

USING SIMULATION AND HYBRID SEQUENCING OPTIMIZATION FOR MAKESPAN REDUCTION AT A WET TOOL

Anna Rotondo
John Geraghty
Paul Young

Enterprise Process Research Centre,
School of Mechanical and Manufacturing Engineering, Dublin City University
Dublin 9, Ireland

ABSTRACT

When rigid scheduling rules apply to wet tools, the development of Cycle Time (CT) optimization strategies becomes a relevant challenge. The impact of sequencing optimization on makespan performance at a wet tool is investigated here by means of a hybrid optimization model that combines an exact optimization approach, based on an efficient permutation concept, and a heuristics, based on Genetic Algorithms (GAs). The model also includes a scheduling module that reproduces the control logics governing wet tools operating in a real semiconductor manufacturing plant and proves effective in generating efficient and detailed schedules in short computational times. The realistic assumptions on which the scheduling module is based allow the simulation of different tool configurations. The results obtained show that significant makespan reductions can be achieved by means of a mere sequencing optimization as parallel processing within the wet tools is better exploited.

1 INTRODUCTION

In manufacturing systems characterized by complex interactions between their components, small variations of operating factors, implemented at any production step, may have a significant impact on the overall system performance. As a result, Cycle Time (CT) reductions obtained at critical production steps could generate considerable productivity improvements and eventually lead to a capacity increase at no investment cost (Quek et al. 2007; Aydt et al. 2008). Hence, CT improvements also become strategic targets for companies that want to maintain competitive advantages, especially when they operate in highly dynamic industries, such as the semiconductor industry (Kuo, Chien, and Chen 2011).

As the overall cleaning process constitutes almost 10% of the operations in a semiconductor wafer manufacturing plant (Kabak, Heavey, and Corbett 2010), wet stations certainly represent critical production steps for the semiconductor wafer manufacturing process. Automated wet tools can be classified as batch chamber tools; they include several chambers, or tanks, each of which can accommodate a batch of wafers, usually made of one or two lots. Due to the inherent complexities and the peculiar scheduling constraints, simulation approaches are usually preferred to model wet tools (Govind and Fronckowiak 2003; Noack et al. 2008; Aydt et al. 2008); simulation models support investigations on the impact of operational settings variations on wet tools performance (Noack et al. 2008). Using an optimization framework for the metallization process, Noack et al. (2008) optimize dispatching rules and virtual queue capacity in front of a wet tool. The effects of various dispatching rules and recipe dedication schemes on wet stations capacity is analyzed by Quek et al. (2007). Recipe dedication optimization at wet tools is also investigated by Aydt et al. (2008); a symbiotic simulation control system is created to dynamically adjust the recipe dedication settings at the available tools with respect to the current WIP level. Govind and

Fronckowiak (2003) investigate the impact of scheduling policies, maximum waiting time for batching operations and maximum internal buffer level on CT at wet tools by means of a hybrid simulation model that combines transient entity modeling and a resident-entity approach. The results obtained suggest that the toolset configuration significantly affects the relationship between the factors investigated and CT. The effects of parallel processing within a wet tool on CT and throughput are analyzed using a flexible simulation model by Mauer and Schelasin (1993). The results reveal how the asynchronous batch processing in integrated tools decouples the inverse proportionality between CT and throughput and, as a result, higher CTs, due to parallel processing, can prove beneficial in terms of throughput and overall run time. The model developed by Mauer and Schelasin (1993) is conceived as a flexible model which can be easily adapted to mimic the behavior of any type of integrated tools, such as cluster tools. Cluster tools are substantially different from wet tools; however, analogies can be made between the two classes of tools. Studies on the effects on CT of parallel processing (Niedermayer and Rose 2003), tools configurations (Koehler et al. 1999; LeBaron and Hendrickson 2000) and sequencing strategies (Dümmmler 1999) conducted for cluster tools can provide useful information on scheduling strategies when analyzing wet tools performances.

Due to the nature of the semiconductor manufacturing process, scheduling problems at wet tools are often combined with scheduling problems at furnaces. Indeed, wet tools operations, such as wafer cleaning or etching, are usually followed by furnace operations for diffusion processes (Gan et al. 2006). The integrated scheduling problem is characterized by relevant complexities due to the different maximum batch sizes at the two batch tools and the possible presence of wait time constraints between the two processes (Scholl and Domaschke 2000). Efficient schedules for the wet etch operations are fundamental to ensure good performance for the furnace operations (Ham and Fowler 2008). Simulation and heuristics are generally used to determine optimal dispatching rules (Scholl and Domaschke 2000; Ham and Fowler 2008) and analyzing the effects of furnace upgrades on the upstream flow (Gan et al. 2006; Yugma et al. 2012).

As several operations are usually performed at a wet tool and parallel processing is allowed, considerable CT savings can be obtained by means of batch sequencing optimization. Mathematical programming approaches and heuristics have been extensively used for solving sequencing and scheduling optimization problems at wet stations. For these problems, as a result of management's suggestions, makespan minimization represents the most common objective function (Geiger, Kempf, and Uzsoy 1997; Zeballos, Castro, and Méndez 2010). Indeed, if the makespan decreases the throughput increases and this avoids that wet tools could constitute a constraint to the factory output (Geiger, Kempf, and Uzsoy 1997). Reducing the makespan also supports an increase in tool capacity and, hence, minimizes the number of tools needed, with obvious advantages in terms of occupied clean room floor space (Geiger, Kempf, and Uzsoy 1997). Moreover, decreasing the makespan leads to a lower inventory and contamination and results in greater profits (Bhushan and Karimi 2003). The observed research trend on wet stations sequencing and scheduling optimization is towards the development of approaches able to deliver nearly optimal solutions in a reasonable time for increasingly larger sized problems. Less rigid assumptions, especially those regarding the material handling system, are also considered in more recent models (Aguirre, Méndez, and Castro 2011; Lee, Lee, and Lee 2007). Tabu search is used by Geiger, Kempf, and Uzsoy (1997); whereas Bhushan and Karimi (2004) develop a simulated annealing based approach to address larger instances of the scheduling and sequencing problem at wet stations with one robot. Bhushan and Karimi (2003) also adopt a Mixed Integer Linear Programming (MILP) approach to generate optimal schedules at a wet station; based on the MILP approach, a two-step heuristic approach is developed. A reformulation of their MILP model is available (Karimi, Tan, and Bhushan 2004). A MILP approach was also recently used by Aguirre, Méndez, and Castro (2011) to model the wet station scheduling problem with several robots; the model proved more efficient than the one developed by Bhushan and Karimi (2003) for solving medium-sized problems. As a more efficient alternative to the MILP approach, (Zeballos, Castro, and Méndez 2011) suggest recourse to Constraint Programming (CP). Finally, Castro et al.

(2011) develop an integrated model, based on MILP and discrete event simulation, for solving large-scale wet station scheduling problems in reasonably short time.

In the studies reviewed here, the sequencing and scheduling problems are generally analyzed as an integrated problem; sequencing optimization affects the scheduling logics and the schedules generated by the models are based on the assumption that a full control of batch transfers within a wet tool is possible. However, this is not always the case in real manufacturing plants as wet tools internal mechanisms are established by the tool vendor and are not easy to modify on the plant. This forces manufacturing engineers to focus on makespan optimization strategies that prescind from variations of the scheduling logics. Following an approach similar to the one adopted by Dümmler (1999) for the optimization of lot sequences at cluster tools, batch sequencing and scheduling are analyzed here as a disjoint problem. Based on the internal workings of a real wet tool, a quasi-analytical scheduling model is developed and coupled with a hybrid sequencing optimization module. The scheduling algorithm generates detailed schedules compatible with the constraints applied; the scheduling logic used proves effective in producing efficient schedules, independent of the sequencing optimization module. The sequencing algorithm rearranges sequences of batches waiting to be processed at a wet tool in order to find the optimal or sub-optimal sequence, that is the sequence that minimizes the makespan; the optimization algorithm consists of a combination of an exact and an approximated approach as full enumeration and GAs are used according to the problem complexity. A sensitivity analysis on the efficacy of GA operators, for selection, mutation and cross-over is also performed. Apart from the nature of the tools analyzed, the problem presented here differs from the problem discussed in Dümmler (1999) as sequencing optimization is performed at a tool level rather than at a station level. The choice of conducting sequencing optimization at a tool level has been made to avoid the observation of misleading results due to the presence of ineffective assignment strategies.

This paper is structured as follows. In section 2, the system modeled in this paper is described and the problem analyzed is defined. Section 3 illustrates the scheduling and sequencing modules of the model developed. The model efficacy in generating optimal results is discussed in Section 4. Finally, conclusions are drawn in Section 5.

2 SYSTEM DESCRIPTION

The wet tool modeled in this study operates in a real wafer manufacturing plant; its schematic layout is shown in Figure 1.

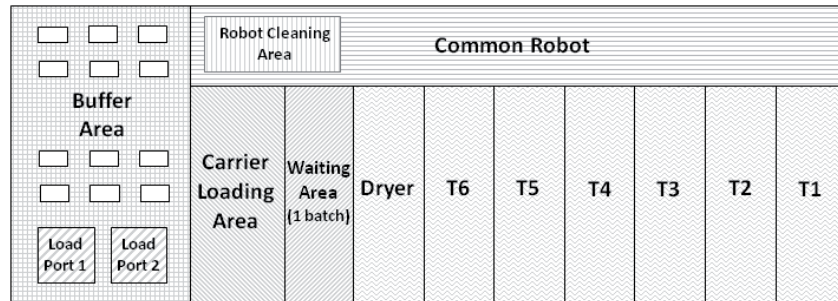


Figure 1: Schematic representation of the wet tool modeled

Conceptually, a wet tool can be considered as a series of tanks (T1-T6) filled with either etchant chemicals, used to etch away the exposed photo-resist from the wafer layers, or deionized water, used to terminate the etchant action. The succession of chemical and water tanks is not necessarily alternating as reported in most research studies (Geiger, Kempf, and Uzsoy 1997; Bhushan and Karimi 2004; Zeballos, Castro, and Méndez 2011); in real plants, consecutive chemical tanks can be observed. For some chemical tanks, overexposure to the etchant will damage the wafers exposed; as a consequence, a Zero-Wait (ZW) constraint is applied; this suggests that the batch must be moved out of the “unsafe” tank when the prescribed exposure time is reached. The remaining tanks will be considered “safe” tanks as overexposure to

the chemicals will not damage the wafers; these tanks can be used as local storage (LS) for a sensible time interval that varies according to the chemicals. A “No-Intermediate Storage” (NIS) constraint also applies to the wet tool modeled as no intermediate buffer is available between any two consecutive tanks. As a consequence, before a batch can leave a tank, the following tank in the corresponding recipe has to be available. A recipe describes the sequence of tanks that a lot should visit; an operation defines the processing times in each tank.

Wafers are processed in lots of the same type. Moreover, two lots waiting in the tool internal buffer and requiring the same operation can be batched together and loaded through a set of mechanical devices onto a carrier. In this paper, the terms batch and carrier will be used indistinctively. For each batch, the internal scheduler will generate a detailed schedule which lists the timestamps in and out of each tank according to the recipe assigned to the batch, the tanks and robot availability. The carrier is moved by the common robot to the first tank prescribed by the recipe or, if this is unavailable, the carrier will be placed in the temporary waiting area, which can be occupied exclusively by unprocessed batches. The handling system will move the carrier across the tanks following the associated schedule, which meets the scheduling constraints illustrated above. Each tank is provided with an internal handler that transfers a carrier from the common robot into the tank and holds the carrier in the tank until the process is complete. After leaving the last tank, any carrier, independent of the recipe, visits the dryer before being unloaded by the mechanical devices that also unbatch the lots. The processed lots, one or two according to the batch size, return into the internal buffer where they are ready to exit the tool through a load port.

In the scheduling problem discussed here, the following assumptions are made:

- The recipes considered do not consist of visiting all the tanks in the same order. This means that the batches can enter the different tanks in a different order. This is known as a “non-permutation” schedule (Karuno and Nagamochi 2003);
- A tank can process one batch at a time;
- The processing/transfer of batches is generally non-pre-emptive; the only exception is made for transfers out of the dryer that are prioritised;
- The robot can move one batch at a time and cannot be used as a temporary storage for any batch;
- The robot is failure free; the tool is subjected to failures and regular preventive maintenance; repouring operations also limit the tanks availability;
- Transfer times within the tool are dependent on the origin and destination;
- Setups are needed for both robot and tanks; indeed, cleaning operations for the robot arms are required post every transfer involving chemical tanks.

The objective of this scheduling problem is to minimize the makespan by means of strategies that require control changes easy to implement in a real manufacturing environment. The approach developed should also generate fully detailed schedules for the batches to be processed; the schedules will include the start and finish times of each batch in the different tanks and a record of the robot movements. Finally, the problem analyzed here consists of a non-permutation scheduling problem with a shared resource, that is the robot, and subjected to the ZW and NIS/LS constraints (Bhushan and Karimi 2004).

3 SIMULATION MODEL DEVELOPMENT

In order to investigate the impact of batch sequencing optimization on makespan at a wet tool, a simulation model of the tool has been developed. The model presents a modular structure that provides high flexibility and adaptability to the objectives of the analyses for which it is used. Two modules are fundamental for the analysis presented here: the scheduling module and the sequencing optimization module.

3.1 Scheduling Algorithm

The scheduling module generates detailed schedules for the batches to be processed at the wet tool based on deterministic input data. Minimum processing times, internal transfer times, which vary according to the origin and destination, and setup times used in the model are deterministic. Repouring operations in

the chemical tanks and preventive maintenance are also modeled following the schedules implemented in the real plant where the modeled tool operates. The control logics implemented in the internal scheduler of the tool modeled, which is described in the previous section, is proprietary to the tool manufacturer and cannot be easily modified in the plant. In the scheduling module, the actual scheduling logics is mimicked by an algorithm based on a trial and error approach that uses fundamental equations to calculate feasible schedules compatible with both scheduling constraints and robot and tanks availability. The equations used are not reported in this paper as their illustration goes beyond the analysis purpose. The trial and error algorithm developed can be considered a conceptual generalization of the JAT algorithm proposed by Bhushan and Karimi (2004). The two algorithms take the ZW and NIS constraints into account while generating a batch schedule; they also consider the robot availability constraint before proceeding with the calculation of the starting time for a batch at the next tank. However, fundamental differences between the two algorithms exist and refer to the complexity of the production scenarios modeled. Different recipes can be considered in the trial and error algorithm and, as a consequence, the FIFO logic used in the JAT algorithm is no more applicable at a tank level; in other words, a non-permutation scheduling problem as opposed to a permutation scheduling problem (Bhushan and Karimi 2004) is analyzed here. Another difference between the two algorithms consists of the consecution between unsafe tanks and safe tanks; in the algorithm developed here irregular successions of tanks, such as two following safe tanks, can be modeled.

```

for i = 1 -> no. batches
    j = first tank to be visited
    while j ≠ dryer
        Calculate  $TS_j$  and  $TF_j$ 
        if robot is available between  $TF_{j-1}$  and  $TS_j$ 
            if tank j is available between  $TS_j$  and  $TF_j$ 
                Update  $\Phi_j$  and  $\Phi_R$ 
                j = j+1
            else Calculate min  $TS_j^*$  at which j is available and restart from
                last safe tank (j = j-x)
            end
        else Calculate min  $TF_{j-1}^*$  at which robot is available and restart from
            last safe tank (j = j-x)
        end
    end
    i = i+1
end

```

Figure 2: Pseudo-code of the scheduling algorithm

The algorithm works as per the pseudo-code above (Figure 2); it considers one batch at a time and completes its schedule before considering the following batch in the sequence. While the schedule for a batch is progressively calculated, the algorithm keeps a record of both the robot availability, Φ_R , and the tanks availability, Φ_M , in order to avoid that more than one batch utilizes the same resource at the same time. The record consists of an ordered sequence of arrays that contain the start time, TS , and the finish time, TF , during which a resource (e.g. robot or tank) is utilized. Generally, TS and TF refer to the start and finish time in a tank. For each tank to be visited, the algorithm starts with checking the robot availability for a time interval that allows the robot to reach the tank, $j-1$, where the batch k is currently kept, at the calculated $TF_{k(j-1)}$, complete its cleaning operation, where required, and be available for the next transfer previously scheduled without any delay. If the robot proves available, Φ_R is updated by inserting an

appropriate time interval that includes $[TF_{k(j-1)}; TS_{kj}]$ so that the chronological order is respected; a similar record is kept for the robot positions corresponding to the relevant timestamps reported in Φ_R . Hence, the tank availability is checked for the time interval $[TS_{kj}; TF_{kj}]$. In case either the robot or the tank are not available, the smallest time at which the unavailable resource becomes available for the required time interval is calculated and, based on that, the start and finish times of the previous unsafe tanks are derived. The iteration restarts from the first safe tank that precedes the tank currently assessed. This is because if the newly calculated TF_{kj^*} for that safe tank, j^* , satisfies the constraint on the overexposure time and there exist a prolonged availability in that tank, the batch can wait in j^* for a longer time than the prescribed processing time and be transferred to the unsafe tanks at the newly calculated times. This will avoid re-starting the schedule calculation from the very first tank in the recipe and will reduce waiting times in the internal buffer. If the tank proves free, Φ_j , which is the Φ_M instance for tank j , is updated and the next tank in the recipe is considered. In case a resource's unavailability forces schedule recalculations, Φ_R and Φ_M are restored to their previous status.

3.2 Sequencing Optimization Algorithm

The sequencing optimization module generates different batch sequences that are assessed by means of the scheduling module; it selects the sequence with the minimum makespan. Contrary to the scheduling logics that is proprietary to the tool's manufacturer and cannot be easily modified, the sequence with which batches are processed by a tool can be controlled by the plant automation system. The possibility of paralleling processing within a wet tool and the difference between the recipes performed in terms of both number of tanks to be visited and succession of tanks constitute a fundamental theoretical base for the applicability of sequencing optimization strategies to the problem.

The sequencing optimization algorithm adopts a hybrid approach. According to the length of the sequence to be optimized and the number of non-identical operations to perform, the optimization approach will be either exact or approximated. The exact approach is based on an intelligent full enumeration approach and consists of assessing all the different sequences that can be generated from the original batch sequence. Relevant batch sequences are those which differ from the previous ones considered in terms of operation sequences. Indeed, two different batch sequences with the exact same operation sequence are perfectly equivalent from a scheduling perspective. This means that when two or more batches in the sequence require the same operation, the number of different sequences to be analyzed proves less than a mere permutation. It can be demonstrated that the number of non-identical permutations in a sequence containing n_i batches of operation i is given by the following equation

$$\# \text{ different sequences} = \frac{\left(\sum_i n_i \right)!}{\prod_i (n_i!)} \quad (1)$$

where the summation at the numerator consists of the sequence length and the denominator represents the number of identical sequences in the full enumeration. The enumeration method used modifies the standard permutation algorithm so that the generation of identical item sequences is prevented. For longer and more complex sequences, the sequencing optimization is based on a genetic algorithm (GA) approach. Using this approach the search space is significantly reduced; different sequences are generated based on the fitness of the parental solutions. A preliminary sensitivity analysis on the efficacy of several GA operators in generating optimal solutions was also conducted and the best performing operators for the problem analyzed here were chosen. The sensitivity analysis was performed on bigger problem sizes; sequences of up to 25 batches were considered; the sequencing algorithm performances proved robust with respect to the problem size. In this paper, small problems are considered as, in the production scenario modeled, longer batch sequences are not observed at a tool level. Two different operators for selection cross-over and mutation were compared and the best performing operators were determined.

The scheduling algorithm was verified by assessing the feasibility of the schedules generated; attention was paid to both the constraints imposed and the utilization of common resources, such as robot and tanks. All the schedules generated by the algorithm proved theoretically correct.

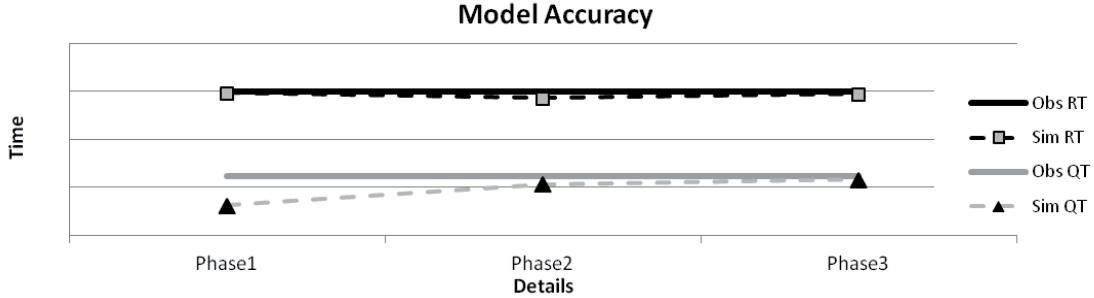


Figure 3: Model accuracy with respect to level of details included in the model

Being inspired by the behavior of real wet tools, the algorithm was validated by comparing the results generated against those observed in the real system. Average Queuing Times (QTs) and Run Times (RTs) were used as validation measures. QT is here intended as the time elapsing between the batch arrival at the buffer upstream the wet tool and the time at which the batch leaves the tool internal buffer; this means the waiting time within the tool is also included in QT. RT is intended as the time that a batch spends inside a tool, excluding the waiting time in the internal buffer. The sum of QT and RT is here called Cycle Time (CT). The calculation of QT required the development of a module for simulating the assignment strategy implemented in the real plant; the illustration of the strategy and the logics behind the assignment module are beyond the scope of this paper. Whereas the QT predictions needed some improvements and further details had to be included in the assignment module, the RT predictions proved very accurate throughout the different steps of the assignment module development (Figure 3). In Figure 3, Phase 1 refers to the initial model; in Phase 2, maintenance and repouring operations, setup times at each tank, maximum waiting times for batching operations and stochastic travel times from the upstream buffer to the tool were included. Finally, during Phase 3, the assignment strategy logic was revised to reproduce the actual one more accurately, variable limits to the internal buffer capacity were also introduced. Figure 3 shows the scheduling algorithm efficacy in reproducing the internal mechanisms of the modeled tool. The QT and RT prediction errors are always less than 0.7% and 1.5%, respectively. In order to evaluate the efficiency of the scheduling logic implemented in the model, the production scenario used in Bhushan and Karimi (2004) to illustrate the impact of rescheduling delayed tasks was simulated. The results obtained show that the scheduling algorithm developed here is able to generate initial solutions that might prove better than those reported in Bhushan and Karimi (2004); in particular, scheduling the 25 jobs reported in one of the examples required 519.6 time units using the Bhushan's algorithm against the 501.5 units estimated by the scheduling algorithm illustrated in this study, thus obtaining 3.5% makespan reduction before the implementation of sequencing optimization approaches. The scheduling algorithm has also shown good performance in terms of computational times as fully detailed batch schedules are generated in about 0.01 seconds. All the experiments were run on a CoreDuo 2.4 GHz processor.

4 RESULTS

The potential benefits deriving from optimizing the sequence of batches entering a wet tool were first investigated using the full enumeration approach. Sequences of 4 to 8 batches were considered; the combinatorial complexity of this problem prevented longer sequences from being investigated as the results would not be available in sensible times for practical implementations. For each sequence length, three different sequences with different operations to be performed were randomly generated (Table 1); the operations were chosen from a pool of 11 different operations (Table 2). These operations were grouped into

three recipes according to the tanks to be visited; recipes with five, three and two tanks were considered. T1 represents the only unsafe tank in the tool; T3 and T4 are complementary chemical tanks where a maximum overexposure of 13 minutes is allowed. T5 and T6 are the rinse tanks associated with T3 and T4; they can be use alternatively according to their availability. All transfer times were set equal to 5 seconds.

Table 1: Operation sequences randomly generated and corresponding min and max makespan

# batches	Operations								Min Makespan	Max Makespan
4	8	11	6	3					97.83	76.17
	10	10	3	4					95.83	74.67
	11	10	5	6					108.00	88.75
5	9	8	11	5	10				108.62	102.83
	7	1	9	2	4				108.62	86.03
	1	3	3	8	4				92.28	81.83
6	3	6	9	3	1	6			134.92	98.37
	11	3	7	3	4	1			132.92	94.33
	3	1	8	5	10	2			121.83	94.33
7	4	7	5	8	6	2	3		151.50	106.50
	11	11	9	2	9	2	7		146.40	114.20
	11	3	11	9	11	7	10		143.78	117.92
8	10	8	6	4	5	5	8	3	197.08	131.00
	2	5	9	10	4	8	3	11	172.95	122.33
	1	9	4	11	7	9	10	6	191.70	136.20

Table 2: Recipes and processing times for the simulated operations

Op.	Processing Times [mins]					
	T1	T2	T3	T4	T5/T6	Dryer
1			3	3	12	12
2			3.5	3.5	12	12
3			4	4	12	12
4	1	12	3	3	12	12
5	0.5	12	3	3	12	12
6	1.5	12	3.5	3.5	12	12
7	1	12	3.5	3.5	12	12
8	1.5	12				12
9	5.7	12				12
10	1	12				12
11	1	12				12

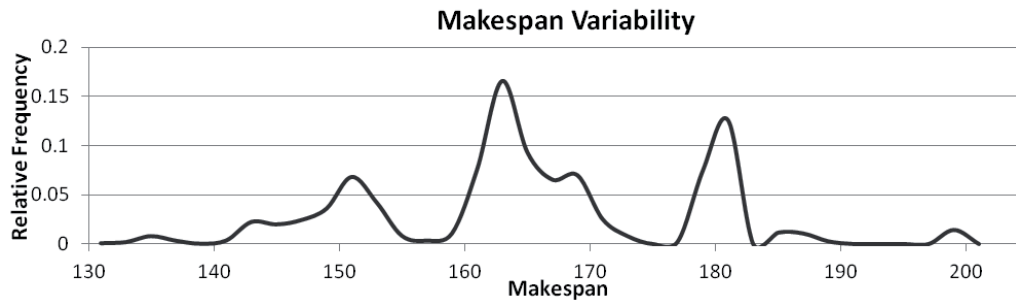


Figure 4: Makespan variations caused by batch sequencing variations

Based on the experiments conducted, a 25% makespan variation between the best and the worst schedules was observed, on average. For the longer sequences, the makespan range reached 33%; Figure 4 shows the makespan variability generated by different batch sequences for a 8 batch schedule. This suggests that considerable advantages can be gained by rearranging the sequence of batches to be processed at a tool without intervening on the scheduling logic and, hence, at no implementation costs.

The GA model efficacy in generating optimal results was assessed by comparing the results obtained using the GA approach against those generated by the full enumeration method. The GA model was successful in generating the absolute optimum for all the sequences analyzed. The GA operators and parameters were kept unchanged throughout the validation experiments; in particular, the population size was kept constant and equal to 20; hence, the success rate in finding the optimum reduced for the larger problem instances. The GA operators used in the validation experiments are the elitist selection, the order 2 crossover (Starkweather et al. 1991) and the displacement mutation. After running the model 10 times per sequence, an average success rate equal to 80% was observed for sequences shorter than 7 batches. A 30% success rate for 8 batch sequences suggested the need for increasing the population size for larger problem instances; additional experiments performed for the 8 batch sequences showed that a larger population size increases both the success rate and the convergence speed to the absolute optimum (Figure 5). For the non-successful optimization repetitions, the relative deviations of the GA solutions from the absolute optima were lower than 9% on average. For the worst case scenario, which consists of a 17% deviation from the optimum, a 12% makespan reduction was still obtained. The relative deviation, D_r , is calculated as follows

$$\Delta_r = \frac{GA\ solution - Opt}{Max - Opt} \quad (2)$$

where Max represents the maximum makespan observed in the full sequence enumeration.

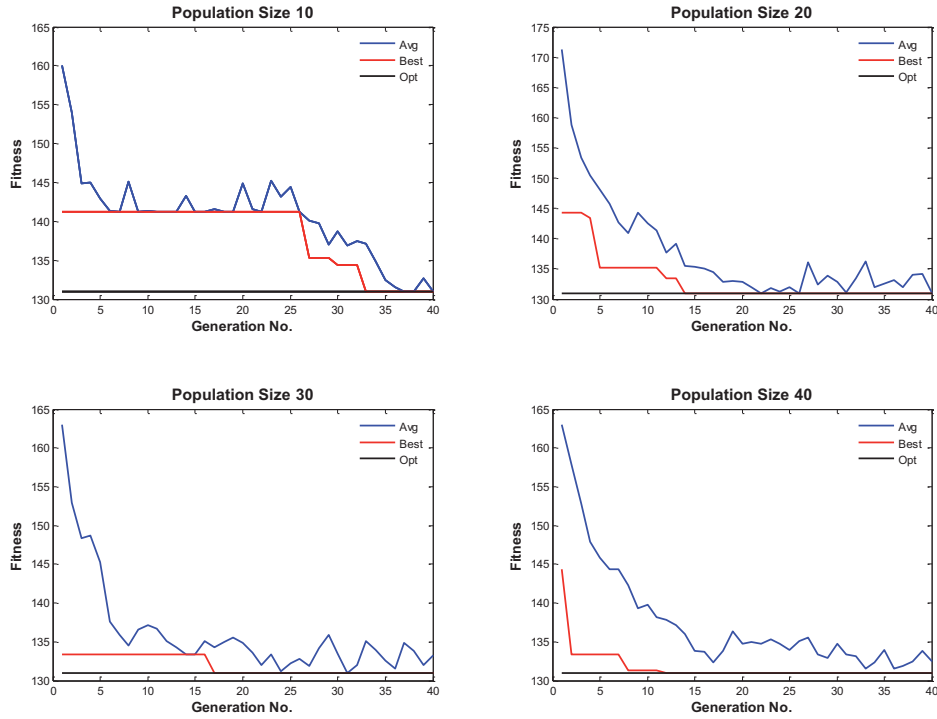


Figure 5: GA convergence for different population sizes (8 batch sequence).

It is worth noting that the analysis on GA performances and consequent decisions on GA parameters should take into account not only the sequence length. Indeed, the sequence complexity also depends on factors such as the number of non-repetitive permutations (Equation 1), the frequency of the optimal solution and the number of different solutions. These last two factors can be assessed only with a full enumeration of the sequence permutations and, hence, an a priori GA parameters' optimization proves difficult. The quality of the initial solution, measured in terms of relative deviation from the optimum, did not significantly impact the results; paradoxically, worse initial solutions seem to generate slightly better results.

Finally, keeping the GA parameters unchanged, the computational times for the GA approach slightly increased with the number of batches in the sequence as a result of the higher complexity of the scheduling operations (Table 3). The computational time savings with respect to the full enumeration method are significant and increase with the sequence length.

Table 3: Average computational times for different sequence lengths [seconds]

# batches	Full Enumeration	GA	# batches	Full Enumeration	GA
4	0.71	1.07	7	47.35	2.09
5	1.44	1.38	8	1,071.74	2.38
6	7.01	1.88			

5 CONCLUSIONS

The criticality of wet-etching in semiconductor wafer fabrication makes makespan optimization at wet tools a fundamental source of productivity improvements for the overall manufacturing system. Complex dynamics govern wet tools. The peculiar scheduling constraints applied suggest that, in order to be easy to implement, makespan optimization strategies should avoid variations of the scheduling logics. This is because of both the rigidity of fundamental scheduling rules and the frequent circumstance that software controlling the wet tools behavior is proprietary to the tools vendor, hence relevant variations in the control logics would prove expensive and not immediately implementable. The effects of sequencing optimization on makespan at a wet tool are investigated in this study by means of a simulation model. This model consists of a scheduling algorithm combined with a sequencing optimization module. The scheduling algorithm incorporates fundamental scheduling constraints and is inspired by a real wet-etch tool scheduling logic. Comparisons with scheduling algorithms proposed in the literature show that the trial-and-error scheduling algorithm is effective in generating efficient and detailed batch schedules; the higher efficiency of the algorithm could prove advantageous for the generation of good initial solutions for optimization problems that involve scheduling logics alterations. The sequencing optimization module, which couples intelligent permutation and GAs, intends to exploit the possibility of parallel processing that is a peculiar feature of wet tools. The order with which batches are introduced in a wet tool significantly impacts the time needed to process the batches waiting at the tool as the makespan range achieves 33% for the problem instances analyzed here.

Future research will focus on the development of alternative CT optimization strategies that do not interfere with the scheduling logics. As previous studies have demonstrated the potential benefits deriving from assignment strategy optimization, various assignment heuristics, easy to implement in a real system, will be developed and assessed using both the assignment module and the trial-and-error algorithm. The assignment module, that also governs virtual batching at the common buffer upstream the wet station, will also be used to optimize maximum waiting times per lot/operation before the assignment decision is made (See also Murray et al. 2008).

ACKNOWLEDGEMENTS

This work has been funded under the National Development Plan 2008-2013 by the Irish Research Council for Science Engineering and Technology and carried out in collaboration with the Enterprise Ireland & IDA sponsored Technology Centre, Irish Centre for Manufacturing Research.

REFERENCES

- Aguirre, A. M., C. A. Méndez, and P. M. Castro. 2011. A novel optimization method to automated wet-etch station scheduling in semiconductor manufacturing systems. *Computers & Chemical Engineering* 28:883-888.
- Aydt, H., S. Turner, W. Cai, M. Low, P. Lendermann, and B. Gan. 2008. Symbiotic simulation control in semiconductor manufacturing. In *Proceedings of Computational Science-ICCS 2008*: 26-35.
- Bhushan, S., and I. A. Karimi. 2003. An MILP approach to automated wet-etch station scheduling. *Industrial & Engineering Chemistry Research* 42(7):1391-9.
- Bhushan, S., and I. A. Karimi. 2004. Heuristic algorithms for scheduling an automated wet-etch station. *Computers & Chemical Engineering* 28(3):363-79.
- Castro, P. M., A. M. Aguirre, L. J. Zeballos, and C. A. Mendez. 2011. Hybrid mathematical programming discrete-event simulation approach for large-scale scheduling problems. *Industrial & Engineering Chemistry Research* 50(18): 10665-10680.
- Dümmler, M. A. 1999. Using simulation and genetic algorithms to improve cluster tool performance. In *Proceedings of the 1999 Winter Simulation Conference*, Edited by P.A.Farrington, H.B. Nembhard, D.T. Sturrock and G.W. Evans, 875-879. Phoenix, Az, USA, Inc.
- Gan, B. P., P. Lendermann, K. P. Te Quek, B. van der Heijden, C. C. Chin, and C. Y. Koh. 2006. Simulation analysis on the impact of furnace batch size increase in a deposition loop. In *Proceedings of the 38th Winter simulation Conference*, Edited by L.F. Perrone, F.P. Wieland, J. Liu, B.G. Lawson, D.M. Nicol and R.M. Fujimoto, 1821-1828. Monterey, California, USA, Inc.
- Geiger, C. D., K. G. Kempf, and R. Uzsoy. 1997. An improved MILP formulation for scheduling an automated wet-etch station. *Journal of Manufacturing Systems* 16(2):102-16.
- Govind, N., and D. Fronckowiak. 2003. Resident-entity based simulation of batch chamber tools in 300 mm semiconductor manufacturing. In *Proceedings of the 2003 Winter Simulation Conference*, Edited by S. Chick, P.J. Sanchez, D. Ferrin, D.J. Morrice, 1398-1405. New Orleans, Louisiana, USA, Inc. .
- Ham, M., and J. W. Fowler. 2008. Scheduling of wet etch and furnace operations with next arrival control heuristic. *The International Journal of Advanced Manufacturing Technology* 38(9):1006-17.
- Kabak, K. E., C. Heavey, and V. Corbett. 2010. Performance improvement for a wet bench tool. In *Proceedings of the 2010 Winter Simulation Conference*, Edited by B. Johansson, S. Jain, J. Montoya-Torres, J. Hugan and E. Yücesan, 2586-2593. Baltimore, Maryland, USA, Inc.
- Karimi, I. A., Z. Y. L. Tan, and S. Bhushan. 2004. An improved formulation for scheduling an automated wet-etch station. *Computers & Chemical Engineering* 29(1):217-24.
- Karuno, Y., and H. Nagamochi. 2003. A better approximation for the two-machine flowshop scheduling problem with time lags. *Algorithms and Computation* 309-18.
- Koehler, E. J., T. M. Wulf, A. C. Bruska, and M. S. Seppanen. 1999. Evaluation of cluster tool throughput for thin film head production. In *Proceedings of the 1999 Winter Simulation Conference*, Edited by P.A.Farrington, H.B. Nembhard, D.T. Sturrock and G.W. Evans, 714-719. Phoenix, Arizona, USA, Inc.
- Kuo, C. J., C. F. Chien, and C. D. Chen. 2011. Manufacturing intelligence to exploit the value of production and tool data to reduce cycle time. *IEEE Transactions on Automation Science and Engineering* 8(1):103-11.
- LeBaron, H. T., and R. A. Hendrickson. 2000. Using emulation to validate a cluster tool simulation model. In *Proceedings of the 2000 Winter Simulation Conference*, Edited by J.A. Joines, R.R. Barton, K. Kang and P.A. Fishwick, 1417-1422. Orlando, Florida, USA, Inc.
- Lee, T. E., H. Y. Lee, and S. J. Lee. 2007. Scheduling a wet station for wafer cleaning with multiple job flows and multiple wafer-handling robots. *International Journal of Production Research* 45(3):487-507.

- Mauer, J. L., and R. E. A. Schelasin. 1993. The simulation of integrated tool performance in semiconductor manufacturing. In *Proceedings of the 25th Winter Simulation conference*, Edited by G.W. Evans, M. Mollaghasemi, E.C. Russel, W.E. Biles, 814-818. Los Angeles, California, USA, Inc.
- Murray, S., J. Geraghty, P. Young and S. Sievwright. 2008. Time-limited next arrival heuristic for batch processing and setup reduction in a re-entrant environment. In *Proceedings of the 2008 Winter Simulation Conference*, Edited by S.J. Mason, R.R. Hill, L. Mönch, O. Rose, T. Jefferson, J.W. Fowler, 2109-2117. Miami, Florida, USA, Inc.
- Niedermayer, H., and O. Rose. 2003. A simulation-based analysis of the cycle time of cluster tools in semiconductor manufacturing. In *Proceedings of the 15th European Simulation Symposium*.
- Noack, D., B. P. Gan, P. Lendermann, and O. Rose. 2008. An optimization framework for waferfab performance enhancement. In *Proceedings of the 40th Winter Simulation Conference*, Edited by S.J. Mason, R.R. Hill, L. Mönch, O. Rose, T. Jefferson, J.W. Fowler, 2194-2200. Miami, Florida, USA, Inc.
- Quek, P. T., B. P. Gan, S. L. Tan, and C. L. Peng. 2007. Analysis of the front-end wet strip efficiency performance for productivity. In *Proceedings of International Symposium on Semiconductor Manufacturing, 2007, ISSM 2007*, 1-4.
- Scholl, W., and J. Domaschke. 2000. Implementation of modeling and simulation in semiconductor wafer fabrication with time constraints between wet etch and furnace operations. *IEEE Transactions on Semiconductor Manufacturing* 13(3):273-7.
- Starkweather, T., S. McDaniel, K. Mathias, D. Whitley, C. Whitley. 1991. A comparison of genetic sequencing operators. In *Proceedings of the 4th International Conference on Genetic Algorithms*, Edited by R.K. Belew, L.B. Booker, 69-76. San Diego, California, USA.
- Yugma, C., Dauzère-Pérès, S., Artigues, C., Derreumaux, A., Sibille, O. 2012. A batching and scheduling algorithm for the diffusion area in semiconductor manufacturing. *International Journal of Production Research* 50 (8): 2118-2132.
- Zeballos, L. J., P. M. Castro, and C. A. Méndez. 2010. A CP-based approach for scheduling of automated wet-etch stations. In *Proceedings of the 2nd International Conference on Engineering Optimisation*. Lisbon, Portugal.
- Zeballos, L. J., P. M. Castro, and C. A. Méndez. 2011. Integrated constraint programming scheduling approach for automated wet-etch stations in semiconductor manufacturing. *Industrial & Engineering Chemistry Research* 50:1705-15.

AUTHOR BIOGRAPHIES

ANNA ROTONDO is a postdoctoral research in the School of Mechanical and Manufacturing Engineering at Dublin City University. She holds a MS degree in management engineering from Politecnico di Bari, Italy, and received her PhD in manufacturing engineering from Dublin City University, Ireland. Her research interests are in modeling, analysis and optimization of complex manufacturing systems, with particular application to semiconductor manufacturing. She collaborates with the Irish Centre for Manufacturing Research to improve the process of developing the models to address specific tactical and strategic decisions. Her e-mail address is anna.rotondo@dcu.ie.

JOHN GERAGHTY is a lecturer in the School of Mechanical and Manufacturing Engineering at Dublin City University, Ireland. He holds a B.Eng., a M.Eng. (by research) and a PhD in Industrial Engineering from the University of Limerick, Ireland. The focus of his research for both his M.Eng. and PhD theses was on inventory and production control strategies in manufacturing systems subject to variability in cycle time and demand distributions. He is a founding member of the Enterprise Process Research Centre at DCU. His research interests include modeling and analysis of semiconductor fabrication systems, lean manufacturing, supply-chain management and operations excellence. His e-mail address is john.geraghty@dcu.ie.

PAUL YOUNG is the director of the Enterprise Process Research Centre at Dublin City University, Ireland. His PhD (Trinity College Dublin, 1991) concerned condition monitoring of turning and was followed by the measurement and modeling of vehicle HVH in Japan. The design, analysis, modeling and monitoring of manufacturing processes was the main focus of his Post-Doctoral research work with the Advanced Manufacturing Research Centre in University College Dublin. In 1998 he was appointed to the faculty in Dublin City University where, since 2001 his research is mainly in the application of modeling and analysis to improve the performance of complex manufacturing systems, while maintaining some machine design and modeling. His e-mail address is paul.young@dcu.ie.