



Contents lists available at ScienceDirect

## Int. J. Production Economics

journal homepage: [www.elsevier.com/locate/ijpe](http://www.elsevier.com/locate/ijpe)

# An integrated plant capacity and production planning model for high-tech manufacturing firms with economies of scale

Chaug-Ing Hsu <sup>a,\*</sup>, Hui-Chieh Li <sup>b</sup>

<sup>a</sup> Department of Transportation Technology and Management, National Chiao Tung University, 1001 Ta Hsueh Road, Hsinchu 30010, Taiwan ROC

<sup>b</sup> Department of Marketing and Logistics Management, Ta Hwa Institute of Technology, Qionglin, Hsinchu County 307, Taiwan ROC

## ARTICLE INFO

### Article history:

Received 27 September 2005

Accepted 9 September 2008

Available online 10 January 2009

### Keywords:

Supply chain network design

Simulated annealing (SA)

Economies of scale

## ABSTRACT

This study developed a nonlinear mixed integer programming (MIP) model for high-tech manufacturer to determine the optimal supply chain network design. The impacts of economies of scale on the optimal capacity and the production amount are also explored. A heuristic solution approach, based on simulated annealing (SA), is developed to solve the optimal problem. An example of a wafer foundry company is provided to demonstrate the application of the model. Results show that when determining the production amount for multiple plants, a large-sized capacity plant with low capital costs and low production costs has a high priority for fulfilling the capacity due to not only having the higher capability to satisfy the customer demand but also the advantage of saving costs. The results show that the benefits brought about by centralized production are larger than the increased transportation cost. The results also show without using many small-sized capacity plants combined with high utilization, operating few larger-sized capacity plants with lower utilization is more cost effective for the manufacturer as long as the customer demand is large enough to offset the high capital cost.

© 2009 Elsevier B.V. All rights reserved.

## 1. Introduction

It has been recognized that operating cost economies of scale are associated with facility size and utilization (Cohen and Moon, 1990). For high-tech manufacturers, the investment in capacity usually involves high capital costs. Economies of scale allow manufacturers with a large-sized capacity to operate more economically than those with a small-sized capacity. However, it could result in high production costs if the market demand is insufficient to realize economies of scale, and if the capacity utilization of the manufacturing plant is low. The manufacturer can also invest many small-sized capacity plants instead. Though there is a good chance of reaching a full-capacity production, the production cost may not be

minimized due to lack of economies of scale and incapability of applying advanced technology in these plants. Moreover, for a manufacturer operating multiple plant sites in different regions, additional complicating factors need to be taken into account, such as investment conditions in different regions and physical distribution problems between customers and plant sites. The former involves capital and variable production costs, while the latter affects customers' satisfactions and outbound costs of the product.

Determining the capacity for plants is fundamental to the manufacturer's long-term planning, while the production assignment among plants is classified as medium- and short-term decisions (Santoso et al., 2005). Manufacturer's total production cost depends mainly on capacity utilization of all the plants, which is the result of the decisions, i.e. plant capacity and production assignment among the plants. Moreover, market demand governs total production amount. In other words, market demand,

\* Corresponding author. Tel.: +886 3 5731672; fax: +886 3 5720844.

E-mail address: [cihsu@mail.nctu.edu.tw](mailto:cihsu@mail.nctu.edu.tw) (C.-I. Hsu).

capacity and production amount of the plants determine manufacturer's total production cost. The wafer fabrication manufacturing is characterized with high capital cost and the impacts of economies of scale on production cost are obvious. The final product of the manufacturer is die, which can be produced in either 12-, 8- or 6-inch wafers. Due to the complexity production process, the variable production cost per wafer and capital cost of a 12-inch wafer plant is the highest than the others. However, operating 12-inch wafer fabrications (FABs) is a more economical alternative than the others due to the lower production cost per die when there is larger customer demand to yield higher capacity production. The capacity utilization and plant capacity determines the total production cost of the manufacturer. In other words, whether or not the manufacturer reaches a minimized cost depends on the allocation of total production between different-sized capacity plants for all different market demand. This study aims to determine the optimal location, capacity and production of the plants to minimize total cost by considering the production characteristics of wafer fabrication industry and economies of scale.

Past studies have investigated plant-loading issues in which the problems are characterized by fixed facility capacity and given facility locations. Cohen and Moon (1990) applied regression analysis to investigate the impacts of production economies of scale, manufacturing complexity, and transportation costs on supply chain facility network. Brimberg and ReVelle (1998) formulated a bi-objective plant location model for analyzing the trade-off between total cost and the portion of the market to be served. In this study, partial satisfaction of demand is considered rather than serving all demand in the traditional plant location problem. The weighting method approach is investigated for obtaining efficient solutions of the model. Jayaraman and Pirkul (2001) extended the plant location problem to incorporate the tactical production–distribution problem for multiple commodities. In the model, a facility or warehouse is constrained to serve one single customer. Miranda and Garrido (2004) proposed a simultaneous approach to incorporate the inventory control decision with typical facility location models. Cohen and Moon (1991) formulated a MIP model to determine the optimal assignment of product lines and volumes to a set of capacitated plants. In the model, the capacity of plants is given and fixed, and the production cost function exhibits concavity with respect to each product line volume reflecting economies of scale. Moreover, correlation and regression analyses are employed to analyze the relationship between the cost parameters. The results indicate that focused plants arise in situations with high economies of scale. However, although the capacity utilization of different-sized plants will result in various influences on their total cost, its extent is seldom discussed. Moreover, the models constructed are usually large-scale linear or nonlinear MIP formulations, which are difficult to solve. Therefore, these studies focused mainly on developing an approximation procedure and compared the efficiency of their proposed heuristics with others.

Regarding the supply chain management field, a lot of issues have been extensively discussed. One of those

discussed the integration and coordination of different functions or participants in the chain, such as buyer–vendor coordination (e.g., Tzafestas and Kapsiotis, 1994; Viswanathan and Piplani, 2001), production–distribution coordination (e.g., Goetschalckx et al., 2002) and inventory–distribution coordination (e.g., Fu and Piplani, 2004; Chen et al., 2001). Other studies developed supply chain network design models in which different factors are considered. Nagurney et al. (2002) and Nagurney and Toyasaki (2005) considered many decision-makers and their independent behaviors in the supply chain and developed an equilibrium model of a competitive supply chain network, in which transportation links are associated with different costs. Tsiakis and Papageorgiou (2008) constructed a mixed integer linear programming (MILP) model to determine the optimal configuration of a production and distribution network subject to operational and financial constraints. Guinet (2001) focused on the multi-site production planning problem by primal-dual approach to examine the workshop scheduling problem. Levis and Papageorgiou (2004) formulated a mathematical programming model to study the long-term capacity-planning problem under uncertainty in pharmaceutical industry. There are few previous studies on supply chain design models considering the high capital costs invested in the capacity by high-tech manufacturers, and how capacity utilization of different-sized plants affects the total average production cost.

In a different line of research, issues in the supply chain of high-tech manufacturing industries have been discussed. Julka et al. (2007) established the current state of research in multi-factor models for capacity expansion in the manufacturing industry. The weakness and strength of past research and opportunities to future studies are also summarized. Nazzal et al. (2006) presented a comprehensive framework for strategic capacity expansion of production equipment so as to cut down cycle times. Chou et al. (2007) evaluated alternative capacity strategies in semiconductor manufacturing under uncertain demand and price scenarios. In the paper, the capacity planning is defined as the preparation for plant transition in anticipation of new process and new product. There are also papers conducting operational level issues regarding high-tech manufacturing industries. Chen et al. (2005) considered production scheduling planning, and developed a capacity-planning system, which considered the capacity and capability of equipment for multiple semiconductor manufacturing FABs. In their paper, “capacity” refers to the upper threshold of the load on an operating unit and “capability” refers to a specific processing capability of a machine, respectively.

Past studies have focused on the manufacturing production issues in which the decisions involve either long-term or short-term. There are few studies putting emphasis on issues about plant capacity and production allocation among plants, in which the long-term and short-term decisions are integrated and treated as endogenous decision variables, taking into account of location, market demand and economies of scale especially for plants in wafer fabrication industry. The distinguishing features of the study are the comprehensive consideration

of economies of scale and capacity utilization, which are due to different-sized and locations of the plants deployed by the manufacturer in the long run and the production assignment for different market demand in the short run. In addition, this study develops the supply chain design model, which determines not only the capacity and production amount of each plant, but also the active vendors, inbound raw material flows from vendors to plants and the market served by each plant. The model specifically applies MIP formulations and attempts to minimize the average total cost per unit product subject to constraints such as satisfying customer demand in various geographic regions, relationships between supply flows and demand flows within the physical configuration as well as the production limitation of different-sized plants.

The remainder of this paper is organized as follows. Section 2 formulates the supply chain network design model. Section 3 demonstrates the development of the algorithm to solve the proposed models. A case study about a leading wafer foundry company in Taiwan is provided in Section 4 to illustrate the application of the model. Finally, Section 5 presents a summary of the findings of the study.

**2. Model formulation**

This study aims at designing a supply chain network model for a manufacturer who operates multiple plants in different regions. As in Vidal and Goetschalckx (1997), the supply chain design problem in this study can be described as: given the customer demand, determine the capacity and production amount for each of the manufacturing plants, and the amount of raw material/final products shipped from vendors/manufacturing plants to manufacturing plants/customers. In other words, the supply chain network design problem is formulated to determine the optimal flows between alternatives from the upper echelons and those of the lower echelons, and the supply flows of all alternatives at different echelons. This study denotes a supply chain network as a directed graph where there are various echelons such as raw material vendors, manufacturing plants and customers and each of them involves different numbers of alternatives, which are located in different regions. In the supply chain network, the relationship between the inbound and outbound flows of an alternative in an echelon, and the relationship between total supply flows among echelons and customer demand are further explored and identified as none of the one-to-one problem. In addition, “capacity” in the paper represents the maximum production amount, i.e. number of pieces of wafer per month. The model formulation follows the production characteristics of the wafer fabrication industry.

Consider a supply chain network  $G(N,A)$ , where  $N$  and  $A$  represent the set of nodes and the set of links, respectively, in a directed graph  $G$ . Let  $k$  denote a specific echelon in a supply chain. From the uppermost echelon to the customer echelon,  $k = 0, 1, \dots, s$ , where  $k = 0$  represents raw material vendors,  $k = s$  is customers, respec-

tively. Let  $n_k$  denote a node at echelon  $k$ , which can also refer to an alternative at echelon  $k$ ,  $n_k \in N$ . A specific node can be classified as a demand or a supply node according to the emanation of the link. When the link is outgoing/incoming from/to a node, then the node can be described as a supply/demand node and the outflows/inflows are also called the supply/demand flows of that node.

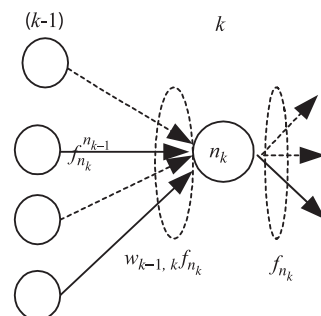
Let  $f_{n_k}$  represent the supply flows of a node at echelon  $k$ ,  $n_k$ , which is transferred from its demand flows. This study considers the fact the demand and supply flows of a specific node might not be an equal amount and let  $w_{k-1,k}$  denote the demand flows of a node at echelon  $k$  from its upper echelon ( $k-1$ ) on condition that the supply flow is one unit. Consequently, the demand flows of a node at echelon  $k$  can be estimated by  $w_{k-1,k}f_{n_k}$ . Furthermore, the following condition must be satisfied, i.e. the demand flows of node  $n_k$  are the sum of the flows from the nodes at its upper echelon ( $k-1$ ),  $n_{k-1}$ . The above description can be expressed as

$$w_{k-1,k}f_{n_k} = \sum_{\forall n_{k-1}} f_{n_{k-1}} \delta_{n_k}^{n_{k-1}} \quad \forall n_k \tag{1}$$

where  $f_{n_{k-1}}$  represents the flows between a node at the upper echelon,  $n_{k-1}$ , and a node at the lower echelon,  $n_k$ , and  $\delta_{n_k}^{n_{k-1}}$  is an indicator variable representing whether  $n_{k-1}$  is an active node to  $n_k$ , that is to say, if the node at echelon ( $k-1$ ),  $n_{k-1}$ , supplies or transports raw materials/intermediate products/finished goods to the node at its lower echelon,  $n_k$ , then  $\delta_{n_k}^{n_{k-1}} = 1$  and the amount transported is  $f_{n_{k-1}}$ , otherwise,  $\delta_{n_k}^{n_{k-1}} = 0$ . Fig. 1 illustrates the relationship mentioned above. The solid lines in Fig. 1 show that the nodes at the upper and lower echelons are connected; in other words,  $\delta_{n_k}^{n_{k-1}}$  is 1, while the dotted lines show that  $\delta_{n_k}^{n_{k-1}}$  equals to 0. The supply flows of the node at echelon  $k$ ,  $f_{n_k}$ , can also be represented as the sum of the flows between a node at echelon  $k$ ,  $n_k$ , and nodes at its lower echelon,  $n_{k+1}$ , such as

$$f_{n_k} = \sum_{\forall n_{k+1}} f_{n_{k+1}} \beta_{n_k}^{n_{k+1}} \quad \forall n_k \tag{2}$$

where  $\beta_{n_k}^{n_{k+1}}$  is an indicator variable representing whether a node at the lower echelon,  $n_{k+1}$ , is a demand node for node at its upper echelon,  $n_k$ . If node  $n_{k+1}$  demands or orders raw material or products from a node at its upper echelon  $k$ ,  $n_k$ , then  $\beta_{n_k}^{n_{k+1}} = 1$  and the quantity demanded is  $f_{n_{k+1}}$ ; otherwise,  $\beta_{n_k}^{n_{k+1}} = 0$ . To sum up, variables  $\delta_{n_k}^{n_{k-1}}$  and



**Fig. 1.** The relationship of flows between nodes at the upper echelon and nodes at the lower echelon.

$\beta_{n_{k+1}}^{n_k}$  can be used to explain, respectively, whether nodes at the upper and lower echelons are active alternatives of the node at echelon  $k$ ,  $n_k$ , while variables  $f_{n_k}^{n_{k-1}}$  and  $f_{n_{k+1}}^{n_k}$  represent flows between  $n_k$  and the active nodes at their upper and lower echelons, respectively, in a supply chain. In addition, capacity constraints are imposed on the supply flows of nodes at echelons, that is,  $f_{n_k} \leq v_{n_k}$ , where  $v_{n_k}$  represents the capacity of  $n_k$ .

The discussion so far has dealt with the relationship and the flows between a node at echelons and nodes at its upper and lower echelons. Next, this paper explores and formulates the total supply flows at different echelons and the total customer demand. Similar to the expression of Eq. (1), the total customer demand is satisfied only if the supply flows at its upper echelon ( $s-1$ ),  $\sum_{\forall n_{s-1}} f_{n_{s-1}}$ , equals to  $w_{s-1,s} \sum_{\forall n_s} f_{n_s}$ , where  $f_{n_s}$  represents the demand of customer  $n_s$ ; similarly, the supply flows at the upper echelon ( $s-2$ ) can be estimated by the relationship between the demand flows and the supply flows at echelon ( $s-2$ ), that is,  $\sum_{\forall n_{s-2}} f_{n_{s-2}} = w_{s-2,s-1} w_{s-1,s} \sum_{\forall n_s} f_{n_s}$ ; and the supply flows at echelon ( $s-3$ ) can be expressed as  $\sum_{\forall n_{s-3}} f_{n_{s-3}} = w_{s-3,s-2} w_{s-2,s-1} w_{s-1,s} \sum_{\forall n_s} f_{n_s}$  and so on. Then, the generalized relationship between the total supply flows at echelon  $k$  and the total customer demand in the supply chain network can be formulated as

$$\sum_{\forall n_k} f_{n_k} = \prod_k^s w_{k,k+1} \sum_{\forall n_s} f_{n_s} \quad \forall k \quad (3)$$

where  $\prod_k^s w_{k,k+1}$  represents the total supply flows at echelon  $k$  when the customer demand is one unit. Eq. (3) shows that the total supply flows at the echelons are different and dependent upon customer demand.

The decision maker in this study is a manufacturer who operates multiple high-tech manufacturing plants and serves customers in different regions. Therefore, the upper echelon of the manufacturing plants can be defined as the raw material vendors, and the lower echelon as the customers. This study assumes that the firm's procurement and outsourcing decisions are centralized at the corporate headquarter rather than in the individual manufacturing plants. Let  $\hat{k}$  be the echelon of manufacturing plants and  $n_{\hat{k}}$  represent a specific plant, respectively, while  $\hat{k} - 1$  is the echelon of raw material vendors and  $n_{\hat{k}-1}$  represents a specific vendor, respectively. Consequently, echelon  $\hat{k} + 1$  refers to customers, which can also be represented as echelon  $s$ .

Let  $f_{n_k}^{n_{k-1}}$  and  $f_{n_s}^{n_k}$  represent the amount of raw material shipped from vendor  $n_{k-1}$  to plant  $n_k$  and the amount of final products produced by plant  $n_k$  and transported to customer  $n_s$ , respectively. Let  $f_{n_k}$  be the production amount at plant  $n_k$ , which is constrained by the plant's capacity,  $f_{n_k} \leq v_{n_k}$ . Then, the relationship between the total amount of raw material required by manufacturing plant  $n_k$  and the amount of raw material shipped from all the raw material vendors can be revised according to Eq. (1) as

$$w_{k-1,\hat{k}} f_{n_k} = \sum_{\forall n_{k-1}} f_{n_k}^{n_{k-1}} \delta_{n_k}^{n_{k-1}} \quad \forall n_k \quad (4)$$

Moreover, the relationship between the amount of production produced by plant  $n_k$  and the amount of products shipped from plant  $n_k$  to customer  $n_s$  can be formulated according to Eq. (2), as

$$f_{n_k} = \sum_{\forall n_s} f_{n_s}^{n_k} \beta_{n_s}^{n_k} \quad \forall n_k \quad (5)$$

Since the total customer demand must be satisfied, the following condition must hold, that is to say,  $\sum_{\forall n_k} f_{n_k} = \sum_{\forall n_k} \sum_{\forall n_s} f_{n_s}^{n_k} \beta_{n_s}^{n_k}$ . Furthermore, the capacity utilization of manufacturing plant  $n_k$  can be defined as  $Y_{n_k} = f_{n_k} / v_{n_k}$ , where  $f_{n_k}$  and  $v_{n_k}$  are decision variables in this study.

Costs in this study can be classified as inbound, fixed, production and outbound costs. Inbound costs include raw material purchase and transportation costs, which relates to the movement of the flows from the vendors to the plants. The fixed costs represent all expenses required for the manufacturer to search and contract with active vendors. Production costs incorporate both the capital cost and the variable production cost, where the capital cost includes costs related to the purchasing and installation of related equipments, and plant construction and land rental fee, etc. and differ among plants due to different locations and capacity of the plants. The variable production cost includes those paid for input factors other than raw materials, such as labor, utility and insurance, etc. Outbound costs refer to the costs related to transporting the final products from the plants to the customers.

The production costs of plant  $n_k$ ,  $L_{n_k}$ , can be formulated as follows:

$$L_{n_k} = C(v_{n_k}) + c(v_{n_k}) f_{n_k} \quad \forall n_k \quad (6)$$

where  $C(v_{n_k})$  and  $c(v_{n_k})$  represent, respectively, the capital costs and variable production costs, which depend on capacity,  $v_{n_k}$ . Eq. (6) shows that the production cost increases with the increase in the production amount,  $f_{n_k}$ . Furthermore, the average production cost per unit product of plant  $n_k$  can be expressed as

$$\frac{L_{n_k}}{f_{n_k}} = \frac{C(v_{n_k})}{f_{n_k}} + c(v_{n_k}) \quad \forall n_k \quad (7)$$

On the contrary, Eq. (7) shows that the average production cost per unit product decreases with the increase in production amount; however, the extent depends upon the capital cost with respect to the capacity. Although a manufacturing plant with large-sized capacity could operate efficiently, the manufacturer experiences a high average production cost when the production amount is low, and as a result the high capital cost cannot be absorbed as shown in Eq. (7). Fig. 2 shows the relationship between the average production cost and the production amount for different sizes of capacity, where  $v_{n_k}^1$ ,  $v_{n_k}^2$ ,  $v_{n_k}^3$  are three different-sized capacities for a manufacturing plant  $n_k$ ,  $v_{n_k}^1 < v_{n_k}^2 < v_{n_k}^3$  and  $C(v_{n_k}^1) > C(v_{n_k}^2) > C(v_{n_k}^3)$ .

As shown in Fig. 2, the average production cost per unit product decreases as the production amount for each of the three sizes of capacity increases. And, when the

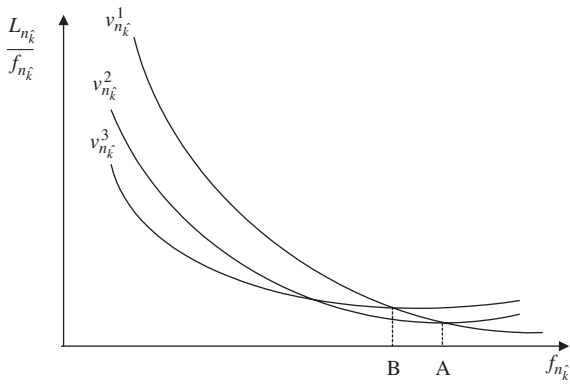


Fig. 2. The relationship between the average production cost and the production amount.

production amount is low, the largest-size capacity  $v_{n_k}^1$  yields the highest average production cost among the three different-sized capacities. Nevertheless, as the production expands, the average production cost for the largest-size capacity  $v_{n_k}^1$  decreases with a higher decreasing rate. When the production amount exceeds point A in Fig. 2, the average production cost for the largest-size capacity  $v_{n_k}^1$  reaches the lowest among the three different-sized capacities, implying that a manufacturing plant with the largest-size capacity  $v_{n_k}^1$  is the most advantageous in terms of saving costs. Since the production amount depends on the demand, a large-sized capacity is recommended when there is a large demand.

For a manufacturer operating multiple plants, the impact of economies of scale on the total average production cost depend not only on the total customer demand, but also on the production assignments among manufacturing plants. Furthermore, the total average production cost per unit product for the manufacturer,  $H_{\hat{k}}$ , can be formulated as

$$H_{\hat{k}} = \frac{\sum_{v_{n_k}} L_{n_k}}{\sum_{v_{n_k}} f_{n_k}} = \frac{\sum_{v_{n_k}} C(v_{n_k}) + c(v_{n_k})f_{n_k}}{\sum_{v_{n_k}} f_{n_k}} \quad (8)$$

The denominator in Eq. (8) is fixed and given due to the fact that the total amount of production is constrained in order to satisfy the total customer demand.

Let  $V_{n_{k-1}}$  be the fixed cost of the manufacturer with raw material vendor  $n_{k-1}$ , which is independent as to the amount of raw material procured and the total number of plants served since the cost is mainly the results of the manufacturer searching and contracting with the vendor. Then, the total fixed cost for the manufacturer can be expressed as  $\sum_{v_{n_{k-1}}} V_{n_{k-1}} \gamma_{n_{k-1}}$ , where  $\gamma_{n_{k-1}}$  is an indicator variable representing whether vendor  $n_{k-1}$  is an active vendor for the manufacturer, and where  $\gamma_{n_{k-1}}$  can be obtained based on the results of  $\delta_{n_k}^{n_{k-1}}$  for all manufacturing plants, that is,  $\gamma_{n_{k-1}} = \min\{1, \sum_{v_{n_k}} \delta_{n_k}^{n_{k-1}}\}$ . If  $\gamma_{n_{k-1}}$  is equal to 1, then vendor  $n_{k-1}$  is an active vendor for the manufacturer; otherwise,  $\gamma_{n_{k-1}}$  is equal to 0. Then, the

average fixed cost per unit raw material for the manufacturer can be formulated as

$$V_{\hat{k}-1} = \frac{\sum_{v_{n_{k-1}}} V_{n_{k-1}} \gamma_{n_{k-1}}}{\sum_{v_{n_{k-1}}} \sum_{v_{n_k}} f_{n_k}^{n_{k-1}} \delta_{n_k}^{n_{k-1}}} \quad (9)$$

The unit of Eq. (9) is measured based on the raw material. To be based on products, Eq. (9) can be revised according to Eq. (3), and shown as

$$\bar{V}_{\hat{k}-1} = \left( \prod_{k-1}^s w_{k,k+1} \right) V_{\hat{k}-1} \quad (10)$$

where  $\bar{V}_{\hat{k}-1}$  represents the average fixed cost per unit product for the manufacturer.

With regards to the raw material purchase costs, the main influences affecting average raw material purchase cost are unit raw material production cost and a reasonable payment ratio determined by the vendor. Regardless of the reasonable payment ratio, the unit raw material purchase cost is low when the market for that raw material is high due to the fact that there exist economies of scale. To simplify the problem, this study assumes two influences, i.e. unit raw material production cost and payment ratio, are exogenous, and denote unit raw material purchase price from vendor  $n_{k-1}$  by  $p_{n_{k-1}}$ . Therefore, the purchase price paid to vendor  $n_{k-1}$  can be shown as  $p_{n_{k-1}} \sum_{v_{n_k}} f_{n_k}^{n_{k-1}} \delta_{n_k}^{n_{k-1}}$ . Summing up all the purchase costs paid to all the vendors, the total raw material purchase cost for the manufacturer can be formulated as

$$\sum_{v_{n_{k-1}}} p_{n_{k-1}} \sum_{v_{n_k}} f_{n_k}^{n_{k-1}} \delta_{n_k}^{n_{k-1}} \quad (11)$$

The transportation costs in this study capture the costs resulting from the spatial distance between two locations. High-tech products are usually characterized as having a high market value and depreciate quickly. A nearby raw material vendor is usually selected, since the raw material delivery time is short and as a result the inventory can be kept at a low level. Transportation costs decrease if the plant and the vendor are located in the same location because the distance is short and a low-cost transportation mode, i.e. trucks, can be employed. Consequently transportation costs are high if the locations of the plant and the vendor are a long distance apart from each other. Let  $t_{n_k}^{n_{k-1}}$  represent the average unit-distance transportation cost per unit raw material between the locations of vendor  $n_{k-1}$  and plant  $n_k$ . The transportation cost for transporting raw material from vendor  $n_{k-1}$  to plant  $n_k$  can be expressed as  $d_{n_k}^{n_{k-1}} t_{n_k}^{n_{k-1}} f_{n_k}^{n_{k-1}}$ , where  $d_{n_k}^{n_{k-1}}$  and represents the average distance from the location of vendor  $n_{k-1}$  to plant  $n_k$ . Then, the total raw material transportation cost for the manufacturer can be formulated as

$$\sum_{v_{n_k}} \sum_{v_{n_{k-1}}} d_{n_k}^{n_{k-1}} t_{n_k}^{n_{k-1}} f_{n_k}^{n_{k-1}} \delta_{n_k}^{n_{k-1}} \quad (12)$$

Eq. (12) shows that the transportation costs vary with the different combinations of  $(n_{k-1}, n_k)$  due to the different

distances and average transportation cost per unit of raw material between vendor  $n_{k-1}$  and plant  $n_k$ .

Summing up Eqs. (11) and (12), the total inbound cost for the manufacturer can be shown as

$$\begin{aligned} & \sum_{\forall n_{k-1}} \sum_{\forall n_k} p_{n_{k-1}} f_{n_k}^{n_{k-1}} \delta_{n_k}^{n_{k-1}} + \sum_{\forall n_k} \sum_{\forall n_{k-1}} d_{n_k}^{n_{k-1}} t_{n_k}^{n_{k-1}} f_{n_k}^{n_{k-1}} \delta_{n_k}^{n_{k-1}} \\ & = \sum_{\forall n_{k-1}} \sum_{\forall n_k} (p_{n_{k-1}} + d_{n_k}^{n_{k-1}} t_{n_k}^{n_{k-1}}) f_{n_k}^{n_{k-1}} \delta_{n_k}^{n_{k-1}} \end{aligned} \quad (13)$$

Dividing Eq. (13) by the total amount of raw material supplied by raw material vendors,  $\sum_{\forall n_{k-1}} f_{n_{k-1}}$ , the average inbound cost per unit of raw material for the manufacturer,  $H_k^{k-1}$ , can be formulated as

$$H_k^{k-1} = \frac{1}{\sum_{\forall n_{k-1}} f_{n_{k-1}}} \sum_{\forall n_{k-1}} \sum_{\forall n_k} (p_{n_{k-1}} + d_{n_k}^{n_{k-1}} t_{n_k}^{n_{k-1}}) f_{n_k}^{n_{k-1}} \delta_{n_k}^{n_{k-1}} \quad (14)$$

The average inbound cost per unit product for the manufacturer can be obtained as  $\bar{H}_k^{k-1} = (\prod_{k=1}^s w_{k,k+1}) H_k^{k-1}$ .

As suggested by Chopra (2003), products with a high value are suitable for a delivery network with direct shipping, that is, products are shipped directly from the plant to the customer. In addition, manufacturing plants with a large production can serve many customers, yet this may lead to high transportation costs. To provide better service and reduce transportation costs, the plants are advised to serve nearby customers. The outbound cost for the manufacturer can be expressed as  $\sum_{\forall n_k} \sum_{\forall n_s} d_{n_s}^{n_k} t_{n_s}^{n_k} f_{n_s}^{n_k} \beta_{n_s}^{n_k}$ , where  $t_{n_s}^{n_k}$  represents the average unit-distance transportation cost per unit product between the location of the plant  $n_k$  and customer  $n_s$ . Moreover, the average outbound cost per unit product for manufacturer,  $T_s^k$ , can be shown as

$$T_s^k = \frac{1}{\sum_{\forall n_k} f_{n_k}} \sum_{\forall n_s} \sum_{\forall n_k} d_{n_s}^{n_k} t_{n_s}^{n_k} f_{n_s}^{n_k} \beta_{n_s}^{n_k} \quad (15)$$

where  $\sum_{\forall n_k} f_{n_k}$  represents the total production amount. Since the total customer demand must be satisfied,  $\sum_{\forall n_k} f_{n_k} = \sum_{\forall n_s} f_{n_s}$ .

The total average cost per unit product for the manufacturer is the sum of the average inbound, fixed, production and outbound costs in the entire supply chain, and can be formulated as

$$\bar{V}_{k-1} + \bar{H}_k^{k-1} + H_k + T_s^k \quad (16)$$

From the discussions above, the nonlinear MIP model for the supply chain network design can be formulated as follows:

$$\text{Min } \bar{V}_{k-1} + \bar{H}_k^{k-1} + H_k + T_s^k \quad (17a)$$

s.t.

$$w_{k-1,k} f_{n_k} = \sum_{\forall n_{k-1}} f_{n_{k-1}}^{n_k} \delta_{n_k}^{n_{k-1}}, \quad \forall n_k \quad (17b)$$

$$f_{n_k} = \sum_{\forall n_s} f_{n_s}^{n_k} \beta_{n_s}^{n_k} \quad \forall n_k \quad (17c)$$

$$\sum_{\forall n_k} f_{n_k} = \sum_{\forall n_k} \sum_{\forall n_s} f_{n_s}^{n_k} \beta_{n_s}^{n_k} \quad (17d)$$

$$\sum_{\forall n_k} f_{n_k}^{n_{k-1}} \delta_{n_k}^{n_{k-1}} \leq S_{n_{k-1}} \quad \forall n_{k-1} \quad (17e)$$

$$Y_{n_k} = \frac{f_{n_k}}{v_{n_k}}, \quad \forall n_k \quad (17f)$$

$$v_{n_k}, f_{n_s}^{n_k}, f_{n_k}^{n_{k-1}} \geq 0 \quad \forall n_k \quad \forall n_{k-1} \quad (17g)$$

$$\delta_{n_k}^{n_{k-1}} = 0 \text{ or } 1 \quad \forall n_k \quad \forall n_{k-1} \quad (17h)$$

$$\beta_{n_s}^{n_k} = 0 \text{ or } 1 \quad \forall n_k \quad \forall n_{k-1} \quad (17i)$$

Eq. (17a) is the objective function that minimizes the total average cost per unit product. Eq. (17b) states that the amount of raw material requested by plant  $n_k$  is the sum of the amount of raw material provided by its active vendors. Eq. (17c) defines that the amount of products shipped from plant  $n_k$  to the customers is equal to the amount of production. Eq. (17d) constrains the total production amount to meeting the total customer demand. Eq. (17e) is the supply limit of raw material vendor  $n_{k-1}$ , where  $S_{n_{k-1}}$  represents the maximum amount of raw material supplied by vendor  $n_{k-1}$ . Eq. (17f) defines the capacity utilization of manufacturing plant  $n_k$ . Eq. (17g) constrains the decision variables  $v_{n_k}, f_{n_s}^{n_k}$  and  $f_{n_k}^{n_{k-1}}$  to be non-negative. Finally, Eqs. (17h) and (17i) define the decision variables  $\delta_{n_k}^{n_{k-1}}$  and  $\beta_{n_s}^{n_k}$  to be binary. The decision variables are  $v_{n_k}, f_{n_k}, f_{n_s}^{n_k}, f_{n_k}^{n_{k-1}}, \delta_{n_k}^{n_{k-1}}$  and  $\beta_{n_s}^{n_k}$ . That is, the manufacturer can apply the model to optimally decide the size of the capacity as well as the production amount for all manufacturing plants, the amount of raw material from the vendors to plants, and which manufacturing plants should produce how much production to serve customers in different regions. Furthermore, the optimal capacity utilization of plants and the optimal number of active raw material vendors for the manufacturer can also be obtained from the model.

### 3. Algorithm

This study have made trial runs to examine the difficulty of solving the problem. The solutions of the manufacturing echelon include not only the capacity of plants but also the production amount of each plant. There are various combinations of wafer fabs for plants. To verify the optimal production allocation, the study tests a variety of initial values of production amounts of plants. After several trials, we found that it is hard and time-consuming to find an optimal solution. Though the problem can be solved easily by linearizing the nonlinear relationship in the model, the impacts of economies of scale on the optimal solutions may not be explicitly captured. Due to the complexity in solving a nonlinear programming problem, approximate methods are required. The available heuristic approaches include primal-dual, genetic algorithms (GA) and simulated

annealing (SA), etc. The main advantage of GA is less susceptible to getting stuck at local optima than gradient search methods. However, adapting GA tends to be computationally expensive (Mishra et al., 2003). In the study, SA algorithm proposed by Kirkpatrick et al. (1983) is applied due to its simple implementation and speed. The simulated annealing (SA) is a method for obtaining good solution in difficult optimization problems (Paik and Soni, 2007). There have been studies in various fields applying the SA-based heuristics to solve the problems (e.g., Dowsland et al., 2007; Paik and Soni, 2007; Yan and Luo, 1999; Loukil et al., 2007). The SA algorithm is based on Metropolis et al. (1953), which was originally proposed as a means of finding the equilibrium configuration of a collection of atoms at a given temperature. It has been extensively used in solving very large-scale integration (VLSI) layout and graph partitioning problems. The major advantage of the SA algorithm over other local search methods is the ability to avoid becoming trapped in the local optimal. The SA algorithm employs a random search, which not only accepts changes that decrease the objective function, but also accepts some changes that increase it. And, the latter are accepted with a probability of Boltzmann distribution. In this section, we first develop an approach to generate an initial solution, and then use the SA algorithm to develop the heuristic for improving the initial solution.

### 3.1. Initial solution (INIT)

Since local improvement methods must start from a feasible solution, this study develops a heuristic to generate good initial solutions. Based on the characteristics of production with economies of scale, the average production cost per unit product may be reduced if there is more production assigned to a plant with larger-sized capacity. An incremental rule, in which the production amount is incrementally assigned to manufacturing plants until total customer demands are satisfied, might be used to investigate the relationship between the assigned production amount among various plants and the total average cost per unit product. Besides, the manufacturing echelon is the most value-added echelon for high-tech products, and the production cost is usually high in the high-tech manufacturing supply chain. In other words, the total cost cannot reach a minimum without a well-designed production plan for the various plants. The heuristics is described as follows.

*Step 1.* Randomly determine the capacity for plant  $v_{n_k}$ , for all  $n_k$ , such that the sum of the capacity of all plants must exceed the total customer demand,  $\sum_{v_{n_k}} v_{n_k} \geq \sum_{v_{n_s}} f_{n_s}$ . Set a value,  $m$ , representing the incremental amount of production that can be assigned to the plants at each iteration;

*Step 2.* Assign the amount of production for the plant;

2.1. Calculate the total average production cost per unit product for manufacturer,  $H_k$ , when the incremental amount of production,  $m$ , is assigned to manufacturing plant  $n_k$ ;

2.2. Find an optimal plant  $n_k^*$  with the minimum value of  $H_k$ . Assign the incremental amount of production,  $m$ , to plant  $n_k^*$ , and update the production amount for plant  $n_k^*$ ,  $f_{n_k^*} = f_{n_k^*} + m$ ;

2.3. Calculate the remaining customer demand and the unfulfilled capacity for each of the manufacturing plants;

*Step 3.* If the demands of all customers are met, then go to Step 4 and output the optimal production amount and the respective amount of raw material requested by the plants; else go to Step 2;

*Step 4.* Determine the active vendors, the amount of raw material from the vendors to the plants and the amount of products from the plants to the customers;

4.1. Set the values  $u$  and  $v$  representing the incremental amount of raw material shipped from the vendors to the plants and the incremental amount of products shipped to the customers from the plants, respectively;

4.2. Calculate  $H_k^{k-1} + \bar{V}_{k-1}$  and  $T_s^k$  when  $u$  and  $v$  are, respectively, assigned to links  $(n_{k-1}, n_k)$  and  $(n_k, n_s)$ ;

4.3. Find the optimal links  $(n_{k-1}, n_k)^*$  and  $(n_k, n_s)^*$  with the minimum values of  $H_k^{k-1} + \bar{V}_{k-1}$  and  $T_s^k$ , respectively. Assign  $u$  to link  $(n_{k-1}, n_k)^*$  and  $v$  to link  $(n_k, n_s)^*$  and update the flows between the links using  $f_{n_k^{k-1}} = f_{n_k^{k-1}} + u$  and  $f_{n_s} = f_{n_s} + v$ ;

4.4. Calculate (1) the unfulfilled amount of raw material of the plants and (2) the unsatisfied demand of the customers;

*Step 5.* If the customer demands and the amount of raw material requested by all plants are met, then stop the algorithm; else go to Step 4.

The complexity of determining the capacity of each plant in Step 1 is  $O(1)$ . Note that in each iteration, the total average production cost per unit product has to be calculated when the incremental production amount is assigned to the plants. The number of iterations is positively related to the setting of incremental production amount  $m$ , given the plant capacity. The computation times are longer with a smaller setting of  $m$  as compared with the capacity of plants. Let  $M$  be the total number of the plants. Then, the complexity of *INIT* is  $O(M^2)$ .

### 3.1. Simulated annealing (SA)

The values of the SA algorithm parameters are obtained by classical experiments, i.e. trial and error tests, and the values are also compared to those in Yan and Luo (1999) and Loukil et al. (2007). The parameters include (1) down\_hill\_move ratio  $\leq 0.5$ , where down\_hill\_move ratio is determined by the number of inferior solutions divided by the number of moves; (2) accept\_ratio  $\leq 0.5$ , where the accept\_ratio is obtained by the number of current accepted solutions divided by the number of moves; (3) initial temperature  $T_0 = 99$ , decreasing ratio of temperature is 0.99 and the stop temperature  $T = 0.1$ ; (4) the maximum number of moves at each temperature = 100;

and (5) the maximum number of down\_hill\_moves at each temperature = 50. Conditions (1)–(3) are the stop criterions for the SA. Conditions (4) and (5) are the stop criterions for the Metropolis algorithm (Metropolis et al., 1953). Referring to Heragu and Alfa (1992) and Yan and Luo (1999), the SA algorithm can be described as follows.

*Step 0.* Employ *INIT* to find an initial feasible solution,  $S$ , and calculate its objective value,  $z(S)$ , where

$$z = \bar{V}_{k-1} + \bar{H}_k^{k-1} + H_k + T_s^k.$$

*Step 1.* At temperature  $T_x$ , implement the Metropolis algorithm;

1.1. Randomly choose a plant and alter its capacity from the initial solution. Apply Steps 2–5 in *INIT* to find a good adjacent solution  $S'$  and calculate its objective value,  $z(S')$ ;

1.2. Determine whether the new solution is accepted;

1.2.1 Calculate the difference between the objective function of  $S$  and  $S'$ ,  $\Delta = z(S') - z(S)$ .

1.2.2 If  $\Delta \leq 0$ , then  $S = S'$ ; else randomly generate a variable  $y \sim U(0,0.99)$ . If  $\exp^{-\Delta/T_x} \geq y$ , then  $S = S'$ ; else go to Step 1;

1.2.3 If the stop criterions of the Metropolis algorithm are satisfied, then go to Step 2; else go to Step 1;

*Step 2.* If the stop criterions of the SA algorithm are satisfied, then go to Step 3; else let  $x = x+1$  and  $T_{x+1} = 0.99T_x$ , and go to Step 1;

*Step 3.* Output the optimal sizes of capacity as well as the production amount for all plants, i.e.  $v_{n_k}^*$  and  $f_{n_k}^*$ , the active raw material vendors as well as the amount of raw material from the raw material vendors to the plants, i.e.  $\delta_{n_k}^{*n_{k-1}}$  and  $f_{n_k}^{*n_{k-1}}$  and also which plants along with their allocated production volumes to serve the customers different regions, i.e.  $f_{n_{k+1}}^{*n_k}$ , and  $\beta_{n_s}^{*n_k}$ , respectively.

#### 4. Case study

A case study of T-company, which specializes in wafer foundry in the semiconductor industry and has its headquarters in Taiwan, is used herein to demonstrate the application of the proposed models. The final product of T-company are dies, which represents the starting form of

an integrated circuit (IC) and can be produced in either 12-, 8- or 6-inch wafers. Because some of T-company's operating costs and customer demand data are unavailable, the annual report data in Taiwan Semiconductor Manufacturing Company (TSMC) (2004) were employed to estimate them. Regarding customer demand, T-company has customers from six major areas, North America, Taiwan, Europe, Japan, Korea and Hong Kong, and the monthly customer demand for dies for the coming year totals approximately  $2 \times 10^8$ . The demands from these customers in six areas are  $6.4 \times 10^7$ ,  $5.7 \times 10^7$ ,  $3.6 \times 10^7$ ,  $3.2 \times 10^7$ ,  $8.4 \times 10^6$  and  $7.5 \times 10^6$  dies, respectively. The alternatives of plant size are 12-, 8- and 6-inch FABs, and they produce an average of 40,000, 35,000 and 30,000 pieces per month, respectively. Note that each size of FAB can only produce its particular size of wafers due to the complexity of the technology employed in the manufacturing process of wafers. One piece of 12-inch wafer is 2.25 times the square area of an 8-inch wafer; and 4 times that of a 6-inch wafer. Regardless of the yield, the number of dies produced by one piece of 12-, 8- and 6-inch wafer are 1233, 514 and 210, respectively, based on the 0.11  $\mu\text{m}$  process technology. To unify, the capacities of 12-, 8- and 6-inch FABs are revised based on the number of dies produced by one piece of wafer, that is  $4.9 \times 10^7$ ,  $1.8 \times 10^7$  and  $6.3 \times 10^6$  dies, respectively. Therefore, constructing a 12-inch FAB is more beneficial for satisfying customer demand in terms of dies as compared with the other two sizes of FABs.

The alternative locations to build the FABs are Taiwan (Hsinchu), Taiwan (Tainan), Shanghai, USA and Singapore. The capital cost per month for different-sized FABs can be estimated by the total costs for the FAB construction plus equipment set-up and the maximum usage period of the FAB. The data on the capital cost and the variable production costs for different-sized FABs in different regions are listed in Table 1.

As shown in Table 1, the capital and the variable production costs for different-sized FABs for some locations, i.e. USA and Singapore are higher than in others, because of higher commodity price indexes in these regions. Moreover, the costs for T-company to operate a 12-inch FAB is higher than the other two sizes, due to the high capital and variable production costs as shown in Table 1. Considering the final products of T-company are dies, the average production cost per unit product of plant  $n_k$  as shown in Eq. (7), can be further revised in terms of

**Table 1**

The alternative sizes of FAB and base production parameters for plants in different locations.

Plant in different location, $n_k$	6-inch		8-inch		12-inch	
	Capital cost ( $10^3$ US\$)	Variable production cost (US\$/wafer)	Capital cost ( $10^3$ US\$)	Variable production cost (US\$/wafer)	Capital cost ( $10^3$ US\$)	Variable production cost (US\$/wafer)
Taiwan (Hsinchu)	1865	205	3000	323	10,000	515
Taiwan (Tainan)	1850	204	2978	321	9900	513
Shanghai	1900	208	3005	327	10,032	520
USA	2100	212	3085	335	10,090	523
Singapore	2030	215	3078	335	10,065	525



dies, which yields

$$\frac{L_{n_k}}{f_{n_k} u(F_{n_k})} = \frac{C(v_{n_k})}{f_{n_k} u(F_{n_k})} + \frac{c(v_{n_k})}{u(F_{n_k})} \tag{18}$$

where  $u(F_{n_k})$  represents the number of dies that one piece of wafer produces when the size of FAB for plant  $n_k$ , is  $F_{n_k}$ , and  $F_{n_k} = 6$ -, 8- and 12-inch, respectively. For example, if T-company decides to operate a 12-inch FAB in the USA, then  $u(12_{USA}) = 1233$ . Taking the base production parameters of Taiwan (Hsinchu) as an example, the relationship between the average production cost per die and the production amount for different-sized FABs can be further explored, and is shown in Fig. 3.

As shown in Fig. 3, the average production cost per die for different-sized FABs decreases as the production amount increases, but at different rates. The average production cost decreases at an increasing rate when the production amount is rather small; however, as the production becomes larger, the average production cost exhibits a constant number. In addition, all different-sized FABs are characterized as having a high production cost due to low-capacity utilization when the production amount is small, such as less than  $5 \times 10^6$  dies. However, the average production cost per die of a 12-inch FAB is the highest when the least capacity utilization is combined with the most expensive capital cost. Fig. 3 also shows that there is an advantage of an 8-inch FAB over a 6-inch FAB when the production amount exceeds  $3 \times 10^6$  dies. Even though the curve representing the average production cost per die of a 12-inch FAB lies above those of the 6- and 8-inch FABs as shown in Fig. 3, the 6- and 8-inch FABs cannot satisfy the market due to capacity constraints when the demand exceeds  $6.3 \times 10^6$  and  $1.8 \times 10^7$  dies, respectively. Furthermore, as the production amount of a

12-inch FAB approaches its full-capacity, a 12-inch FAB yields the lowest cost among all those with full-capacity production, implying there are economies of scale in the wafer foundry industry.

The main raw materials for producing wafers include silicon wafers, chemical source, photoresist and specialty gases (TSMC, 2004). To simplify the study, chemicals are selected as a raw material. The amount of chemicals required to produce one piece of 6-, 8- and 12-inch wafers can be estimated as 0.585, 0.475 and 0.450 L, respectively (TSMC, 2004). Considering the difference in the amount of chemical resource required in producing one piece of different-sized wafers, then the total amount of chemical source required by plant  $n_k$  can be further revised as  $w_{k-1,k}(F_{n_k})f_{n_k}$ , where  $w_{k-1,k}(F_{n_k})$  represents the amount of chemical source required to produce one piece of wafer when the size of FAB for plant  $n_k$  is  $F_{n_k}$ . There are four chemical source vendors in the market, namely, Merck, Chem Sources (CS), Tai-Young High Tech (TY) and Chemical Sources International (CSI); both Merck and TY have plant sites in Taiwan, while CS and CSI are located in the USA. Table 2 shows the initial values of the base procurement parameters.

This study further assumes that the transportation cost per unit flow shipped is measured based on the weight and distance. The weight of one die after packaging can be approximately estimated as weighing 300 mg, i.e.  $3 \times 10^{-4}$  kg. According to the Taiwan Institute of Economic Research (TIER) (2004), the average transportation charges per kg are approximately US\$2.8–4. The transportation cost per kg between two locations can be further estimated by unit-distance transportation cost per kg and the distance between them, and the difference in unit-distance transportation cost between two locations is due to the transportation mode employed. Tables 3(a)

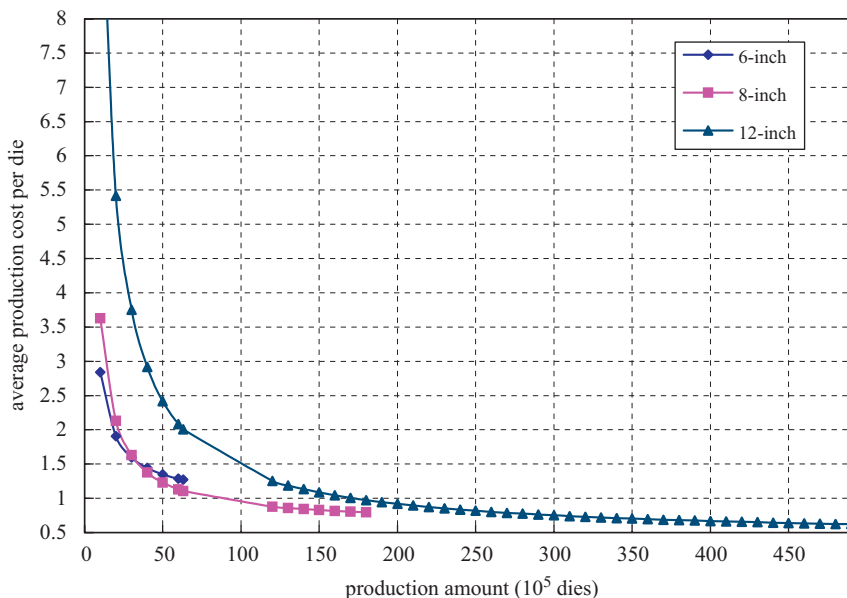


Fig. 3. Average production cost per die vs. production amount.

and (b) show the transportation cost per kg between locations of chemical source vendors and plants, and between locations of plants and customers in six areas, respectively.

The model is programmed using Visual C++, a computer-modeling program developed by Microsoft, based on the developed heuristic algorithm. Tables 4–6 summarize the initial solution results.

In the case study, since Taiwan's government provided incentives for developing high-tech industry, and since there is a large local customer demand due to economies of agglomeration in the semiconductor industry, T-company chose to construct 12-inch FABs in Taiwan, and they are located in Hsinchu and Tainan. Due to the large demand, the T-company also operates 12-inch FABs in Shanghai and the USA, respectively, as well as an 8-inch FAB in Singapore, as shown in Table 4. The capacity utilizations of the four 12-inch FABs are 100%, while that of the 8-inch FAB is 42%. Because of the low capacity of a

6-inch FAB, it is not employed when there is large demand. This can also be explained by the fact that the T-company expanded its capacity to operate more 12-inch FABs rather than 6-inch FABs. Table 4 also shows that four 12-inch FABs have the lowest average production cost per die, approximately US\$0.62, while it is 1.06 US\$/die for an 8-inch FAB. In addition, the total average production cost per die for T-company is US\$0.65. These results imply that because of the high customer demand, the manufacturer can operate its plants with large-sized capacity, combined with full-capacity production, thereby lowering the production cost. These results also imply that when determining the production amount for multiple plants, plants with large-sized capacity have a high priority over others for filling the capacity, not only due to the high capability of satisfying the customer demand but also because they provide greater cost savings. Finally, Table 4 shows the amount of chemical supplies required by each manufacturing plant, with the total amount required being 100,642 L.

Table 5 shows the initial results of the procurement decisions including the optimal active chemical source vendors, the procured amount of chemicals, and the amount of chemicals shipped from the vendors to the plants. As shown in Table 5, the optimal active vendors include Merck, TY and CSI, and each serves different plants. This is because the unit chemical source purchase costs offered by these vendors are relatively low. Although there is a high fixed cost with CSI, this high fixed cost per unit chemical source is reduced if the procurement amount is large. The advantage from this low unit cost outweighs the disadvantage of the high fixed cost. Since the distance between two alternatives can be reflected by the transportation cost, active vendors tend to serve the plants nearby. With reference to Table 3(a), the distance from vendors Merck and TY to the two plants in Taiwan are due to the fact that they are all located in Taiwan; therefore, FABs in Taiwan are mainly served by Merck and TY rather than by CSI. The average inbound cost per die is US\$4.63  $\times 10^{-3}$ .

Table 6 shows the initial results of the relationship between the plants and customers in six areas. Since in this study the customer is not constrained to be served by one single plant, some customers are served by more than one manufacturing plant, such as customers in North America, Taiwan, Europe and Korea as shown in Table 6. For example, customers in North America are served by

**Table 2**

The initial values of base procurement parameters.

Chemical source vendor, $n_{k-1}$	Fixed cost, $V_{n_{k-1}}$ (US\$)	Unit chemical source purchase price, $p_{n_{k-1}}$ (US\$/L)	Maximum amount supplied, $S_{n_{k-1}}$ (L)
Merck	53	5.50	40,000
CS	98	5.55	45,000
TY	82	5.51	37,500
CSI	112	5.49	41,000

**Table 3(a)**

The transportation cost per kg between locations of chemical source vendors and plants.

Manufacturing plant, $n_k$	Chemical sources vendor, $n_{k-1}$			
	Merck	CS	TY	CSI
Taiwan (Hsinchu)	2.8	6.1	2.8	4.9
Taiwan (Tainan)	2.9	6.2	3.0	4.4
Shanghai	4.1	6.6	4.5	5.5
USA	3.5	4.3	4.4	4.9
Singapore	5.5	5.1	5.0	5.1

**Table 3(b)**

The transportation cost per kg between manufacturing plants and customers in six areas.

Customer in different areas, $n_c$	Manufacturing plant, $n_k$				
	Taiwan (Hsinchu)	Taiwan (Tainan)	Shanghai	USA	Singapore
North America	5.5	5.8	6.4	3.2	6.0
Taiwan	2.6	2.7	2.9	6.2	3.4
Europe	6.1	6.2	6.3	4.5	6.0
Japan	3.3	3.3	3.9	4.8	4.3
Korea	3.2	3.3	4.0	4.7	4.4
Hong Kong	2.9	3.0	2.7	5.0	4.0

Unit: US\$/kg.

Unit: US\$/kg.

**Table 4**  
The initial results of the plants.

Location of the plant	Taiwan (Hsinchu)	Taiwan (Tainan)	Shanghai	USA	Singapore
Size of FAB, $F_{n_k}$	12-inch	12-inch	12-inch	12-inch	8-inch
Capacity, $v_{n_k}$ (wafer)	40,000	40,000	40,000	40,000	35,000
Production amount, measuring in wafer, $f_{n_k}$	40,000	40,000	40,000	40,000	14,825
Production amount, measuring in die, $f_{n_k} u(F_{n_k})$	$4.9 \times 10^7$	$4.9 \times 10^7$	$4.9 \times 10^7$	$4.9 \times 10^7$	$7.6 \times 10^6$
Chemical sources required, $w_{k-1,k}(F_{n_k})/f_{n_k}$ (L)	23,400	23,400	23,400	23,400	7042
Capacity utilization, $Y_{n_k}$	100%	100%	100%	100%	42%
Average production cost per die (US\$/die)	0.62	0.62	0.63	0.63	1.06
Total production amount	$2 \times 10^8$ (Dies)				
Total average production cost per die	0.65 (US\$/die)				
Total amount of chemical source required	100,642 (L)				

**Table 5**  
The initial results of the relationship between the plants and chemical source vendors.

Active vendor	Plants served	Amount of chemical sources shipped, $f_{n_k}^{n_{k-1}}$ (L)	Total amount of chemical sources supplied, $\sum v_{n_k} f_{n_k}^{n_{k-1}} \delta_{n_k}^{n_{k-1}}$ (L)
Merck	Taiwan (Tainan)	9300	40,000
	Shanghai	7300	
	USA	23,400	
TY	Taiwan (Hsinchu)	23,400	37,500
	Taiwan (Tainan)	14,100	
CSI	Shanghai	16,100	23,142
	Singapore	7042	

The sum of average fixed and inbound cost per die  $4.63 \times 10^{-3}$  (US\$/die).

FABs in Taiwan (Hsinchu) and USA, while customers in Hong Kong are served by the FAB in Shanghai. In Taiwan, the majority of customers are served by FABs in Taiwan due to their relative low transportation and production costs. In addition to the amount of customer demand, the main reason that customers are served by different FABs lies in the distance between the customer and the FAB. For example, customers in North America are served by the FAB in the USA, while customers in Taiwan are served by the FAB in Taiwan. These results imply that to reduce the outbound cost, the product should be shipped from a plant to a customer with the shortest distance between them. These results also imply that for a wafer foundry company the benefits brought about by centralized production are larger than the increased transportation costs by decentralized production. The manufacturer may adopt a production strategy with centralized production in plants with large-sized capacity and then ship the product to customers in different regions. Summing up the average production, the inbound and outbound costs

**Table 6**  
The initial results of the relationship between the plants and customers in six areas.

Customer in different areas, $n_s$	Customer demand, $f_{n_s}$ (dies)	Manufacturing plants, $n_k$	Amount of products shipped, $f_{n_s}^{n_k}$ (dies)
North America	$6.4 \times 10^7$	Taiwan (Hsinchu)	$1.5 \times 10^7$
		USA	$4.9 \times 10^7$
Taiwan	$5.7 \times 10^7$	Taiwan (Tainan) Shanghai	$1.5 \times 10^7$ $4.2 \times 10^7$
Europe	$3.6 \times 10^7$	Taiwan (Hsinchu) Singapore	$2.8 \times 10^7$ $7.6 \times 10^6$
Japan	$3.2 \times 10^7$	Taiwan (Tainan)	$3.2 \times 10^7$
Korea	$8.4 \times 10^6$	Taiwan (Hsinchu)	$6.3 \times 10^6$
		Taiwan (Tainan)	$2.1 \times 10^6$
Hong Kong	$7.5 \times 10^6$	Shanghai	$7.5 \times 10^6$

Average outbound cost per die  $1.16 \times 10^{-3}$  (US\$/die).

per die in Tables 4–6, the total average cost per die of T-company can be calculated as US\$0.65579, with the portion of production cost being 99%, the highest of them all. This implies that the wafer foundry industry shows production with economies of scale, and that the production is the most valued-added in the entire supply chain. Therefore, the manufacturer must be aware of the impact on the total cost of capacity utilization of plants with different-sized capacity.

So far, this study has conducted a numerical example for T-company specializing in wafer foundry in the semiconductor industry. Next, this study will further explore the influences of changes in key parameters on

the average production cost, and the optimal capacity and optimal production amount of the plants.

In this study, the capacity utilization of different-sized FABs significantly influences the total average production cost per die. As Table 4 shows, a minimum cost is yielded for T-company if four 12-inch FABs reach their full-capacity production. This study takes FABs in the USA and Singapore as an example to explore the impact of the production amount of different-sized FABs on the total average production cost per die. As shown in Table 4, the total production amount of FABs in the USA and Singapore are  $5.66 \times 10^7$  dies, with the FAB size in the USA and Singapore as 12- and 8-inch, respectively. Under the assumption that the FAB size of other plants and their production amount remain the same, the number of dies produced by the USA and Singapore are negatively related. Fig. 4 shows the total average production cost per die vs. the production amount in the USA.

As shown in Fig. 4, as the production amount produced by the USA increases, the average production cost per die for both USA and T-company decreases, which also includes Singapore for exhibiting increased average production cost. This is because the number of dies produced by one piece of 12-inch wafer exceeds that of an 8-inch wafer. Also, the impact of the increased cost in Singapore on the total cost can be offset by the decreased cost in the USA. In other words, the total production cost can be reduced if more production is assigned to a plant with large-sized capacity. This result also implies that with the existence of economies of scale, the mechanism for determining the production amount for the plants with different-sized capacity is to assign the most demands to plants with the largest-sized capacity, and to assign the remaining ones to those with small sized. However, in addition to the size of capacity, the other important factor affecting the assignments of production is the amount of customer demand. As stated, when the customer demand is small, using manufacturing plants

with large-sized capacity may lead to high production cost. Consequently, a plant with a large-sized capacity is preferred when the demand is large. Table 7 shows the optimal sizes of FABs and the capacity utilization of the plants with different amounts of customer demand.

As shown in Table 7, there is a high total average production cost when the amount of customer demand is low. Referring to Table 1, the advantages of Hsinchu and Tainan in Taiwan are low capital and variable production costs, which provided T-company with the incentives to construct two 12-inch FABs in those regions, as shown in Table 7. The optimal FAB size is 12-inch for Hsinchu and Tainan with full-capacity production regardless of the customer demand. This result implies that the determination of where to operate a plant with large-sized capacity lies in the labor costs involved as well as the corresponding land rental, expenses for equipment installation, etc. This finding also shows that high-tech manufacturers can operate a large-sized capacity plant in regions with adequate and low-cost supplies and where low-paid skilled workers are available, or if the governments provide incentives, such as rent or tax holiday, to the high-tech industry. Table 7 also shows that not all FABs are being operated until the amount of customer demand exceeds  $17.99 \times 10^7$  dies. For example, T-company needs only three plants when the customer demand is  $12.85 \times 10^7$  dies. This result implies that when the demand is extremely low, it is not necessary to operate all of the plants, and in addition a location with a very high cost is not suggested. As the customer demand increases, in addition to two 12-inch FABs, T-company will start to operate a smaller-sized FAB at locations with high capital and variable production costs. For example, when customer demand is  $11.82 \times 10^7$  dies, then there are three 12-inch FABs, i.e. Hsinchu, Tainan and Shanghai with 100% capacity utilization and one 6-inch FABs in the USA, with a utilization as low as 25%. This implies that a company may incur high costs to operate a large-sized capacity at a

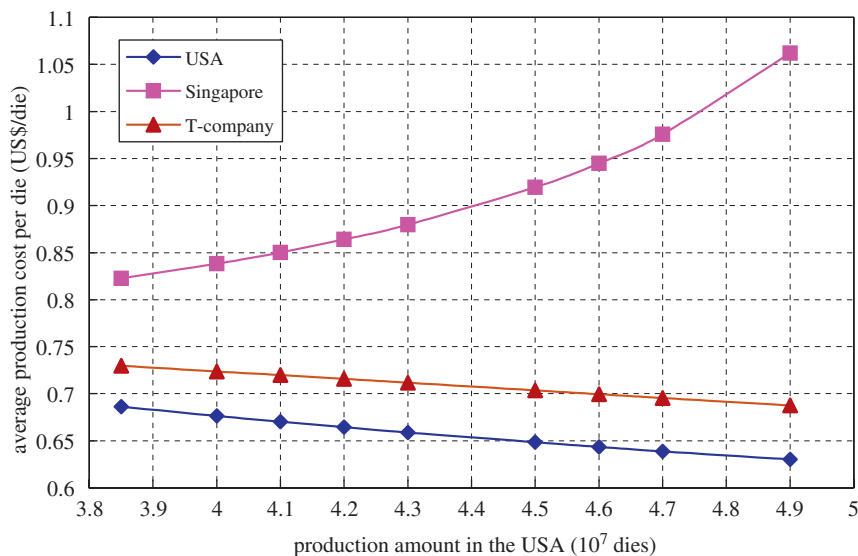


Fig. 4. Average production cost per die vs. production amount in the USA.

**Table 7**

The optimal size and capacity utilization of FABs with different amounts of customer demand.

Customer demand (10 <sup>7</sup> dies)	The total average production cost per die (US\$/die)	Taiwan (Hsinchu)		Taiwan (Tainan)		Shanghai		USA		Singapore	
		(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
		11.57	0.6823	12	100	12	100	12	94	–	–
11.82	0.6869	12	100	12	100	12	100	6	25	–	–
12.85	0.6813	12	100	12	100	–	–	12	60	–	–
15.42	0.6549	12	100	12	100	12	100	8	34	–	–
17.99	0.6530	12	100	12	100	12	100	12	65	–	–
20.56	0.6386	12	100	12	100	12	100	12	100	8	46
20.82	0.6387	12	100	12	100	8	60	12	100	12	100
21.07	0.6387	12	100	12	100	8	74	12	100	12	100
21.33	0.6386	12	100	12	100	8	89	12	100	12	100
21.46	0.6387	12	100	12	100	8	96	12	100	12	100
21.59	0.6522	12	100	12	100	12	100	12	100	12	38
21.85	0.6036	12	100	12	100	12	100	12	100	12	43
22.10	0.5560	12	100	12	100	12	100	12	100	12	48
23.13	0.5066	12	100	12	100	12	100	12	100	12	69
24.41	0.5024	12	100	12	100	12	100	12	100	12	95

(1): size of FAB.(2): percentage capacity utilization.

region with relative high capital and production costs when there is not large demand. As the customer demand increases, the tendency to operate 12-inch FABs increases; even though there is not enough demand to enable full-capacity production for all 12-inch FABs.

This study further compares the cost efficiency of operating one large-sized capacity plant from that of operating few small-sized capacity plants. As shown in Table 7, the optimal FAB size in Singapore is 12-inch once customer demand exceeds  $21.59 \times 10^7$  dies. Assuming the manufacturer alters the strategy of operating one 12-inch FAB in Singapore to two 8-inch FABs instead, while the FAB sizes and the capacity utilizations of the plants in the other regions remain the same as those in Table 7. Fig. 5 shows the comparison of the results between operating one 12-inch FAB and two 8-inch FABs in Singapore with customer demand ranging from  $21.59 \times 10^7$  to  $23.13 \times 10^7$  dies.

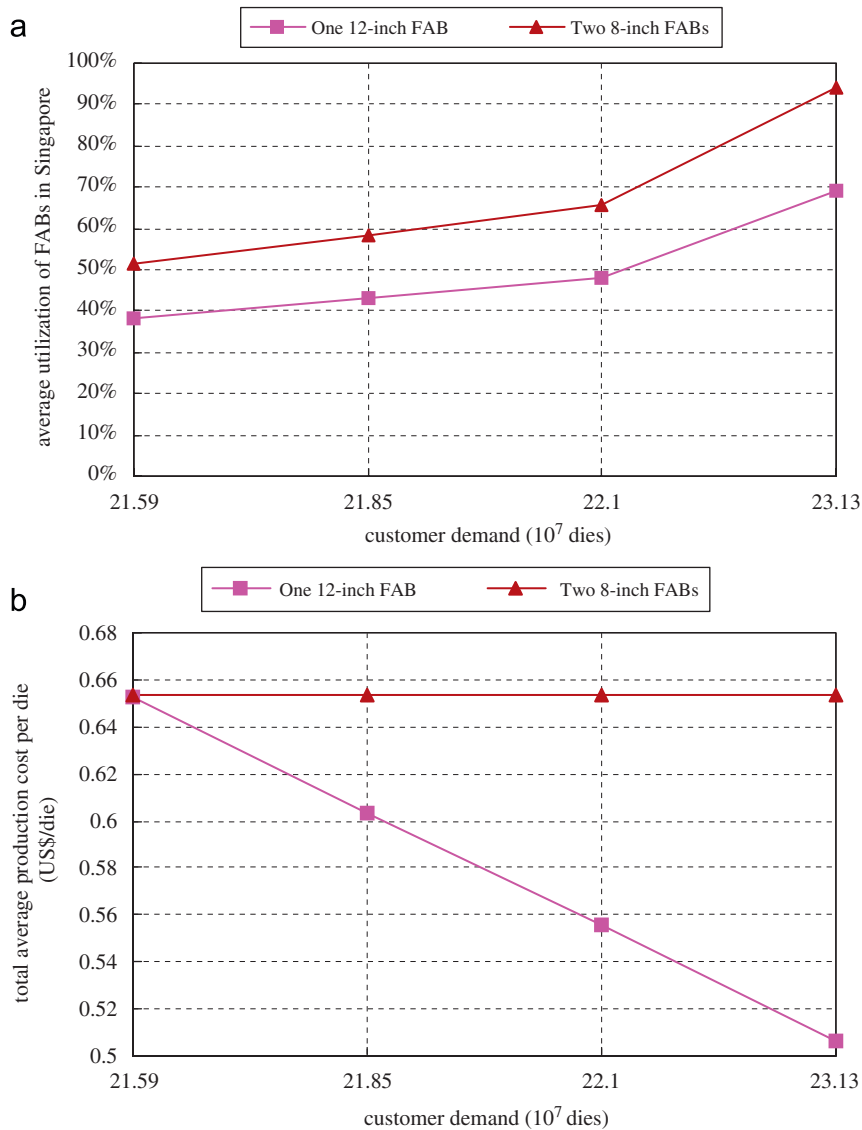
As shown in Fig. 5(a), the capacity utilizations of FABs when operating either one 12-inch FAB or two 8-inch FABs increase when there is an increase in customer demand. The capacity utilizations of the two strategies increase from 38% to 69% and from 51% to 94%, respectively. However, the 12-inch FAB is not of high utilization as the two 8-inch FABs, i.e. the utilization of the 12-inch FAB is smaller than that of two 8-inch FABs. Though operating the smaller-sized FABs results in a high-capacity utilization value, the strategy of operating two 8-inch FABs yields higher production cost as shown in Fig. 5(b). Moreover, the advantage of operating the 12-inch FAB over two 8-inch FABs increases as customer demand increase.

## 5. Conclusions

Past studies have extensively investigated plant location issues. Most of these studies dealt with the problem

by constructing MIP models in which the plant is constrained to serve a single warehouse or customer. These studies focused mainly on developing an approximation procedure and on the efficiency of the proposed heuristics. However, the impact of high capital cost on the optimal plant capacity, and the production amount among the plants of different-sized were rarely discussed. This study constructed a nonlinear MIP model to minimize the average total cost per unit product subject to constraints such as satisfying customer demand in various geographic regions, relationships between supply flows and demand flows within the physical configuration, and the production limitation of different size plants. This study showed how economies of scale can be considered in solving the capacity and production problems. This study also showed that the capacity utilization as well as the production amount in the short run, and the size of capacity of multiple plants in the long run are related, and that those two factors influence the total cost. Note that the study focused on plant capacity and production planning issues in wafer fabrication industry, the findings may only be applied to wafer fabrication manufacturers.

A case study of T-company in Taiwan, which is the world's largest dedicated semiconductor foundry, was provided to demonstrate the application of the proposed models. The results show that because of the high customer demand, the manufacturer can operate the plants with large-sized capacity combined with full-capacity production, thereby lowering the production cost. Since the government of Taiwan provided incentives for developing the high-tech industry, and since there is a large local customer demand due to the economies of agglomeration in the semiconductor industry, T-company' core operations are based in FABs in Taiwan. The results of this study also show that when determining the production amount for multiple plants, those with large-sized capacity combined with low capital and variable production costs have a higher priority in filling this capacity,



**Fig. 5.** Comparisons of the results between operating one 12-inch FAB and two 8-inch FABs in Singapore. (a) Average utilization of FABs in Singapore vs. customer demand. (b) Total average production cost per die vs. customer demand.

compared to those with small-sized capacity combined with relative high capital and variable production costs. The reason is not only they have higher capability to satisfy customer demand, but also they are more cost effective. Although there is a trade-off between production cost economies and transportation cost diseconomies, the results show that the benefits in terms of cost savings for the wafer foundry company brought by centralized production are larger than the increased transportation cost as a result of decentralization. Therefore, this finding suggests that the manufacturer may adopt a production strategy of centralizing production in manufacturing plants with large-sized capacity and then shipping the products to customer in different regions.

As to raw material procurement, the results show that vendors providing low purchase price are mostly active

even there exists a high fixed cost, i.e. searching the active vendor is difficult. Since there is low unit purchase cost, a large-amount procurement from the vendor brings great advantage for the manufacturer thereby outweighing the disadvantage of high fixed cost. Also, active vendors tend to serve manufacturing plants that are nearby. In addition, to reduce outbound costs, the product should be shipped from a manufacturing plant to a customer that is located within a short distance. The results also show that the production cost is the highest of all costs for a wafer foundry. This finding implies that the wafer foundry industry shows production with economies of scale, and that this production is the most valued-added in the entire supply chain. Therefore, the manufacturer must be aware of the impact on the total cost of capacity utilization by its manufacturing plants with

different-sized capacity. The results also show that a location with very high capital and variable production costs is not recommended when the demand is extremely low. Moreover, as long as the customer demand is large enough to offset the high capital cost, operating a large-sized capacity plant is recommended even if the capacity utilization is low. In other words, operating a few small-sized capacity plants with high-capacity utilization may not yield a decreased production cost as compared with that from operating merely one larger-sized capacity plant with lower capacity utilization.

In sum, the integrated plant capacity and production model developed in this study provides a highly effective tool that enables high-tech manufacturers to evaluate the expansion, contraction, or reallocation of their production tasks among various plants. This model can be applied to investigate the relationship between the total average production cost and production allocation when deciding whether to operate a new manufacturing plant, or to develop a new technology, such as 0.09  $\mu\text{m}$  process technology, which may lead to larger production. Due to yield issues (dies per wafer) of bigger size wafers, and advanced technology involved in operating larger sized wafer FABs, the key findings of the paper may only be related to the wafer fabrication industry. However, the implication of the finding is worth noted for other manufacturing settings. When the transportation cost of the product is relatively small as compared to its production cost, the finding suggests a production strategy of centralizing production in large-sized capacity plants and then shipping the products to customer in different regions. And the strategy is appropriate when the production facilities have huge setup costs and rather low variable production costs.

## Acknowledgements

The authors would like to thank the National Science Council of the Republic of China for financially supporting this research under Contract no. NSC 93-2416-H-009-005.

## References

- Brimberg, J., ReVelle, C., 1998. A bi-objective plant location problem: cost vs. demand served. *Location Science* 6, 121–135.
- Chen, J.C., Chen, C.W., Lin, C.J., Rau, H., 2005. Capacity planning with capability for multiple semiconductor manufacturing fabs. *Computers & Industrial Engineering* 48, 709–732.
- Chen, F., Federgruen, A., Zheng, Y.S., 2001. Coordination mechanisms for a distribution system with one supplier and multiple retailers. *Management Science* 47, 693–708.
- Chopra, S., 2003. Designing the distribution network in a supply chain. *Transportation Research E* 39, 123–140.
- Chou, Y.C., Cheng, C.T., Yang, F.C., Liang, Y.Y., 2007. Evaluating alternative capacity strategies in semiconductor manufacturing under uncertain demand and price scenarios. *International Journal of Production Economics* 105, 591–606.
- Cohen, M.A., Moon, S., 1990. Impact of production scale economies, manufacturing complexity, and transportation costs on supply chain facility networks. *Journal of Manufacturing and Operations Management* 3, 269–292.
- Cohen, M.A., Moon, S., 1991. An integrated plant loading model with economies of scale and scope. *European Journal of Operational Research* 50, 266–279.
- Dowland, K.A., Soubeiga, E., Burke, E., 2007. A simulated annealing based hyperheuristic for determining shipper sizes for storage and transportation. *European Journal of Operational Research* 179, 759–774.
- Fu, Y., Piplani, R., 2004. Supply-side collaboration and its value in supply chains. *European Journal of Operational Research* 152, 281–288.
- Goetschalckx, M., Vidal, C.J., Dogan, K., 2002. Modeling and design of global logistics systems: a review of integrated strategic and tactical models and design algorithms. *European Journal of Operational Research* 143, 1–18.
- Guinet, A., 2001. Multi-site planning: a transshipment problem. *International Journal of Production Economics* 74, 21–32.
- Heragu, S.S., Alfa, A.S., 1992. Experimental analysis of simulated annealing based algorithms for the layout problem. *European Journal of Operational Research* 57, 190–202.
- Jayaraman, V., Pirkul, H., 2001. Planning and coordination of production and distribution facilities for multiple commodities. *European Journal of Operational Research* 133, 394–408.
- Julka, N., Baines, T., Tjahjono, B., Lendermann, P., Vitanov, V., 2007. A review of multi-factor capacity expansion models for manufacturing plants: searching for a holistic decision aid. *International Journal of Production Economics* 106, 607–621.
- Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P., 1983. Optimization by simulated annealing. *Science* 220, 671–680.
- Levis, A.A., Papageorgiou, L.G., 2004. A hierarchical solution approach for multi-site capacity planning under uncertainty in the pharmaceutical industry. *Computers and Chemical Engineering* 28, 707–725.
- Loukil, T., Teghem, J., Fortemps, P., 2007. A multi-objective production scheduling case study solved by simulated annealing. *European Journal of Operational Research* 179, 709–722.
- Metropolis, N., Rosenbluth, M., Rosenbluth, A., Teller, A., Teller, E., 1953. Equation of state calculations by fast computing machines. *Journal of Chemical Physics* 21, 1087–1092.
- Miranda, P.A., Garrido, R.A., 2004. Incorporating inventory control decisions into a strategic distribution network design model with stochastic demand. *Transportation Research E* 40, 183–207.
- Mishra, A., Dutta, P.K., Ghosh, M.K., 2003. A GA based approach for boundary detection of left ventricle with echocardiographic image sequence. *Image and Vision Computing* 21, 967–976.
- Nagurney, A., Dong, J., Zhang, D., 2002. A supply chain network equilibrium model. *Transportation Research E* 38, 281–303.
- Nagurney, A., Toyasaki, F., 2005. Reverse supply chain management and electronic waste recycling: a multitiered network equilibrium framework for e-cycling. *Transportation Research E* 41, 1–28.
- Nazzari, D., Mollaghasemi, M., Anderson, D., 2006. A simulation-based evaluation of cycle time reduction in Agere Systems wafer fabrication. *International Journal of Production Economics* 100, 300–313.
- Paik, C.H., Soni, S., 2007. A simulated annealing based solution approach for the two-layered location registration and paging areas partitioning problem in cellular mobile networks. *European Journal of Operational Research* 178, 579–594.
- Santoso, T., Ahmed, S., Goetschalckx, M., Shapiro, A., 2005. A stochastic programming approach for supply chain network design under uncertainty. *European Journal of Operational Research* 167, 96–115.
- Taiwan Institute of Economic Research, 2004. *International Economic Data* <<http://english.tier.org.tw/english.htm>>.
- Taiwan Semiconductor Manufacturing Company, 2004. 2004 Annual Report <[http://www.tsmc.com/english/e\\_investor/e02\\_annual/e02\\_annual.htm](http://www.tsmc.com/english/e_investor/e02_annual/e02_annual.htm)>.
- Tsiakis, P., Papageorgiou, L.G., 2008. Optimal production allocation and distribution supply chain networks. *International Journal of Production Economics* 111, 468–483.
- Tzafestas, S., Kapsiotis, G., 1994. Coordinated control of manufacturing/supply chains using multi-level techniques. *Computer Integrated Manufacturing Systems* 7, 206–212.
- Vidal, C.J., Goetschalckx, M., 1997. Strategic production–distribution models: a critical review with emphasis on global supply chain models. *European Journal of Operational Research* 98, 1–18.
- Viswanathan, S., Piplani, R., 2001. Coordinating supply chain inventories through common replenishment epochs. *European Journal of Operational Research* 129, 277–286.
- Yan, S., Luo, S.C., 1999. Probabilistic local search algorithm for concave cost transportation network problem. *European Journal of Operational Research* 117, 511–521.