Optimization Problem
and Capacity Allocation
Process in a
Semiconductor Company

M. Bucolo*, F. Di Grazia*, G. Strano*,
and G. Tomarchio*†

1 Introduction

In structures with a complex organization, as enterprises, public administration agencies, decisions are continuously taken, and they can have modest or noteworthy relevance and they can cause short or long terms effects [1]. The individual or collective capability of taking such decisions represents one of the main factors affecting the performance and the competitiveness of an organization [2]. On the other hand, in the current socioeconomic environment, the high speed of the economic processes and the deep transformations occurring in the production modes and in the economical relationships have given more and more importance to the information transfer[3]. Fortunately the adoption of technologies for low cost storage memories with high connectivity made a huge amount of data to be available and stored, but the open issue is the development of strategies for the planning and the control in enterprises with a complex organization, which allow the transformation of such information into knowledge[4]. The planning system is usually connected to the management control system, which has the aim to drive the enterprise management towards the achievement of the planned objectives and to implement the opportune corrective actions. This strict integration motivates why, usually, the term "systemofplanningandcontrol" is used [5]. In semiconductors market where competition is played on the ability to innovate and where the cost of production increases with the number of products to manage and decreases with the

* DIEES Università degli Studi di Catania email: maide.bucolo@diees.unict.it
† STM Microelectronics Catania Site email: giusy.tomarchio@st.com
time to be produced, the process of production planning is difficult to be managed. It is done usually at the level of product or its aggregation, and allows to define the productive capacity needed to support the market demand [6]: a proper planning thus maximize the production capacity adapting it to demand trend (Just In Time) and optimizing the use of production plants. A Decision Supporting System (DSS) [7][8] transforms the data coming from the analysis tools by a business intelligence framework to forms of active support, by using methods and mathematical models for the resolution of complex problems which arise in the management of enterprises and structured organizations and by giving opportune control strategies. The aim of this paper is to create a DSS for the optimization of the supply-chain to find the best production capacity management for the Analog, Power and MEMS group (APM) of STMicroelectronics, that is described in section 2. STMicroelectronics is one of the worldwide biggest semiconductor firm and is include by five groups that have different business models along the five dimensions that characterize the sales in term of sale regions (America, Europe, Asia Pac, Emerging Markets, Japan), market segments (product application environment: industrial, automotive, power supply and so on), sale channels (class of customers: key customers, retailers and so on), product families (aggregated level or Business Unit and product level). The need of the optimization of the logistic chain and thus of having models and management tools for the medium-term planning and for the capability analysis is particularly critical in relation to logistic-productive systems of high complexity, of polycentric kind, i.e. systems which are distributed on several establishments and several markets and which are characterized by an high level of investments in technologies of high automation and by an intensive use of the productive capability [9]. In this perspective, the process of medium-term production planning is devoted to the determination of a planning of logistic and productive flows which is optimal, i.e. such that the total cost (the sum of the costs of supply, transformation, stock and distribution) is minimized [8]. On the other hand, as reported in Section 3, in order to make the optimal capacity allocation suitable for the practical adoption, it is necessary that it is also feasible, i.e., it satisfies the logic and physical constraints which derive from the limited productive capability, from specific technological conditions, from the configuration of the logistic network and from conditions of subjective nature as expressed by the responsible of the planning process. Both single objective model, solved by the Simplex Method(SM) [10], and multiobjective model [11], solved by the Genetic Algorithm (GA), have been developed and deployed in the system by an automatic process integration. In this framework, a fundamental role is played by both the choice of the functional to be optimized and the model definition. All the results, shown in Section 4, are evaluated and compared through suitable adopted performance indexes and visualized through complex network architectures, that can represent start points for future study focuses on the development of an innovative multiobjective approach based on the graph theory.
2 Case Study

The supply chain can be defined, in a quite large sense, as a network of organized units connected and interdependent which operate in a coordinated way to manage, control and optimize the flow of material and information which go from the suppliers to the final clients after passing from the supply, transformation and distribution subsystems of an enterprise. Semiconductor firms, as well as any other indirect supplier, are situated in a supply chain that starts from the final customer and ends with the raw material supplier. The company sales are obviously the result of what the final user requires propagated along the whole supply chain and between the final customer and the semiconductor firm, several entities act with different functions as other manufacturing firms or distributors that represent the direct customers of the semiconductor firm. In particular the study here presented is focused on the APM group, that consists of 20 Business Units (BU) sharing 58 plants located all over the world for the production of 168 different macro-packages, that are the sets of products assembled using the same technology. The management of the productivity capacity process in the APM group has a fundamental role in the entire management of production process, it comes to be implemented by all planning cycles (Long, medium and short term) being a key element. The one of the main issue within the APM group is the management of the allocation process, whose aim is the allocation of the capacity into the BUs belonging to the group. The organizational structure of APM provides that each BU is managed as a freestanding entity, with specific degree of autonomy (i.e chooses its targets, its strategies) and starting from common guidelines regarding the ordinary management, takes its decisions, in particular, in relation with the economic-financial aspects, presenting their own budgets. The Central Planning is responsible for the management of capacity allocation, considering both the capacity of the production sites and the various BU requests. The production process is high critical because different products, in terms of BU, technology and application, can share the same package, therefore they are manufactured in the same production-lines by the same machineries adequately equipped. The route of production management is often broken down into two stages: in the front end where the product is processed through the silicon diffusion and the back end which handles the packaging and testing. The analysis here presented get from the back end macrophase in relation to the year 2006, considering all the four quarters, allowing, thus, the analysis throughout the entire year. The case study concerns the capacity allocation management of all the products related to five macro-packages that have been selected from the 20 macro-packages available, by applying the Pareto analysis. They cover the 40% of total production of APM group. As reported in Table 1, each selected macro-package is used by a different number of BU, indicated in the third column, and can be manufactured a different number of plants, indicated in the fourth column.
Table 1. Subset of macro-package chosen through Pareto approach.

<table>
<thead>
<tr>
<th>Macropackage</th>
<th>CODE</th>
<th>DESCRIPTION</th>
<th>nBU</th>
<th>nPlants</th>
</tr>
</thead>
<tbody>
<tr>
<td>TO 220 PWR</td>
<td>14</td>
<td>'TO 220 PWR'</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>'SO 8'</td>
<td>50</td>
<td>9</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>'SO 14'</td>
<td>6C</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>'TO 220 D2PACK'</td>
<td>BQ</td>
<td>7</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>'DPAK'</td>
<td>CP</td>
<td>7</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

3 Optimization Strategies

Generally, the optimization process involves the use of specific techniques to determine the most cost effective and efficient solution to a problem or design for a process, usually trying to maximize or minimize one or more of the process specifications, while keeping all others within a fixed range [8][10]. In the first phase it is fundamental to construct a mathematical model, feasible with the process that have to be optimize and with constrains that have to be respected. Given the set of the feasible solutions \( X \), and a real objective function \( f(x) \) definite for each element of \( X \), it is required to find the element of \( X \) which maximizes (or minimizes) an objective function, as \( \max_{x \in X} f(x) \), that is subjected to \( m \) constraints \( g_i(x) \leq 0 \), \( i = 1, 2, \ldots, m \). Specific hypothesis on \( X \), \( g_i(x) \), \( f(x) \) bring to different classes of optimization problems. Particularly in Linear Programming Problem \( X \) is a particular subset of the \( n \)-dimensional space \( \mathbb{R}^n \) and, \( g_i(x) \) and \( f(x) \) are both modelled by linear expressions [10]. In general the problem could not have feasible solutions, in this case it is called inadmissible; but if it admits solution, it is possible to find the optimal solution using various numerical methods, one of the most adopted is the Simplex Method. The proposal paradigm assumes that the decisional process has absolutely rational logic structure, who decides expresses his preferences into an objective function and chooses the best solution in terms of profit. More frequently decisions come as results of compromises and balances between different performance criterions, in this case solutions are provided by Multiobjective Optimization Problem [11][12]. Supposed to have \( k \) objectives \( f_i(x) \)(\( i = 1, 2, \ldots, k \)) the function that have to be optimize can be written as:

\[
F(x) = (f_1(x), f_2(x), \ldots, f_k(x)) \text{ with } x \in \Omega \text{ and } h(x) \leq 0 \text{ } c(x) = 0
\]

where \( x \) is a solution (decision vector) defined in \( \Omega \), and the terms \( c(x) \) and \( h(x) \) are, respectively, equality and inequality constrain of the problem. The multiobjective optimization problem is based on the research of the "non dominated" solution that is when there is not any other solution which guarantees all the objectives functions at the same time. A non dominated solution is often called as pareto solution and the set of all the pareto solutions is called pareto set [13]. A Multiobjective Optimization Problem need to be scalarized that means that the \( k \) objective functions are combined in a single one, different ways are shown in literature to do this [8][11]. Both these approaches have been considered to model the resource allocation problem related to the case study: a single objective linear model solved by
the simplex optimization method and a multiobjective model implemented through
the genetic algorithm. In the case study considered, the main constraints are the
production limit of the sites, each plant cannot produce more than its maximum
production capacity. Supposing to have m BUs sharing n Plants, each BU has a
request of capacity that have to be sanctified, and Plant has a availability that can
not be overcame. The central planning have to determine how much capacity, for
each Plant, is possible to be shared by the BUs in order to maximize the profit
established as price-cost that involves both the production phase and the logistic
for each product. In the single objective model, this optimization problem can
be referred to a transportation problem [10] and considering the generic "i" BU that
produces in the "j" plant (P_j), can rewritten as follows:

\[ \text{max } f(x) = \text{max} (\sum_{ij} c_{ij} x_{ij}) \]

where \( f(x) \) is the object function, \( c_{ij} \) is the profit that the BU_j can obtained
by producing in the \( P_j \) and the decisional variable \( x_{ij} \) is in term of the number of
products. All the \( c_{ij} \) elements are collected in a \( mn \) matrix \( C \), the Payoffs Matrix.
Moreover this problem is subjected to the following constraints:

\[ \sum_{j} x_{ij} - C_{b_i} = 0 \forall i = 1, ..., m \]
\[ \sum_{j} x_{ij} - C_{p_j} = 0 \sum_i x_{ij} = C_{p_j} \forall j = 1, ..., n \]
\[ x_{ij} \geq 0 \forall i = 1, j \]

where \( C_{b_i} \) is the total request of the "i" BU and \( C_{p_j} \) as the available capacity
of the "j" Plant. The model assumes that BU_i request distribution over all the
plants is \( C_{b_i} \), and generic plant \( P_i \) could not produce, in respect of all the BU_i
request a number of products bigger than its maximum capacity \( C_{p_j} \). The values
of \( C_{b_i} \) e \( C_{p_j} \) are expressed in quartely averaged kpcs per day. The problem admits
a solution when the sum of the BU_i request (\( C_{b_i} \)) is equal to the sum of the Plant
capacity (\( C_{p_j} \)):

\[ \sum C_{p_j} = \sum C_{b_i} \]

meanwhile if the sum of the BU_i request (\( C_{b_i} \)) is bigger that the sum of the
Plant capacity (\( C_{p_j} \)), a fictitious plant (\( P^* \)) is defined to realize the equality of the
two terms, as follows:

\[ \sum C_{p_j} < \sum C_{b_i} \quad C_{p^*} = \sum C_{b_i} - \sum C_{p_j}. \]

The Multi objective model has been developed considering \( m \) objectives
related to request of each BU, \( n \) objectives related to maximal level of production
for each Plant and one objective related to global profit, as follows:

\[ F(x) = (f_1(x), ..., f_m(x), f_{m+1}(x), ..., f_{m+n}(x), f_{m+n+1}(x)) \]

with \( x \in \Omega \) is the vector of the decision variables and the total number of
objectives is \( m + n + 1 \). The "i" objective function for \( i = 1...m \), is estimated as
the difference between production and request of each BU,
\[ f_i(x) = \left| \sum_{j}^{n} x_{ij} - C_{bi} \right| \forall i = 1, ..., m. \]

The "i" objective function for \( i = m + 1 \ldots n \), is estimated as the difference between production and capacity of each Plant,

\[ f_i(x) = \left| \sum_{j}^{m} x_{ij} - C_{pj} \right| \forall j = 1, ..., n. \]

The last objective function \( m + n + 1 \) consider the total production profit:

\[ f_{m+n+1}(x) = \frac{1}{\sum c_{ij} - x_{ij}}. \]

To maintain the continuity with the past allocation decisions, the variables \( x \) have been chosen in the range \( 0 \leq x \leq 2a^* \) where \( a^* \) is the maximum value among the elements \( a_{ij} \) of the actual allocation matrix \( A \) of the APM group, related to the year 2006. To scalar the problem all the \( m + n + 1 \) objective functions has been considered as components of an unique vector, and this fitness vector has been minimized by the GA using as optimal selection the following parameters for all the case studies.

<table>
<thead>
<tr>
<th>PARAMETERS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>NIND</td>
<td>individuals for population</td>
</tr>
<tr>
<td>MAXGEN</td>
<td>Maximum number of generation</td>
</tr>
<tr>
<td>GGAP</td>
<td>Generation gap</td>
</tr>
<tr>
<td>PRECI</td>
<td>Binary precision</td>
</tr>
</tbody>
</table>

4 Results

To compare the performance of the proposed models, two suitable performance indexes \((I, I_p)\) have been defined, for the estimation of the profit increase in relation to the produced quantities the percentage of introduced increase. The couples \((F_{old}, P_{old})\) and \((F_{new}, P_{new})\) are related to the object function values and produced unit quantities evaluated respectively through the method actually adopted at the APM central planning and with the two introduced models, so it is possible to introduce, the following parameters:

\[ I = \frac{I_{new}}{I_{old}} \quad I_p = \frac{I_{new} - I_{old}}{I_{new}} \cdot 100 \]

where \( I_{new} \) and \( I_{old} \) are defined as \( I_{new} = F_{new}/P_{new} \) and \( I_{old} = F_{old}/P_{old} \). The results express through the indexes \((I, I_p)\) related to the single objective model, the for each quarter of the year and for all the five macro-packages selected, are reported respectively in the Tables 2-3. In Tables 4-5 the values of indexes \((I, I_p)\) are related to the multi-objective models, moreover the genetic algorithm approach offers the possibility to the expert to select a favourite solution within admissible set. The results related to the first model show that is possible to reach significant improvements from the 3% (MP ”14”, quarter IV ) to 47% (MP ’6C’). quarter
Table 2. $I$ values of each Macro-package (MP) in the four different quarters, for one objective method.

<table>
<thead>
<tr>
<th>MP</th>
<th>I QUARTER</th>
<th>II QUARTER</th>
<th>III QUARTER</th>
<th>IV QUARTER</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>0.9</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>50</td>
<td>1.2</td>
<td>1.2</td>
<td>1.1</td>
<td>1.2</td>
</tr>
<tr>
<td>6C</td>
<td>1</td>
<td>1.2</td>
<td>1.1</td>
<td>1.9</td>
</tr>
<tr>
<td>BQ</td>
<td>0.9</td>
<td>1</td>
<td>0.9</td>
<td>1</td>
</tr>
<tr>
<td>CP</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3. $I_p$ values of each Macro-package (MP) in the four different quarters, for one objective method.

<table>
<thead>
<tr>
<th>MP</th>
<th>I QUARTER</th>
<th>II QUARTER</th>
<th>III QUARTER</th>
<th>IV QUARTER</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>-8.2</td>
<td>-2.8</td>
<td>-1.6</td>
<td>3.2</td>
</tr>
<tr>
<td>50</td>
<td>14.5</td>
<td>10.9</td>
<td>12.6</td>
<td>15</td>
</tr>
<tr>
<td>6C</td>
<td>2.7</td>
<td>16.1</td>
<td>8.8</td>
<td>4.7</td>
</tr>
<tr>
<td>BQ</td>
<td>-7.4</td>
<td>-2.7</td>
<td>-6.9</td>
<td>0.1</td>
</tr>
<tr>
<td>CP</td>
<td>0.9</td>
<td>1.5</td>
<td>-4.5</td>
<td>0.3</td>
</tr>
</tbody>
</table>

IV), and in almost all the case the performances improvement is around the 10%. The results related to the second model show that is possible to reach significant improvements from the 6% (MP “50”, quarter I) to 17% (MP ’50’, quarter IV), and in almost all the case the performances improvement is around the 7%. Looking at global performance degree, the sum of the improvements of all quarters for all macropackages, reaches an increasing profit of 25% and 6% by using respectively the single-objective model and the multi-objective model. As shown in the tables by first model it is possible to reach an average improvement higher than second one, but it have to be taken into account that the results are obtained by fixing the same set of GA parameters for all the case studies and it represents a limitation for the performances of the multiobjective model. However the multiobjective model provides a mathematical description of the allocation problem more close to the expert view, due to the competitive nature of problem and an important result is complementariness between two models: where first single objective model seems to be less efficient, the potentialities of multiobjective approach increases.

Finally, to obtain an easy visualization of the allocation problem, the result have been read through the graph theory approach. Network theory is the natural framework for the mathematical treatment of complex networks as in this case where two different agents are represented as bipartite graphis. A bipartite network is a triple $G = (T,\perp,E)$ where $T$ and $\perp$ are two disjoint sets of nodes, respectively the top and bottom nodes, and $E \subseteq T \times \perp$ is the set of links of the network [14]. The difference with classical (unipartite) networks lies in the fact that links exist only between top nodes and bottom nodes. In this study the graph theory has been used exploiting the concept of weighted graph, in which the nodes are the BUs and Plants of each macro-package and the links related to top-bottom (BUi-Pj)
associations are weighted in term of percentage of capacity used for production of macropackage. As example the capacity allocation of the macropackage ’50’, IV quarter, is reported in Figures 1a-b-c by three network representations each related to a different case: not optimized allocation network (Figure 1a), single objective optimized allocation network (Figure 1b) and multi objective optimized allocation network (Figure 1c). In particular, the improvement arisen for the optimized allocation networks, single objective and multi objective optimized allocation networks, is equal to the 17% and 15% respectively. Furthermore, looking at the Figures 1b-c, optimized networks does not have new links in respect of those appearing in the not optimized network, Figure 1a, thus only a reallocation of the previous capacity have been performed. Anyway some links have been removed to permit a better performance on profit improvement. This proposed representation not only give a fast way to visualize the results but could represent a new way to cope with the allocation problem considering as objective functions the network parameters (degree distribution, clustering coefficient and characteristic path length)[15] .

5 Conclusions

A Decision Supporting System have been provided with the aim to cope with the issue of capacity allocation laying on the supply-chain organization of APM Group of STMicroelectronics. Two suitable mathematical models have been used to analyze and solve those planning problems, one based on a functional maximization of profit and the other standing on multi-objective optimization techniques implemented exploiting the computational ability of genetic algorithms. Both methods provided an improvement in the profit quantitatively characterized through suit-
Figure 1. Capacity allocation network for the macropackage ’50’, in the original state (a) with the application of one-objective linear method (b) and with the application of multi-objective method (c).

able adopted performance indexes. The first model reach an average improvement higher than second one, however the multiobjective model provides a mathematical description of the allocation problem more close to the expert view, due to the competitive nature of problem. An important result is the complementariness between two models: where first single objective model seems to be less efficient, the potentialities of multiobjective approach increases. The results have been also represented exploiting the base concept of the network theory and in particular the bipartite graphs structures satisfying the system requirement of having two different agents: BUs and Plants. This strategy not only give a fast way to visualize the results but could represent the starting point for further studies to manage the allocation problems in an enterprise production’s networks considering as ob-
jective functions the classical network parameters (degree distribution, clustering coefficient and characteristic path length).
Bibliography


