Korean J. Chem. Eng., **30**(1), 27-32 (2013)
DOI: 10.1007/s11814-012-0122-0

Developing an integrated capacity planning framework for production processes and demand supply chains

Jun-Hyung Ryu†

Department of Nuclear & Energy System, Dongguk University, Gyeongju 780-814, Korea
(*Received 5 February 2012 • accepted 28 July 2012*)

Abstract−This paper presents a new capacity planning framework integrating internal production processes and external demand supply chains. Since planning is concerned with preparing how resources are arranged before future situations are realized, the planning problem should consider the entire aspect of a company, including internal and external aspects. Most previous works focused on only one aspect by simplifying the other. The proposed framework is formulated as a mixed integer programming (MILP) problem with an aim to establishing a rigorous decision-making tool. Lessons learned during their industrial implementations are discussed with some remarks.

Key words: Integration, Process Capacity Planning, Supply Chain Planning, MILP

INTRODUCTION

Planning is concerned with preparing resources in preparation for the future with an aim to maximize the profit of a company. A planning problem is thus concerned with the entire aspect of the company, including internal production processes as well as external demand supply chains. Here the internal production processes are typically determined by planning the capacity of production processes over multiple time periods. The production capacity planning is critical since it directly determines the production amounts and their delivery dates. Production planning has been an active area of research; see, for example [1,2]. On the other hand, external demand supply chain planning is also an important issue that should not be underestimated, as profitability and customer relationships are mainly dependent upon it. Numerous studies have been conducted including state-of-the-art reviews [3,4], multiperiod planning [5], and even regulatory factors in the capacity expansion of supply chains [6].

In the previous work, the capacity of production processes was simplified in overall planning problem. For example, in capacity planning over multiple time periods, the capacity of a process, $R_{i,t}$ is determined by the following constraint:

$$b_{l,i}R_{l,t}+INV_{l,i,t}\geq S_{l,i,t}+INV_{l,i,t+1} \quad \forall l, i, t \tag{1}$$

where $S_{l,i,t}$ denotes a demand which is allocated to process l of product i at time t and $b_{l,i}$ is the performance coefficient of product i in the plant l. The discrete decision of expanding the capacity is realized by the constraint with discrete variable representing whether to expand the current capacity at a certain time t. For instance, the capacity of plant l at time t, $R_{l,t}$ is expanded from the previous time period as much as the capacity expansion amount of $CE_{l,t}$.

$$R_{l,t}=R_{l,t-1}+CE_{l,t} \quad \forall l, t \tag{2}$$

However, it is unrealistic to assert the capacity of complex processes represented by only a few aggregated constraints as (1) and (2). A separate rigorous planning framework should be employed to evaluate the actual capacity of the production processes.

Chemical processes consist of a series of multiple units of tools connected in various configurations. A tool or a machine in the process can do multiple unit operations. The operator shifts the role of the tool to meet the target of the process. The capacity of the process is thus determined by deciding which tool is playing which role during the limited time periods. The weekly or monthly production target is set and the corresponding managing schedules such as maintenance are prepared while meeting the target.

Production capacity can be determined as a combination of three parts: product, process and tool as shown in Fig. 1. A product is ready after being through a series of processes. The sequence and condition of these processes are distinctive in terms of product. The actual process is operated in a tool, a machine. A tool may operate a number of distinctive processes and a process can be operated in distinctive tools, respectively. Therefore a key issue of process capacity planning is how to reflect the presence of a large number of combinations of which tool to make which product in which process for meeting the production target. The target is determined from the external demand supply chain planning.

There is a clear difference between detailed production capacity and enterprise-wide demand supply chain. Both of them aim to maximize the profit of the company, They do not deal with totally different problems but their perspectives are different. Therefore, it is

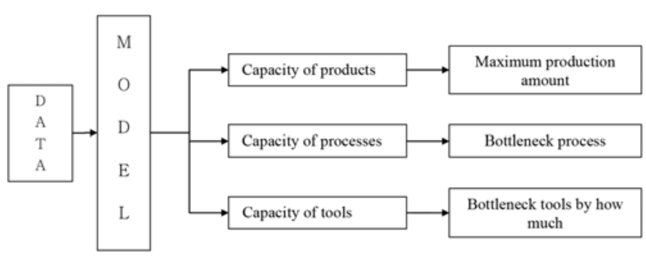

Fig. 1. Process capacity representation.

†To whom correspondence should be addressed.
E-mail: jhryu@dongguk.ac.kr

necessary to have a communication link between the production capacity planning and demand supply planning. That is, the production capacity should be computed again for a certain set of products using the result of demand chain planning. This paper aims to meet the need by proposing an integration framework.

This paper employs a semiconductor manufacturing process and its supply chain to illustrate the proposed framework. Semiconductor manufacturing, particularly the fabrication process, involves many chemical and process processes such as etching and chemical vapor deposition (CVD). The industry is under severe competition, it is a good candidate to illustrate the impact of integration as an effort to remain competitive. The rest of this paper is as follows: At first, a mathematical production capacity planning model is addressed. The integrating framework of capacity planning and supply chain planning is then proposed with industrial implementation case. The industrial results will be discussed with some remarks.

MATHEMATICAL CAPACITY PLANNING MODEL

A process capacity planning model can be mathematically constructed by encapsulating the actual processes. In this paper, complex fabrication processes will be employed as an illustration in terms of product, process and tool use a similar way as [7]. Generally speaking, fabrications processes are divided into major processes such as etch, chemical mechanical polishing (CMP), photolithography, chemical vapor decomposition (CVD) in terms of characteristics. Tools are mainly divided in terms of the process technology and manufacturer. The information necessary for the proposed mathematical framework is as follows: product process plan and step sequence, available tool information, and production target. Then two optimization problems are constructed. At the first problem, the number of wafers to operate at each manufacturing step is computed. At the second, the capacities of individual processes are maximized by using the result of the first problem.

The objective function of the first problem is to maximize production of all products:

$$\max \sum_{p} \mathrm{PROD}_p \tag{3}$$

The following constraints are considered. At first, the output amount from step i is computed by multiplying the starting material amount with its corresponding yield:

$$\mathrm{OUT}_i = \sum_j \mathrm{yield}_i \times \mathrm{Start}_{i,j} \quad \forall \mathrm{i} \tag{4}$$

Basic material balance should be also met: Amount of flow at the each manufacturing process stage should be the same of the sum of material in hand and material coming out of it. All amounts at each step i are transferred to its next steps after process:

$$\mathrm{OUT}_i = \sum_{k \in NE_{i,k}} \mathrm{FLOW}_{i,k} \quad \forall \mathrm{i} \tag{5}$$

where NE_{ik} denotes a set of the next steps of step i.

The input amount at step i is the sum of all its previous steps:

$$\sum_{k \in PR_{i,k}} \mathrm{FLOW}_{i,k} = \sum_j \mathrm{Start}_{i,j} \quad \forall \mathrm{i} \tag{6}$$

where $PR_{i,k}$ denotes a set before step k for product i.

During everyday operations, a malfunction or an error during an operation can result in a defect. Such loss is also considered in terms of loss rate, which is computed based on the analysis of the previous statistical data. On the other hand, setup cost is not directly considered but included in the processing times because the changes are so frequent.

In terms of flow computation, reentrant flow rates are considered as a rework factor which is determined from the statistical history data. Rework amount at step i is the multiplication of output i and rework ratio:

$$\mathrm{Flow}_i = \mathrm{output}_i \times \mathrm{rwratio}_{i,r} \quad \forall \mathrm{i}, \mathrm{r} \in \{\text{rework steps}\} \tag{7}$$

The final production amount is the output at the last step:

$$\mathrm{PROD}_p = \mathrm{OUT}_l \quad \forall \mathrm{p}, l\text{: the last step of product p} \tag{8}$$

From the solutions of the above problem, we can obtain a number of wafers to operate at each step to obtain a wafer in the end. In the next, another optimization problem is formulated to maximize the capacity of individual processes:

$$\max \sum_{i \in CG} \mathrm{PCap}_i + \mathrm{BNCAP}_i \tag{9}$$

Here the total time for production should be less than the available time for which tools can operate. The available time of the tool for the process group is the same or bigger than necessary time for production.

$$\sum_{i \in (PGg)} \mathrm{Start}_i \times \mathrm{ST}_{i,j} = (\mathrm{TL}_{g,j} - \mathrm{TL}^+_{g,j}) \times \mathrm{util}_j \quad \forall \mathrm{g}, \mathrm{j} \tag{10}$$

where $ST_{i,j}$ is the processing time and $TL_{g,j}$ denotes the integer variable allocating tool group g to process j. TL^+_{gj} is the redundant number of tools after being assigned to process j and $util_j$ is the utilization rate of the tool.

The number of tool groups for operations should be arranged into an integer:

$$\sum_g \mathrm{TL}_{g,j} = \mathrm{TLQty}_j \tag{11}$$

where $TL_{g,j}$ is nonnegative integer and $TLQty_j$ denotes the total available time for tool at station group j.

At all steps, the amount at step i in tool j should be the product of the bottleneck capacity and the result of the first optimization problem:

$$\sum_j \mathrm{Move}_{i,j} = \mathrm{Start}_i \times \mathrm{BNCAP}_i \quad \forall \mathrm{j} \in \{\text{unchecked process group}\} \tag{12}$$

$$\sum_j \mathrm{Move}_{i,j} = \mathrm{Start}_i \times \mathrm{PCap}_i \quad \forall \mathrm{j} \in \{\text{checked process group}\} \tag{13}$$

where $BNCAP_i$ denotes the maximum capacity of the unchecked bottleneck process group and $PCap_i$ is the maximum capacity of the checked bottleneck process group.

All process groups should have a bigger capacity than the previously obtained upper level value.

$$\mathrm{BNCap}_i \geq \mathrm{prevUPcap} \quad \forall \mathrm{i} \tag{14}$$

$$\mathrm{PCap}_i \geq \mathrm{pCAPUP} \quad \forall \mathrm{i} \in \{\text{checked process group}\} \tag{15}$$

where prevUPCAP is the solution of the $BNCAP_i$ at the previous iteration and pCAPUP is the maximum capacity.

This optimization problem is iterated for all process groups and the corresponding capacities are computed and compared. The corresponding MILP problems are solved using CPLEX and a series of decomposition methodologies are implemented to reduce the computational loads. For the detailed mathematical programming formulations, the reader may refer to the other previous research such as [7,8].

INTEGRATING CAPACITY PLANNING WITH SUPPLY CHAIN PLANNING

The integration can be addressed in the context of hierarchical decision-making structure. After constructing mathematical models encapsulating fabrication processes and developing computation methodologies to compute the solution, it is still early to say the job is finished from an industry's perspective. The model should be incorporated into the existing decision-making frameworks and contributed to explicitly raising the overall performances of the company. With this regards, the proposed framework has been explicitly associated with the overall decision-making systems of semiconductor manufacturing supply chains to make actual contribution. Refer to Fig. 2 for its integration with other systems.

As can be seen in Fig. 2, semiconductor manufacturing decision-making frameworks consist of many supply chain planning systems of distinctive horizons and boundaries that are hierarchically connected each other. The upper level planning is mainly concerned with planning over a long term and wide boundary in an aggregated manner. On the other hand, the detailed and local issues are focused in the low level. The results of a module are used as information for determining different objectives in other modules. The proposed capacity planning framework can be also incorporated to provide the information on the fabrication capacity, which is the bottleneck in the entire supply chain.

Since many tools are often customized in real plants and their capacities cannot be expanded within a short time, the effect of the capacity expansion is only valid for a specific production target. We cannot increase all products except the targeted when we expand a specific capacity of a tool or process. This means that we not only need accurate demand information but also rigorous capacity estimation system to achieve a tangible result. Because of their relationships, the capacity planning system should be incorporated into supply chain decision-making systems.

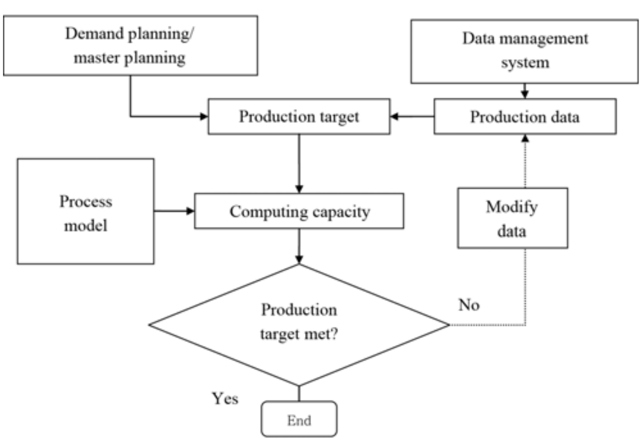

Fig. 2. Proposed integrated planning decision-making structure.

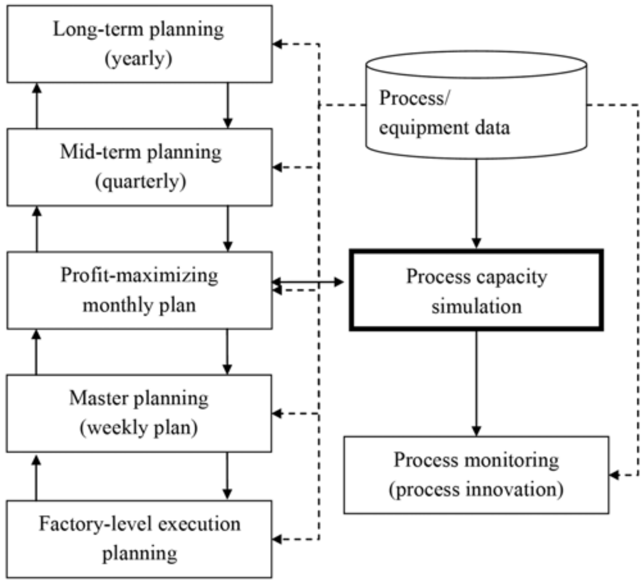

Fig. 3. Integration of capacity planning and overall supply chain decision-making framework: Schematic representation.

As shown in Fig. 2, individual customer requests are summarized in demand planning and they are assigned to the operation target in supply chain planning module with an objective of maximizing the profit of the company. The proposed capacity planning allows us to review whether the production target can be met in the process and to check which tool or unit process is deficient to meet the target. It is also possible to test how the process should be operated to operate different production targets. This simulation function is of great help to review the production planning over a number of business scenarios as shown in Fig. 3.

For example, some decision-making problems take advantage of fabrication capacity as a parameter in their problems. It is suitable to incorporate the proposed framework with the monthly planning module where corporate profit-maximizing strategies are explicitly pursued. The monthly planning which is most often used is concerned with the forthcoming three months operation of the company. The weekly planning deploys fabrication capacity as a constraint but in much simplified types. It does not provide enough information for the adjustment of production targets. Here the proposed planning framework can allow us to review the effect of the target change in terms of fabrication manufacturing. When the demand of the future time period is in the increasing trend and the currently estimated corresponding capacity cannot support it, the company should prepare expanding the capacity before the increasing demand is materialized. It would be also possible to incorporate it with higher modules in the decision hierarchy with the aggregated format to simulate the future capacity with regards to forecast. For the long-term planning for the future decisions, their best performance data can be employed in the proposed capacity framework. The important issue is that these decision-making modules should not be stand-alone but exchange their results between each other to review the feasibility and obtain action items to execute the scenarios. This closed-loop communication feature is an important tool to stay competitive under the challenging business environments.

INDUSTRIAL IMPLEMENTATION CASE

The proposed integration framework has been implemented in a semiconductor company in Korea. Before addressing the result of the proposed framework, here is a brief background of the semiconductor industry. More detailed information on semiconductor manufacturing supply chain can be found in [9].

Commonly called microchip or integrated circuit (IC), “semiconductor” refers to components that provide intelligence functions in various electronic products such as computers, mobile communication equipment, and digital photography. Because of its wide usage, it has enormous markets--as much as $220 billion for the year of 2009. Many companies are thereby eager to enter this industry with its huge economic potential but there are challenging barriers.

The impact of fabrication processes is very significant and its decision-making should be systematically taken into account. The fabrication capacity should be continuously revised and expanded to keep up with varying customer demands. The capacity can be expanded by improving the performance of the bottleneck tools and processes in terms of time and processing amounts to meet a specific production target. One fundamental issue in expanding the capacity is that we have to know the exact capacity of a fabrication process.

Because of its grave impact, increasing attention has been given to semiconductor fabrication capacity planning. Swaminathan [10] provided a tool capacity planning model based on stochastic programming formulation with heuristic based solution methods. Particularly, it is worthwhile to mention the work by IBM researchers [7,11]. They modeled semiconductor production processes into mathematical programming problems and computed tool purchase investment plans based on the model. However, they did not explicitly address how their capacity planning models were associated with the other planning system as a part of the overall decision management problem. Another important point to highlight is that the previous work has not explicitly addressed how industries actually take advantage of their capacity planning, to the author’s knowledge, while the issue of assessing the overall capacity is an everyday issue faced by all semiconductor companies.

Some of the challenges in the semiconductor industry include the following. First, *it is very expensive to enter and survive*. Huge investments must be made continuously, often in terms of billions of dollars. Such large investments can be of little use when they fail to provide products in time because the prices collapse so fast. Second, *it takes very long to manufacture the product*. Turn-around time (TAT), which denotes the total elapsed time from the start until semiconductors are ready for sale, is as long as 30 or sometimes 60 days. Among the four major stages of semiconductor manufacturing processes (fabrication, probe, assembly and test), fabrication consumes the largest portion, roughly more than 85% of TAT. The long processing time is due to the presence of more than 200 manufacturing processes and 100 testing steps. During fabrication, intermediate products in the form of a wafer often have to be operated repeatedly at the same step for the purpose of purification, cleaning, etc. This reentrant feature makes its operation management more complex. Thirdly, *its product life cycles are rapidly shrinking*. The

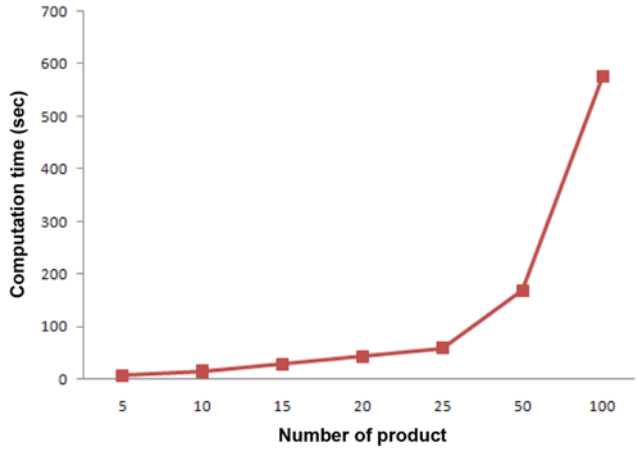

Fig. 4. Computation times of the proposed framework according to the increase of number of products.

price of some products is lowered more than 75% only after two years. Under this circumstance, we can see that more advanced products need to be continuously developed and introduced right on time in the market before competitors do.

The implementation of the proposed framework was carried out in three steps. First, a sample manufacturing line was selected to evaluate the validity of the proposed model. The necessary data of the line for building the above presented mathematical model were prepared after the evaluation of the personnel in the field to check the validity of the data. The capacity configuration according to the proposed model was then compared with the one by a spreadsheet based heuristic method that has been used so far. Fig. 4 shows the computational statistics as the increase of product number. Second, the entire data structures of other remaining lines have been arranged to provide the necessary data for the planning system and are constructed as a database. As the processes are operated, their performances are updated via the database system. Third, the above proposed mathematical model and the database are incorporated into a unified system. The resulting capacity planning framework is graphically illustrated in Figs. 1 and 2. Fourth, semiconductor manufacturing capacities have been expanded.

Based upon the above-described models and business processes, the fabrication capacity estimation framework is constructed. The next will illustrate the proposed work using a simple example: Consider a fabrication line with three products. Since the products are divided in terms of their technology, such as 100 nm (nanometer), 70 nm, etc., each product actually represents a group of further distinctive products. The illustrated result of the proposed framework can be explained in terms of processes as can be seen in Figs. 5 and 6.

By implementing the proposed planning, the capacities of the individual process groups, which are denoted as the separate rectangles, are computed. The minimum unit capacity of all processes as a result of the model denotes the maximum possible capacity of this fabrication line. It is about 78,000 wafers per month in Fig. 5 and *process 8* is the bottleneck which denotes the photolithography process in this case. The products A, B and C denote the product group with the same process technology, respectively.

The result of the proposed framework can be also analyzed in terms of tools as shown in Fig. 6. The y-axis denotes the number

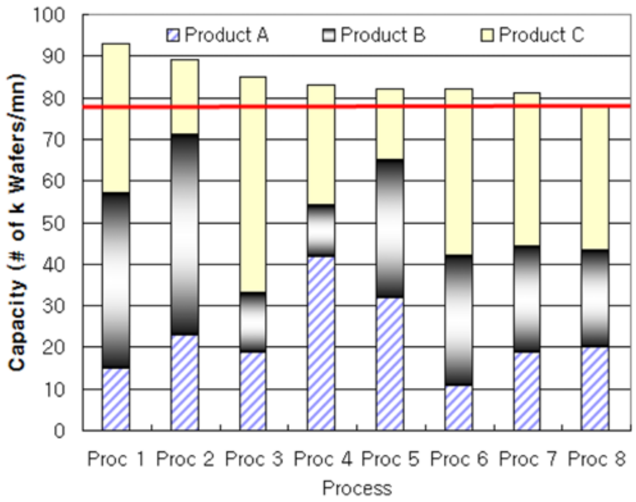

Fig. 5. An illustration of the capacity planning in terms of production amount.

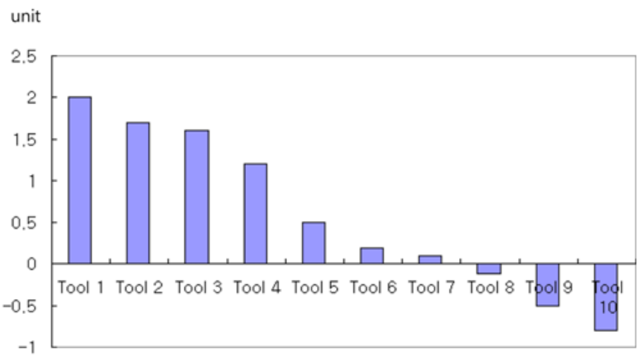

Fig. 6. An illustration of capacity planning in terms of tool.

of units at each manufacturing stage. To meet the current demand, the proposed modeling framework allows us to obtain a number of tools at each stage. By comparing the necessary number of tools (result) and the current number of tools at each stage (current practice), we can discover which stage is in need of expanding the capacity by how many tools. That is, the sufficient or insufficient number of units of the current tools while meeting the demand is obtained. The (−) signed unit denotes that we need more units for the specific tool in the corresponding stage and (+) signed unit represents that we have enough tools for the current stage. Therefore, we can obtain the information on the bottleneck tool by how much. For the case of Fig. 6, the current capacity of the process is as deficient as the 0.5 unit of *tool 9* and 1 unit of *tool 10* to meet the current production target while other tools are redundant. From this result, it can be said that the key issue for the production line is to figure out how to increase the capacity of the tool 9 and tool 10. For instance, it may be thought that these tools are shared with other lines.

A number of remarks can be made from the experience of implementing the proposed framework in practice. First, it is necessary to elaborate on the data preparation system for the proposed fabrication capacity planning framework. For computing capacity planning, a large amount of data should be prepared. The necessary data for the proposed model is transported from data information system involving all the data such as product, process steps, and processing times, operation types and conditions, BOM (Bill of Material), etc. It is an important and challenging issue to handle these data efficiently. In semiconductor companies, a huge amount of statistical data on products and tools has been accumulated mainly to monitor product performances, qualities, and tool utilization rate. It should be guaranteed in terms of accuracy and constant availability. It is thereby based on rigorous data acquisition systems which provide accurate and up-to-date data. The works associated with this in industry are called master data management (MDM). If appropriate data are not available or not updated in response to progress, the result of the model does not represent the actual capacity; the line cannot produce as many wafers as computed and in the end may lose precious financial opportunity. Thus, for the successful management of semiconductor manufacturing, cooperation between participants in various positions is quite essential.

Secondly, the proposed capacity result can be applied in other decision-making problems. We can see whether a certain tool is redundant or deficient from the capacity planning result. For a deficient tool, we can think of (1) purchasing a new tool, which is associated with tool purchase plan, or (2) increasing the process performance by improving the relevant parameters, which is with process innovations. The redundant tools, on the other hand, can be used to operate for other lines which are short of them. The proposed framework thus plays the role of a simulation tool to evaluate the effect of a business strategy to increase the capacity. For instance, we can have the result for multiple cases such as (i) the effect of merging a number of process steps into a single step, (ii) the effect of increase of capacity when tool down time is reduced, or (iii) the effect of installing a new tool from a tool vendor or other lines to respond to the increase of demand.

When we compare the result by the conventional heuristic method and the proposed framework, we notice an interesting issue for comments. As noted earlier, planning is concerned with establishing the target for execution in practice. The target is determined based upon the previous performance data. Here, we have to select what type of data to use among many previous results. When we just take the best parameters, the resulting target is sure to be quite high, which looks great for managers but challenging for the execution side because it is only possible when everything is the best: One deviation in performance may cause the result to be missed. That can be an 'excellent' target but what also should be acknowledged is the small chance of realization. When we only take the average or below average value, we may get a 'negotiable' result in the sense that the target can be met highly probably, but that does not actually represent the maximum capacity the line can do. Therefore, cautious business communications should be in place between the planning and the execution side in a direction to drive the target to be met. To construct a proper decision-supporting tool, the proposed planning is incorporated with the overall supply chain planning frameworks which will be addressed next.

CONCLUDING REMARKS

This paper proposes an integration planning framework of production capacity and supply chain capacity. Since the capacity denotes an actual target for operation in practice, focus should be given to computing them in a systematic decision-making procedure while respecting business in-the-field practices as much as possible. In this paper, the work process flow in the context of linkage between multiple models is highlighted instead of detailed mathematical programming equations. Incorporating industrial practices and academic studies is difficult but necessary work. Such integration is beneficial to both sides. With this regard, there is much to be done from the area of process systems engineering (PSE) communities. This manuscript is in line with this effort.

To survive under challenging economic environments, many issues should be considered to stay competitive; for example, maintaining favorable relationships with customers, and keeping process performing in the best condition to meet the manufacturing target. Because these important issues cannot be separated but approached from the universal perspective of maximizing the overall profit, industries do planning over multiple time periods. Many efforts have been made but still to come because planning is expected to remain an important task for the industry, as the future will be even more challenging.

ACKNOWLEDGEMENT

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2010-0023678).

NOMENCLATURE

Index

p : product (p=1, …, P)
g : tool group (g=1, …, G)
i, k, r, l : manufacturing step
j : station group

Parameters

prevUPCAP : solution of the $BNCapa_{unc}$ at the previous iteration
pCAPUP : max capacity of the process group at the previous iteration
$rwratio_{i,r}$: rework ratio of step i to r
$ST_{i,j}$: processing time at step i at station j
$TLQty_j$: total available time for tool at station group j
$util_j$: utilization rate at j

Variables

$BNCAP_i$: the maximum capacity of the unchecked bottleneck process group
$Flow_{i,k}$: amount from i to k
OUT_l : output amount at the last step l
$PCap_i$: the maximum capacity of the checked bottleneck process group
$pmix_p$: product mix ratio
$PROD_P$: out amount of product p
$Start_i$: inserted amount at step i at tool station j
$TL_{g,j}$: tool processing time at tool group g at station j
$TL^+_{g,j}$: processing time at tool group g at station j
$yield_i$: wafer yield at step i

REFERENCES

1. N. V. Sahinidis, I. E. Grossmann, R. Fornari and M. Chathathi, *Comput. Chem. Eng.*, **13**, 1049 (1989).
2. J. H. Ryu, *Ind. Eng. Chem. Res.*, **45**, 6622 (2006).
3. V. A. Varma, G. V. Reklaitis, G. E. Blau and J. F. Pekny, *Comput. Chem. Eng.*, **31**, 692 (2007).
4. N. Shah, *Comput. Chem. Eng.*, **28**, 929 (2004).
5. J. H. Ryu and E. N. Pistikopoulos, *Ind. Eng. Chem. Res.*, **46**, 8058 (2007).
6. H. C. Oh and I. A. Karimi, *Ind. Eng. Chem. Res.*, **43**, 3364 (2004).
7. S. Bermon and S. Jean Hood, *Interfaces*, **29**, 31 (1999).
8. B. Catay, S. S. Erenguc and A. J. Vakharia, *Comput. Oper. Res.*, **30**, 1349 (2003).
9. J. H. Ryu and I. B. Lee, *Korean Chem. Eng. Res.*, **46**(3), 571 (2008).
10. J. M. Swaminathan, *Eur. J. Oper. Res.*, **120**, 545 (2000).
11. S. Jean Hood, S. Bermon and F. Barahona, *IEEE T. Semiconduct. M.*, **16**, 273 (2003).