvert\}$.
$k =$	Queuing position index, $k \in \{1,...\lvert NO\rvert\}$.
$\Omega =$	Precedence orders sets defining the execution precedence of operations of the same jobs.
$s =$	Simulation replication index, $s \in \{1,...,\lvert SL\rvert\}$.
Parameters	
N	Number of jobs.
NO	Total number of operations.
NO_j	Total number of operations for job j.
NM	Number of machines.
SL	Total number of simulation replications indexed by s.
d_j	Due date of job j.
P'_i	Stochastic processing time of operation id i.
ϕ_i	Probability distribution of processing time of operation id i.
t'_{wj}	Stochastic processing time of operation w of job j.
CE_j	Earliness cost of job j caused by inventory costs.
CT_j	Tardiness cost of job j caused by tardiness in delivering the job.
$V = (v_{ml})_{NM\times N}$	The process matrix, where $v_{ml} \in \{1,...,\lvert N\rvert\}$ denotes that job v_{ml} is processed at machine m in position l.
$B = (b_{jw})_{N\times NM}$	The operation matrix, where $b_{jw} \in \{1,...,\lvert NM\rvert\}$ denotes a machine. Elements $\{b_{j1}, b_{j2}, ..., b_{jNO_j}\}$ in the j^{th} row represent operations of job j. That is the job j is processed orderly on machines $b_{j1}, b_{j2}, ..., b_{jNO_j}$.
Decision Variables	
S_{jw}	The starting time of operation w of job j.
O_{ml}	The starting time of l^{th} job processed on machine m.
C_{jw}	The completion time of w^{th} operation of job j.
Q_{ml}	The completion time of a job processed on l^{th} position of machine m.
X_{ik}	If operation id i is assigned to the k^{th} postion of the dispatching queue, then $X_{ik} = 1$, 0 otherwise.
T'_{js}	The stochastic tardiness time of job j in the simulation replication s.
E'_{js}	The stochastic earliness time of job j in the simulation replication s.
Functions	
f_s	The objective calculation function for the simulation replication s.

In this research, the above model is converted to a SO model, where, during the optimization procedure, values will be obtained from an evaluation model, which will be a metamodel in phase 1 of OO or a simulation model in phase 2 (Ghasemi et al. 2021). Equation 1 is transformed into Equation 5, for use with the evaluation model, where f_s calculates total costs of solution X_{ik} by the simulation replication s. Besides, F defines the fitness value for an SJSSP solution.

$$Min\ F = Min\left(\frac{1}{SL}\sum_{s=1}^{SL} f_s\left(\{\cup_{i=1}^{NO}\cup_{k=1}^{NO} X_{ik}\}, P'_i\right)\right) \tag{5}$$

In each solution, X_{ik}, is a binary variable that denotes if operation i is assigned to the k^{th} position of the dispatching queue, with the above objectives having the following constraints. Each job should be assigned to one of the existing dispatching queue positions as follows:

$$\sum_{k=1}^{NO} X_{ik} = 1, \forall i \in \{1,...,NO\} \tag{6}$$

Moreover, it must be guaranteed that for each existing position in dispatching queue k, at most one operation is assigned. Equation 7 defines this constraint as follows:

$$\sum_{i=1}^{NO} X_{ik} = 1, \forall k \in \{1,...,NO\} \tag{7}$$

Precedence of operations of a job are defined in the set Ω. Consequently, Equations 8 and 9 construct a precedence relationship between two operations i and i' from the same job j in the SJSSP. In other words, when operation i (assigned to the position k) precedes operation i', operation i' must be assigned to a position k' ($k' > k$) on the dispatching queue.

$$X_{ik} \geq X_{i'k'}, \forall i, i' \in \Omega, \forall k, k' \in \{1, ..., NO\}, k < k' \tag{8}$$

$$X_{ik} - (X_{ik} X_{i'k'}) \geq X_{i'k'}, \forall i, i' \in \Omega, \forall k, k' \in \{1, ..., NO\}, k > k' \tag{9}$$

Equation 10 defines the decision variable feature in the model:

$$X_{ik} \in \{0, 1\}, \forall i, k \in \{1, ..., NO\} \tag{10}$$

All in all, Equations 6 to 10 guarantee the feasibility of queue solution X_{ik} by considering both assignment and precedence constraints.

3 PHOTOLITHOGRAPHY AREA

The photolithography process includes three main steps. These steps are coat, expose, and develop. First, the wafer is coated with a thin film of a photosensitive polymer, called the photoresist strip. Then, in the "expose" step, the wafer is exposed with ultraviolet light (UV) in order to print the circuit pattern onto the wafer. This is done using a reticle, which is a chrome patterned glass that defines the circuit pattern. Since the circuits are made up of layers, every wafer passes through the photolithography area up to 40 times to process each layer. Accordingly, there are different layers to be processed on a wafer in the photolithography workstation. This production environment can be seen as a job shop environment, where wafers and layers refer to jobs and operations in the job shop scheduling literature, respectively (Ghasemi et al. 2021).

To process each layer of a wafer, there is a set of eligible machines, where the eligibility essentially refers to considering both machine process capability and machine dedication constraints. In each production shift, the production planner decides the wafer mix going through the photolithography workstation considering different parameters of the system. In this research, we examine the impact of jobs due dates and machines flexibility in producing different layers (operations) of wafers (jobs) on the shop performance in each shift. To perform this analysis, the following assumptions are considered within the SJSSP: 1) In a production shift, an equal number of operations is assigned to all machines. 2) In a production shift, the number of operations for all jobs are equal. 3) Jobs production plan is predefined considering both machine process capability and machine dedication constraints (i.e., the assignment of operations to machines in a shift is predefined considering the mentioned operational constraints). Considering these three assumptions and the SJSSP model presented in Section 2.1, the problem inputs are: Number of Jobs (N), Number of Operations for each job (NO_j), Number of Machines (NM), Due Dates (d_j), Operations Processing Times (t'_{wj}), Cost of Earliness (CE_j), Cost of Tardiness (CT_j), Predefined Operations Assignment to the Machines (B), while the problem output is the total weighted earliness and tardiness costs for a production shift. Ghasemi et al. (2021) implemented ELBSO to this problem and examined the impact of CE_j and CT_j on the total objective value. In this research, we implement the same method, but we use data sets from a semiconductor manufacturing system to retrieve processing time distributions, and we examine the impact of B and Due Dates on the total objective value.

In this research, data was collected for three months from a frontend semiconductor fab with a data set consisting of 929,178 rows of data. The data set analysis in this research has already been published in Ghasemi et al. (2020) for a capacity allocation problem in a photolithography area. Here, we use this data set to parametrize processing times within SJSSP. To clarify, these data sets include the size of lots (TRACKINMAINQTY), photolithography entrance times in each route (TRACKINTIME), and photolithography exit times in each route (TRACKOUTTIME). Using track in and track out times of each lot, the processing time for each lot could be easily calculated. It is worth mentioning here that all data sets are sampled from the Manufacturing Execution System (MES) at the considered semiconductor manufacturer and are verified by them before analysis. Since some lots do not require photolithography in each route, their processing time in the photolithography area equals to zero. Therefore, all jobs with processing times equal to zero are deleted.

4 ELBSO IMPLEMENTATION

As mentioned above, in this research, we implement the ELBSO algorithm proposed by Ghasemi et al. (2020) to the photolithography area considered as an SJSSP. In ELBSO, after developing a DESM of the proposed SJSSP, Figure 1 shows the general structure of the ELBSO algorithm. In the GP preparation phase of the ELBSO algorithm, a data set is produced using the DESM of the SJSSP for different replications and scenarios. Then, using this data set, a GP model is trained to estimate the SJSSP objective value. In phase one, an initial random set *pop* of population size *popsize* is generated. Next, a set *npop* of population size *npopsize* of offsprings is created using a Neighborhood Search Function (NSF). All solution objective values in both *pop* and *npop* sets are evaluated using GP in the selection module, and *popsize* best solutions are selected for the next generation. These steps are replicated until *termination* is achieved. Subsequently, phase two is started by running the simulation model R_p times for each solution $p \in pop$ (R_p defines the number of simulation replications in phase 2 for each solution). In this phase, a method called Simulation Budget Allocation (SBA) is used, however other procedures such as Optimal Computation Budget Allocation (OCBA) or rank and selection could be used. Details of the ELBSO structure are provided in Ghasemi et al. (2021). As suggested by Ghasemi et al. (2021) the ELBSO parameters are set as follows: the solution population size ($popsize = 1000$), offspring solution population size ($npopsize = 2000$), total number of simulation replications to calculate each solution's objective value ($SL = 10^5$), total number of algorithm iterations ($iteration = 100$), and as ELBSO is an OO-based evolutionary SO methods, in the second phase the total number of sub phases executed simulation experiments is equal to six ($nSub = 6$). Moreover, ELBSO will be terminated in 2000 seconds. As the results will vary for each execution of the algorithm, 10 replications of each algorithm are used.

5 EXPERIMENTS RESULTS

According to the research question defined in Section 1 (i.e., examining the impact of job mix selection on the total shop performance), here, we examine the impact of both job due dates and operation routes through machines on the total weighted earliness and tardiness costs of a production shift. This supports the photolithography workstation planning procedure by examining different product mix selection strategies. Thus, considering the discussed case data set, ELBSO is used to tackle SJSSP. As mentioned in Section 3, in each production shift, the production planner selects a mix of jobs to be processed on machines. Then, considering operational constraints (e.g., machine dedication), the operation routes through machines are defined. The question here is that: what is the best strategy in terms of operation routes through machines and job due dates variability to be considered for selecting a job mix for each shift? Here, we consider three strategies for each of the mentioned factors. We use two measures, Total Flexibility Ratio (TFR) and Due Date Standard Deviation (*DDS*) to study this issue. We introduce three TFR categories: TFR *Category* 1: where each machine can process each operation (i.e., the first, second,..., and last operations of jobs which provides a level flexibility); TFR *Category* 2: where machine flexibility level is set at a medium level; TFR *Category* 3: where machine flexibility is set to the lowest level, where each machine is capable of processing a certain operation number similar to flow shop production environments. Due dates are also experimented under three categories: *DDS Category* 1: jobs have equal due dates; *DDS Category* 2: there is a low level of variability in due dates; *DDS Category* 3: there is a considerable variation between jobs due dates.

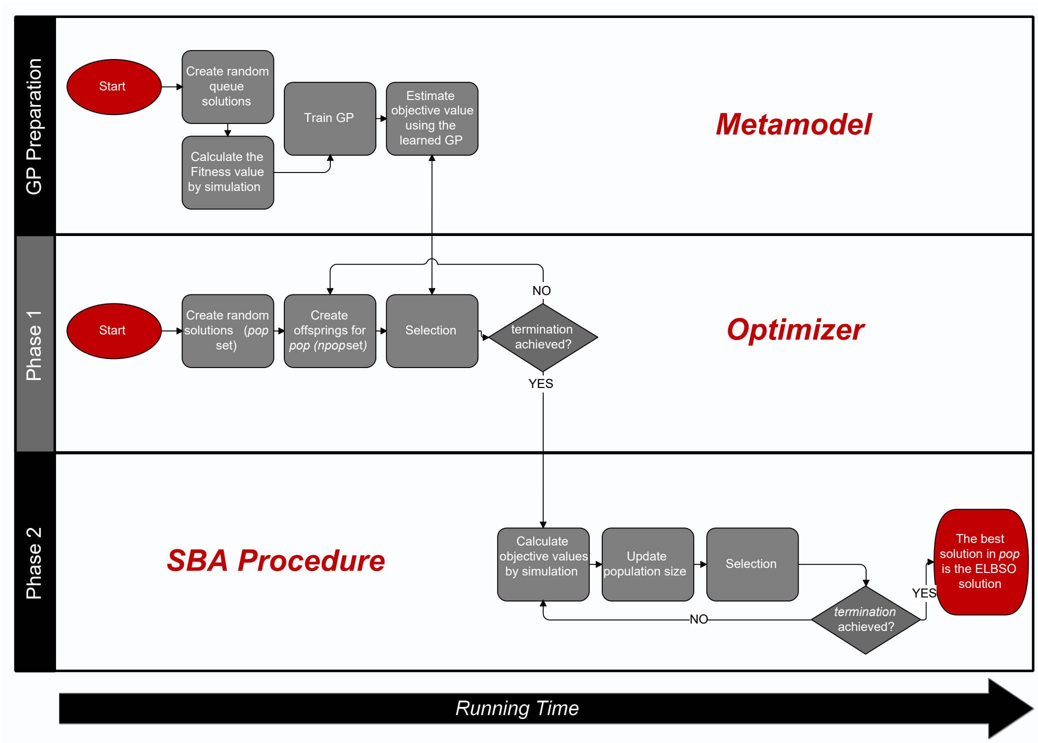

Figure 1: The general framework of the ELBSO algorithm (Ghasemi et al. 2021)

.

To illustrate the experimentation procedure in this research, Table 2 shows an example of an SJSSP considered. In this research, we consider 10 jobs and 10 operations for each job to be processed on 10 photolithography machines. As discussed in Section 3, the processing time of each operation follows the $Gamma[16.98, 2.448]$ distribution. To provide a targeted experimentation enabling us to provide a useful and unique sensitivity analysis on SJSSP, we solve different SJSSPs, which are essentially different in terms of job due dates and operation routes through machines. As shown in Table 2, each job has a specified due date and route between machines. For instance, the first operation of job two should be processed on the second machine (M_2). Consider R_{mj}, as the operation of job j assigned to machine m. Thus, FR_m is calculated as follows:

$$FR_m = \frac{\sum_{j=1}^{N} R_{mj}}{N} \qquad \forall m \in NM \tag{11}$$

where, FR_m defines the *mean* of operation numbers assigned to machine m. Accordingly, FRS_m provides to the standard deviation of operations assigned to machine m as follows:

$$FRS_m = \sqrt{\frac{(\sum_{j=1}^{N} R_{mj} - FR_m)^2}{N-1}} \qquad \forall m \in NM \tag{12}$$

Another metric that we consider in this research is Total Flexibility Ratio (TFR) of the operation numbers assigned to each machine in a particular SJSSP, which is equal to:

$$TFR = \sum_{m=1}^{NM} \left| FR_m - \frac{\sum_{i=1}^{N} i}{N} \right| \tag{13}$$

Table 2: Problems Environment Example with $TFR = 14.6$, $DDM = 68$, and $DDS = 33.2$.

Job id	Due Date	Operations Route									
		M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_{10}
1	30	1	2	4	3	7	8	9	6	5	10
2	40	2	1	3	4	7	8	9	6	5	10
3	25	4	1	3	2	7	8	9	10	5	6
4	55	6	1	3	9	7	8	2	10	5	4
5	75	7	8	3	9	6	1	2	10	5	4
6	105	2	8	3	1	6	9	7	10	5	4
7	95	6	5	3	1	2	9	7	10	8	4
8	35	1	5	4	6	2	9	7	10	8	3
9	105	4	3	1	7	2	9	6	10	8	5
10	115	2	8	1	5	4	9	6	10	3	7
FR_m		3.5	4.2	2.8	4.7	5	7.8	6.4	9.2	5.7	5.7
FRS_m		2.1	2.8	0.9	2.8	2.1	2.3	2.4	1.6	1.6	2.4

To illustrate this equation further, we use Table 2. Accordingly, if operations are assigned to machines where each machine will process a unique operation from a job then $TFR = 0$, i.e., FR_m will equal $(\sum_{i=1}^{N} i)/N$ in all cases. Using Equation 11, 12, and 13, FR_m, FRS_m, and TFR values are calculated, respectively, with calculations of the Due Date Mean (DDM) and Standard Deviation (DDS) of the job due dates. To make it clearer, consider the first machine (M_1) in Table 2, $FR_1 = 3.5$. However, if all operation numbers from one to 10 from different jobs were assigned to M_1, then the FR_1 value would equal 5.5 (i.e., $((1+2+3+\ldots+10)/10 = 5.5)$). Instances exist, where for the example given in Table 2 $FR_m = 5.5$, for example, where half the operations equal five and the other half equal six. In this case, FRS_m for the same machine will show this deviation of operation numbers on a machine. TFR will give information on the FR_m deviation of machines from 5.5. This indicates the total machines flexibility level in processing different operation numbers of jobs in a SJSSP instance.

Table 3 describes nine generated SJSSPs and experiment results in this research. The nine problems are divided into three categories based on their TFR levels. The mean due dates ($DDM = 500$) equate across all none problems, while examining the standard deviation of the due dates (DDS) one can see that they vary. Table 3 shows the result of solving the defined SJSSPs. That includes both the mean and standard deviation of the SJSSP objective value (F) in 10 replications for each problem. Full details for Problem 3 (see Table 4) and the flow time results (see Table 5) from Table 3 are given in Appendix A. Note that the full details of the SJSSP instances and results in this research are available at the following link: Data Sets and Results. As stated earlier the ELBSO ran for 2000 CPU seconds.

6 DISCUSSION

As it is clear from Table 3, the best results obtained for the TFR *Category* 1 job shop environment with the $Mean(F)$ between 807.4 and 2151.4. On the other hand, increasing the due dates variability in each group caused a sharp increase in *Mean* of the objective value. To analyze further the results, we used a Taguchi Design Of Experiment (DOE) method using Minitab application (Minitab 2021). This method is based on a special set of arrays called orthogonal arrays to conduct the minimum number of efficient experiments that could give insights on all factors that affect the performance measure. After executing the test, the main effects plots for both TFR and DDS are shown in Figure 2, where 1 corresponds to TFR *Category* 1, etc and for DDS 1 corresponds to DDS *Category* 1, etc. Figure 2 shows how each factor affects the response characteristics (Signal Noise (SN) ratio, means, slopes, and standard deviations) (here minimizing the total cost objective). A main effect exists when different levels of a factor affect the characteristic differently.

Table 3: SJSSP instances in this research and experiments results.

Machine id	Metric	Problems								
		TFR Category 1			*TFR Category* 2			*TFR Category* 3		
		1	2	3	4	5	6	7	8	9
1	*FR*	5.5	5.5	5.5	7.7	7.7	7.7	1	1	1
	FRS	2.8	2.8	2.8	2.5	2.5	2.5	0	0	0
2	*FR*	5.5	5.5	5.5	5.4	5.4	5.4	2	2	2
	FRS	2.8	2.8	2.8	2.7	2.7	2.7	0	0	0
3	*FR*	5.5	5.5	5.5	4.7	4.7	4.7	3	3	3
	FRS	2.8	2.8	2.8	1.8	1.8	1.8	0	0	0
4	*FR*	5.5	5.5	5.5	4.9	4.9	4.9	4	4	4
	FRS	2.8	2.8	2.8	2.6	2.6	2.6	0	0	0
5	*FR*	5.5	5.5	5.5	5.9	5.9	5.9	5	5	5
	FRS	2.8	2.8	2.8	3	3	3	0	0	0
6	*FR*	5.5	5.5	5.5	5.4	5.4	5.4	6	6	6
	FRS	2.8	2.8	2.8	2.6	2.6	2.6	0	0	0
7	*FR*	5.5	5.5	5.5	5.2	5.2	5.2	7	7	7
	FRS	2.8	2.8	2.8	3.1	3.1	3.1	0	0	0
8	*FR*	5.5	5.5	5.5	5.8	5.8	5.8	8	8	8
	FRS	2.8	2.8	2.8	2.7	2.7	2.7	0	0	0
9	*FR*	5.5	5.5	5.5	5.6	5.6	5.6	9	9	9
	FRS	2.8	2.8	2.8	2.8	2.8	2.8	0	0	0
10	*FR*	5.5	5.5	5.5	4.4	4.4	4.4	10	10	10
	FRS	2.8	2.8	2.8	3.3	3.3	3.3	0	0	0
TFR		0	0	0	6	6	6	25	25	25
DDM		500	500	500	500	500	500	500	500	500
DDS		0	207.3	317.6	0	207.3	317.6	0	207.3	317.6
$Mean(F)$		807.4	1407.2	2151.4	1840.4	2123.3	2711.7	2250.4	2510.1	3244.2
$STD(F)$		137.7	28.4	132.4	97.4	31.9	67.9	202.2	265.6	607.7

For instance, for a factor with two levels, you may discover that one level increases the mean compared to the other level. This difference is a main effect. Accordingly, as it is shown, the best results are achieved when both *TFR* and *DDS* factors are in their first state indicating that *TFR Category* 1 and *DDS Category* 1 are the best job mix strategy in our research. That indicates the best job shop performance occurs when the job mix in each shift is selected in a way that all machines process different operation numbers of jobs and due dates of jobs are as equal as possible. Considering photolithography workstations, it means that, considering operational constraints such as machine dedication, the production planner should select the wafer mix in the way that machines could process different layers from different wafers, and job due dates are as equal as possible. Indeed, our findings are clearly against line production, which refers to processing each layer of wafers on a certain machine.

7 CONCLUSIONS

In this paper, we examined the impact of job mix selection on the performance of the job shop in a production shift. That is one of the main concerns within photolithography workstations in semiconductor manufacturing systems. Accordingly, job due dates and routes between machines are considered as the influencing factors. Thus, different sets of SJSSP using a photolithography case data were proposed. To solve the designed SJSSPs, we used an SO method named ELBSO proposed by Ghasemi et al. (2021). While the analysis had many limitations in that it did not model machine process capabilities, machine dedication constraints, or reticle constraints, and examined a workstation rather than the complete fab, it does provide some guidance in how planners could select jobs for a shift. However, experiment results show that the best job shop performance occurs when each machine in a production shift produces different layers of wafers, and the assigned wafers to a shift have as equal due dates as possible. As an extension to this research, we would like to examine the impact of mask availability constraints on the job mix selection

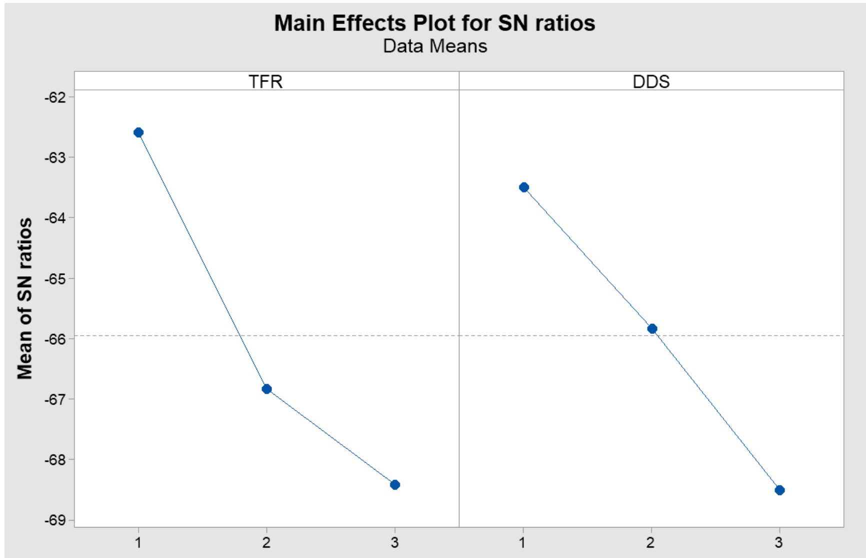

Figure 2: Results main effects plot for the SN ratios.

and qualification of machines. Moreover, we would like to modify ELBSO by using innovative heuristics such as the ones proposed by Beheshtinia and Ghasemi (2018), Beheshtinia et al. (2017).

REFERENCES

Beheshtinia, M. A., and A. Ghasemi. 2018, September. "A multi-objective and integrated model for supply chain scheduling optimization in a multi-site manufacturing system". *Engineering Optimization* 50(9):1415–1433.

Beheshtinia, M. A., A. Ghasemi, and M. Farokhnia. 2017, December. "Supply chain scheduling and routing in multi-site manufacturing system (case study: a drug manufacturing company)". *Journal of Modelling in Management* 13(1):27–49.

Can, B., and C. Heavey. 2012, February. "A comparison of genetic programming and artificial neural networks in metamodeling of discrete-event simulation models". *Computers & Operations Research* 39(2):424–436.

Dunke, F., and S. Nickel. 2020, February. "Neural networks for the metamodeling of simulation models with online decision making". *Simulation Modelling Practice and Theory* 99:102016.

Figueira, G., and B. Almada-Lobo. 2014. "Hybrid simulation-optimization methods: A taxonomy and discussion". *Simulation Modelling Practice and Theory* 46:118–134.

Ghasemi, A., A. Ashoori, and C. Heavey. 2021, July. "Evolutionary Learning Based Simulation Optimization for Stochastic Job Shop Scheduling Problems". *Applied Soft Computing* 106:107309.

Ghasemi, A., R. Azzouz, G. Laipple, K. E. Kabak, and C. Heavey. 2020, January. "Optimizing capacity allocation in semiconductor manufacturing photolithography area – Case study: Robert Bosch". *Journal of Manufacturing Systems* 54:123–137.

Graham, R. L., E. L. Lawler, J. K. Lenstra, and A. H. G. R. Kan. 1979, January. "Optimization and Approximation in Deterministic Sequencing and Scheduling: a Survey". In *Annals of Discrete Mathematics*, edited by P. L. Hammer, E. L. Johnson, and B. H. Korte, Volume 5 of *Discrete Optimization II*, 287–326. Elsevier.

Horng, S.-C., S.-S. Lin, and F.-Y. Yang. 2012, February. "Evolutionary algorithm for stochastic job shop scheduling with random processing time". *Expert Systems with Applications* 39(3):3603–3610.

Horng, S.-C., F.-Y. Yang, and S.-S. Lin. 2012, August. "Embedding evolutionary strategy in ordinal optimization for hard optimization problems". *Applied Mathematical Modelling* 36(8):3753–3763.

Hussain, M. F., R. R. Barton, and S. B. Joshi. 2002, April. "Metamodeling: Radial basis functions, versus polynomials". *European Journal of Operational Research* 138(1):142–154.

Jamili, A. 2019, February. "Job shop scheduling with consideration of floating breaking times under uncertainty". *Engineering Applications of Artificial Intelligence* 78:28–36.

Kleijnen, J. P. C. 2009, February. "Kriging metamodeling in simulation: A review". *European Journal of Operational Research* 192(3):707–716.

Minitab 2021. "Data Analysis, Statistical & Process Improvement Tools".

Mönch, L., R. Uzsoy, and J. W. Fowler. 2018, July. "A survey of semiconductor supply chain models part III: master planning, production planning, and demand fulfilment". *International Journal of Production Research* 56(13):4565–4584.

Pinedo, M. 2016. *Scheduling: theory, algorithms, and systems*. Fifth Edition ed. Cham Heidelberg New York Dordrecht London: Springer. OCLC: 945375528.

Shen, J., and Y. Zhu. 2016, June. "Chance-constrained model for uncertain job shop scheduling problem". *Soft Computing* 20(6):2383–2391.

Trigueiro de Sousa Junior, W., J. A. Barra Montevechi, R. de Carvalho Miranda, and A. Teberga Campos. 2019, February. "Discrete simulation-based optimization methods for industrial engineering problems: A systematic literature review". *Computers & Industrial Engineering* 128:526–540.

Winands, E., I. Adan, and G. van Houtum. 2011, April. "The stochastic economic lot scheduling problem: A survey". *European Journal of Operational Research* 210(1):1–9.

Yang, H.-a., Y. Lv, C. Xia, S. Sun, and H. Wang. 2014. "Optimal Computing Budget Allocation for Ordinal Optimization in Solving Stochastic Job Shop Scheduling Problems". *Mathematical Problems in Engineering* 2014:1–10.

Zhang, J., G. Ding, Y. Zou, S. Qin, and J. Fu. 2019, April. "Review of job shop scheduling research and its new perspectives under Industry 4.0". *Journal of Intelligent Manufacturing* 30(4):1809–1830.

AUTHOR BIOGRAPHIES

AMIR GHASEMI is a Postdoctoral Researcher in the School of Engineering at the University of Limerick. He published papers in the field of Simulation, Optimization, and Machine Learning-based Decision Support Tools for operations, transportation, and supply chain planning. His research interests include designing Simulation, Optimization, and Machine Learning-based Smart Agents in order to replace and/or support the human in decision making applied to both manufacturing and service sectors. His email address is: Amir.ghasemi@ul.ie.

CATHAL HEAVEY is an Associate Professor in the School of Engineering at the University of Limerick. He is an Industrial Engineering graduate of the National University of Ireland (University College Galway) and holds an M. Eng.Sc. and Ph.D. from the same University. He has published in the areas of queuing and simulation modeling. His research interests include simulation modeling of discrete-event systems; modeling and analysis of supply chains and manufacturing systems; process modeling; and decision support systems. His email address is Cathal.Heavey@ul.ie.

A Problem Data

Table 4: Problem 3 for job mix selection.

		Machines										
		1	2	3	4	5	6	7	8	9	10	Due Date
Jobs	1	1	2	3	4	5	6	7	8	9	10	60
	2	10	1	2	3	4	5	6	7	8	9	160
	3	9	10	1	2	3	4	5	6	7	8	220
	4	8	9	10	1	2	3	4	5	6	7	310
	5	7	8	9	10	1	2	3	4	5	6	420
	6	6	7	8	9	10	1	2	3	4	5	530
	7	5	6	7	8	9	10	1	2	3	4	640
	8	4	5	6	7	8	9	10	1	2	3	700
	9	3	4	5	6	7	8	9	10	1	2	810
	10	2	3	4	5	6	7	8	9	10	1	1150
	FR	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	500
	FRS	2.87	2.87	2.87	2.87	2.87	2.87	2.87	2.87	2.87	2.87	

Table 5: Total cost results.

		Replication											
		1	2	3	4	5	6	7	8	9	10	mean(F)	STD(F)
Problems	1	1021	614	715	776	1015	752	780	653	968	780	807.4	137.79
	2	1366	1434	1432	1433	1371	1439	1431	1405	1374	1387	1407.2	28.46
	3	1946	2071	2001	2319	2199	2314	2191	2052	2331	2090	2151.4	132.49
	4	1751	1672	1951	1950	1815	1722	1941	1816	1934	1852	1840.4	97.40
	5	2118	2172	2096	2139	2072	2141	2132	2143	2071	2149	2123.3	31.93
	6	2823	2677	2759	2573	2677	2758	2731	2691	2655	2773	2711.7	67.96
	7	2356	1950	1966	2543	2338	2264	2508	2002	2254	2323	2250.4	202.26
	8	2782	2553	1967	2776	2261	2537	2740	2253	2443	2789	2510.1	265.60
	9	2330	2995	4414	3077	3305	3716	2616	4018	3137	2834	3244.2	607.76