

025-0827

Grid computing for a stochastic product-mix problem in Brazil

BARBASTEFANO, R. G., CEFET/RJ, Av. Maracanã, 229, sala E-504b – Rio de Janeiro

– Brasil – 20271-110. rgb@cefet-rj.br

SOUZA, L. R., CEFET/RJ, luizrrossi@gmail.com

CARVALHO, D., CEFET/RJ, dcarvalho@cefet-rj.br

POMS 23<sup>rd</sup> Annual Conference

Chicago, Illinois, U.S.A.

April 20 to April 23, 2011

# 1 Introduction

A Biaxially Oriented Polypropylene (BOPP) manufacturer has to be capable of producing hundreds of products varying in weight, coloring and cutting in order to meet market requirements and stay competitive. For such a company, a greater portfolio implies in increased production complexity with increasing number of setups [1] and available combinations between products and production lines. As product-mix decisions affect both manufacturing efficiency and productivity, adopting optimization techniques for production planning can help managers increase the quality of decision-making on capacity distribution and product generation.

BOPP production yields stochastic scrap during extrusion and cutting, increasing the difficulty to find an optimal solution for product allocation. This work<sup>1</sup> presents a methodology designed to optimize the stochastic product-mix planning problem of a BOPP company with a portfolio of more than 200 products, manufactured in 6 different production lines within 3 plants located in South America. Considering its production lines and the 3 possible markets, the mixed integer programming problem (MIP) applied to the product-line allocation analysis has at least 2,500 variables and 3,600 restrictions.

Stochastic programming techniques [2] were adopted to account for waste generation during manufacturing. A series of Monte Carlo (MC) experiments [3] were carried out to estimate an optimal solution used to maximize total contribution margin [4]. Scrap configurations were drawn within each experiment, generating MIPs that were solved to determine the maximum contribution related to each product-line allocation. The best product-line allocations were chosen among random configurations in order to find a feasible solution with the greatest likelihood.

A grid computing infrastructure with over 1,200 processors, belonging to the EELA-2 (E-science grid facility for Europe and Latin America) collaborative project [5], was used to run the MC simulation. This infrastructure allowed the execution of 50,000 experiments in 6 processing hours per execution. The equivalent computing effort in a

---

<sup>1</sup>Supported by FAPERJ.

single desktop computer would consume 16 days. The analysis of the diverse executed experiments indicated the possibility of an increase around 10% in the company's total contribution margin using a product-mix determined with the developed technique.

This work is organized as follows: in section 2 the product-mix concept is presented and its importance to industries explained. In section 3 lies a discussion about the adopted methodologies for the research and development of a solution specific for the BOPP company in study. Section 4 presents some data concerning the company and its products. Sections 5 and 6 refer to the development of the mathematical model used to simulate the company's operations with MC experiments and this model's computational implementation over EELA-2 infrastructure, respectively. In section 7 the results are presented with a brief discussion about their impacts. Section 8 concludes the study with authors' observations.

## **2 Product-mix Problem**

Product-mix involves a decision making process in which one's objective is to maximize a profit function (in general, represented by total contribution margin), assigning capacity allocation to a group of products, considering companies' production limitations and the items' demand forecast [1]. A manager seeks the most profitable way of producing and selling goods through the rational usage of productive resources while determining a companies' product-mix. The definition comprises everything from a portfolio determination to be offered to customers to the capacity distribution among products. Marketing departments usually believe that an increase in product assortment yields profit addition to an enterprise. However, a company faces diseconomies of scale in its specialized production of a few goods as its product variety increases, which brings a trade-off between product variety and economies of scale.

A product line development consists of two stages: (a) identify the portfolio to be adopted, based on client necessities and expectations and (b) portfolio selection and as-

assessment, which consists in finding the most profitable way to allocate scarce resources to manufacture products [6]. The present work explores the configuration for the resource allocation in a plastic manufacturing company, presented in section 4. The interested reader may find a discussion on portfolio identification in [7].

Product-mix planning can be thought of as a set of decisions aiming to determine which product-line configuration will yield higher profit, given that the company has more than one production line capable of making a set of products. In other words: in which production line each product shall be manufactured so the company achieves the highest possible return. Very often, managers have to make such decisions relying on generic management softwares that don't take into account the specificities faced by the company.

The product-mix problem is traditionally solved using Linear or Mixed Integer Programming techniques. However, designing a model for a specific company can yield better results. In this work will be used Stochastic Programming as it is a more robust method to solve problems permeated by uncertainty. Stochastic Programming is applied when some of the parameters or variables become known only when the event takes place. Different approaches to solve Stochastic Problems can be found in the literature, ranging from recourse models to external sampling methods. We are interested in the latter, consisting of producing a number of scenarios big enough to be considered representative of the solution space.

### **3 Methodology**

Mathematical modeling and computer simulation techniques were used in this work to solve the product-mix problem of the plastic manufacturing firm in this study. The simulations were carried out with 50,000 MC experiments using a mathematical model designed to represent the company's operations. The results were analyzed with a statistical clustering technique.

The utilization of quantitative models to study enterprise operations has been spread-

ing over the last years. The usage of mathematical modeling in operations management evolved from being only a way to solve day-to-day problems turning into an important source of scientific development, but not without losing its empiric characteristics along this evolution process [8]. Now it is seen the need to put together these two mathematical modeling characteristics as a way to solve real operations management problems and contribute to scientific knowledge.

Quantitative models are “based on a set of variables that vary over a specific domain, while quantitative and casual relationships have been defined between these variables” [8]. The usage of this class of constructs helps to understand observed phenomena in the operations of goods and services as they describe how the events occur in total or part, allowing not only the study of events in the past but also possible future configurations of the system.

The development of computational simulation models presents some advantages when compared with direct experimentation of a system [9]. These advantages lie in three aspects: (a) cost, (b) risk and (c) time. Designing models to study certain phenomena reduces the associated costs as it does not involve the use or acquisition of high value machinery. It also preserves material and human resources that would be affected by constant changes to the ordinary operation parameters for tests. Simulation models allow managers to assess diverse scenarios, even years of operations, in only a few days or weeks, which would not be possible in a direct experimentation process.

Although there is no consensus in the literature concerning the validity of clustering methods, the k-means technique was chosen to analyze the data. Different techniques, including k-means, were applied by [10] to validate a Tibetan skulls classification with good accuracy and vouched for the validity of applying clustering techniques in data analysis. Moreover, [11] showed how clustering can be useful in automatic information searching inside big textual databases using studies of innovation and technological and scientific progress in Europe through patent databases as an example.

Preliminary reports were made and discussed with the company’s managers in dif-

ferent steps of the simulations in order to build a document with useful information able to aid them in developing a production planning. Product allocation, contribution margins and capacity utilization were the information in which the managers showed more interest.

## **4 Case Study**

The company in study manufactures and sells BOPP in Brazilian, Argentinean and international markets. Its products are made in six production lines spread over two Latin American countries, generating a portfolio of more than 200 different plastic films. Packaging manufacturers in many countries are its most important clients.

BOPP films are inputs to other industries and can be applied to the manufacturing of many types of flexible packages, including food packages in direct contact with the contained products because of their inert and nontoxic characteristics. Other uses are labeling and adhesive tapes manufacturing. One of BOPP advantages is that it is one hundred percent recyclable as recent studies show that even metalized plastic food packages can be recycled.

A BOPP production line is capable of producing a wide range of different products, varying in weight, transparency and cutting. Stochastic scrap occurs especially during extrusion and cut processes, but also in the course of machine setups, since plastic films are made in continuous process, even setups are made with the production line working. For this reason, the closer the products' characteristics allocated within a production line, the less scrap will be produced in the transition between products.

Managers from the production department of the studied company work with a product-mix that they believe to be below optimum and looked for consulting while trying to determine how capacity distribution among products could maximize the company's total contribution margin. The products are classified by markets and it may be preferable to manufacture the same product in different plants to reduce customs and transportation

costs when clients are in different countries.

Unlike production scheduling decisions, which are operational and have to be made on a daily or weekly basis, product-mix decisions are tactical and should be revised in a time frame that varies from industry to industry, but in general occur every two or three months. In the course of the project development, it was determined that a 50,000 sample should be used to run the MC simulations for each study cycle. This computational complexity would make it prohibitive for the company to revise its product-mix on feasible time and take the necessary actions to adjust the production. The solution was to resort to EELA-2 computational grid infrastructure. With a large number of computers available it was possible to run experiments and analyze the results to solve the stochastic product-mix optimization problem on reasonable time to comply with company's needs.

## **5 Model**

The model designed during this study was based on descriptions and observations of the company's production process, aiming to best represent the production lines behavior. Restrictions forcing similar products to be allocated in the same production lines were created to increase line specialization, reducing setup times and scrap levels during product transitions. The proposed model enabled the evaluation of different scrap configurations while running MC simulations to determine the product-mix.

## Sets

$I$  – Set of products

$J$  – Set of production lines

$V$  – Set of groupings

## Parameters

$p_i$  – Price of product  $i$

$c_{ij}$  – Cost of product  $i$  at production line  $j$

$d_i$  – Demand of product  $i$

$s_{ij}$  – Scrap of product  $i$  at production line  $j$

$o_{ij}$  – Output of product  $i$  at production line  $j$

$T_j$  – Time available at production line  $j$

$K_{ij}$  – Minimum quantity of product  $i$  at production line  $j$

$A_v$  – Grouping  $v$ ,  $v \in V$

## Variables

$q_{ij}$  – Quantity of product  $i$  manufactured at production line  $j$

$$y_{ij} = \begin{cases} 1, & \text{if product } i, i \in I, \text{ is manufactured at } j, j \in J; \\ 0, & \text{otherwise} \end{cases}$$

$$\text{Maximize: } \sum_{i \in I} \sum_{j \in J} \left( p_i - \frac{c_{ij}}{(1 - s_{ij})} \right) q_{ij} \quad (1)$$

Subject to:

$$\sum_{i \in I} \frac{q_{ij}}{(1 - s_{ij})o_{ij}} \leq T_j, \quad \forall j \in J \quad (2)$$

$$\sum_{j \in J} q_{ij} = d_i, \quad \forall i \in I \quad (3)$$

$$q_{ij} \geq K_{ij}y_{ij}, \quad \forall i \in I, \forall j \in J \quad (4)$$

$$q_{ij} \leq d_i y_{ij}, \quad \forall i \in I, \forall j \in J \quad (5)$$

$$y_{lj} = y_{mj}, \quad \forall l, m \in A_v, \forall v \in V \quad (6)$$

$$q_{ij} \geq 0, \quad \forall i \in I, \forall j \in J \quad (7)$$

$$y_{ij} \in \{0, 1\}, \quad \forall i \in I, \forall j \in J \quad (8)$$

## 6 Computational implementation

Grid computing infrastructures consist of shared resources (computing and storage resources, software, reading instruments like electronic microscopes, among others) that can be geographically dispersed, being organized around Virtual Organizations (VO). This type of infrastructure provides computational power to solve problems that need a great amount of processing operations and can be tackled with disjunctive computing.

MC simulations with thousands of experiments were carried out, having different scrap levels according to probability distributions inferred from historical data. A C program was designed to randomize scrap levels for each experiment. This program built the experiments with data from an input file with operational information such as product cost at each production line, line capacity, selling price of the products, product output per hour and historical average scrap.

A set of scripts (Industry@Grid) were designed to allow automatic execution of a

considerable amount of experiments on EELA-2 infrastructure, reducing the hassle and effort involved in frequent simulations. Industry@Grid profits from the grid infrastructure with three modules: (1) ModelBuilder: responsible for initializing the experiments inside the grid; (2) Runner: runs the experiments to produce results; and (3) Merger: puts together and fetches the results to the user. After execution, the product-mix solution to be implemented is chosen with k-means clustering technique, which separates the solutions in clusters using similarity measures between them. The following sections present a summarized description of each module and the k-means method.

## **6.1 ModelBuilder**

The ModelBuilder prepares the computing environment to run the simulations. The application executes a program in the grid infrastructure with data respective to products, production lines, scrap levels and more, preparing the MIP problems to be solved in batches. Due to the distributed nature of the grid infrastructure, replicas of the data are spread over different geographical areas to minimize transmission times of the data during the execution of the next module. EELA-2 had 28 computational resource centers (RC) spread over institutions in Latin America and Europe [5]. Each RC presents a computer dedicated to temporary data storage (storage elements – SE) that was employed to hold the replicas. This local storage increases efficiency in infrastructure usage.

## **6.2 Runner**

The Runner is responsible for managing execution of the MIP batches created on the previous step. A job is scheduled for execution in the infrastructure for each existing experiment batch. This module determines the best RC for execution and manages fault control and job re-submission. For example, in case 50,000 experiments are created and the user decides that the batch size will be 200 experiments, the Runner will send 250 jobs for computers located in the 28 RCs. These computers access the batches making use of a file containing a list of the experiments to be solved and the ad-

addresses of the SEs that hold the replicas and optimization software – GLPK (available in <http://www.gnu.org/software/glpk/>). At the end of each job, the result of each experiment is transferred to a SE of CEFET/RJ’s RC, located in Rio de Janeiro.

### 6.3 Merger

The Merger consolidates the execution information of all experiments stored in CEFET/RJ’s SE, generating a file with all solutions of the MIP problems regarding all experiments. In case of the 50,000 experiments example, a matrix with 50,000 lines and 1,290 columns (215 products x 6 production lines) is produced. This matrix is transferred outside the grid so the user can analyze the results.

### 6.4 K-means

After fetching the results, the analysis to determine the product-mix configuration is made applying a k-means algorithm. This method was initially proposed by [12] and has several applications, including separating objects into clusters using a measure of distance or similarity between them. Objects can be represented by vectors of dimension  $d$ , represented by  $X = \{x_1; x_2; \dots; x_n\}$ , that will be distributed over  $k$  clusters, with  $k < n$ ,  $S = \{s_1; s_2; \dots; s_k\}$  to minimize the sum of the squares of the distances amidst the objects of each cluster:

$$\min \sum_{i=1}^k \sum_{x_j \in S_i} \|x_j - \mu_i\|^2$$

where  $\mu_i$  is the average of the objects belonging to  $S_i$ . The algorithm proposed by Lloyd (1982) consists of the execution of three steps:

- (a) Take at random a set of  $k$  objects and assign each of them as the average of cluster  $S_i$ .
- (b) Add each object to the cluster with the closest average.
- (c) Recalculate the mean  $\mu_i$  of each cluster.

The algorithm repeats steps (a) and (b) until until no object changes cluster. The product-mix solution is the closest solution to the average of the largest cluster, which is considered to have the greatest likelihood, as the average may be an infeasible solution.

## **7 Results**

The most important outcome is the product-mix configuration that maximizes the total contribution margin for a given scrap configuration. This configuration has two important outcomes: (a) total contribution margin and (b) distribution of production lines capacity.

The product-mix configuration determines which products will be manufactured in each production line. Since the solution may need a few adjustments before implementation, the managers responsible for production and sales review it carefully. Among the factors that can influence changes to the final product-mix configuration is the fact that some plants have little or no experience manufacturing a determined product and the implementation could have a switching cost that is not quantified by the model.

The study of the results allowed a better understanding of how the scrap levels impact operations and profit. Experiments with high levels of waste showed lower levels margin not only because of increased manufacturing costs without increase in sales prices, but also because some products have to be rearranged to comply with capacity restrictions of production lines.

### **7.1 Product Allocation**

The changes in product allocation were assessed through careful analysis of the results of all 50,000 MC experiments. When a product appears at the same production line in at least 95% of the experiments it is considered that it can be removed from the analysis as its allocation isn't affected by changes in operation's conditions. In this work, 69 of the 215 products studied fit this case, which represents 32% of the total products. This approach reduces uncertainties faced by the company while making its product-mix decisions.

Changes in the product-mix configuration must be avoided during factory floor implementation. Once any local modification is made to this configuration, the global result may change. Alterations should be avoided because they may cause a violation of one or more constraints of the model. However, if any refinement is needed for specific reasons or a solution cannot be implemented without any changes, the information about products that appear frequently in the same production lines through different experiments can be used to assess alternatives. In this case the decision maker may have an indication of which line should produce a particular product according to the frequency with which a product appears on certain production lines.

## 7.2 Capacity Utilization

Figure 1 shows the average capacity load of each of the six production lines available. This analysis can help medium and long-term strategic decisions as: line deactivation, investment in new technologies and new facilities.

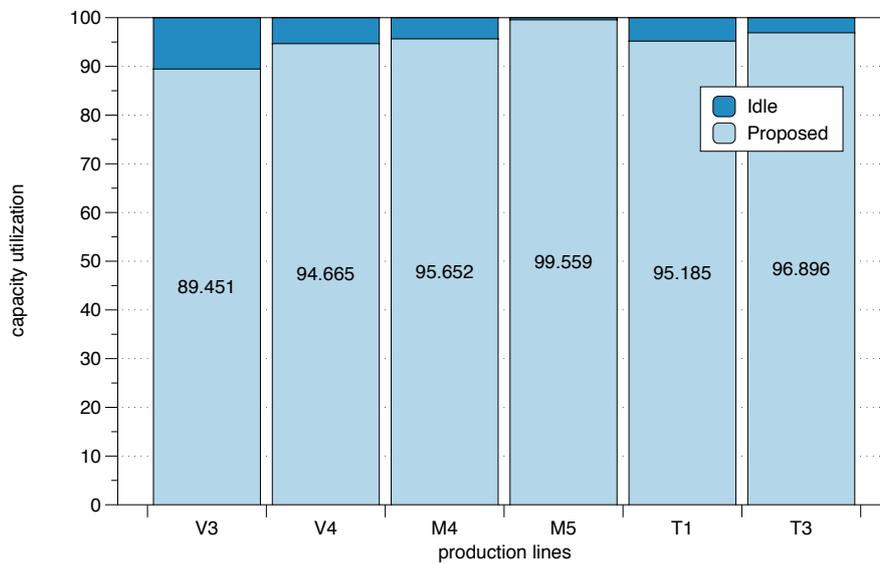


Figure 1: Average load of production lines

The fact that almost all production lines are close to maximum load is consistent with the strategy adopted by the company, seeking to maximize the usage of installed capacity due to the high capital investments and maintenance cost. If the company maintained its

previously adopted product-mix it could cause an overflow of some lines with all plants running on maximum capacity due to demand and scrap fluctuations. The gains obtained with the implementation of the product-mix achieved with the proposed approach are not restricted to the direct increase in total contribution margin, but also the fact that there is a reduction in used capacity. It is estimated that the “extra” capacity can generate a financial gain around 3 to 5% of the margin obtained with the previous product-mix.

The simulations can also help planning capacity expansion decisions. Industry@Grid is a set of tools that allows the manager to infer if the current capacity is enough to meet a given demand with various scrap levels. From a certain demand level on, part of the solutions of the model will be infeasible, indicating that the demand is greater than the installed capacity. Monitoring this indicator, the company can find the demand level in which it should make capacity expansion investments.

### **7.3 Contribution Margin**

The total contribution margin of the company varies with changes in product allocation and scrap levels. Figure 2 shows a histogram of the contribution margins for a simulation containing 50,000 MC experiments. The horizontal axis represents margin, parameterized against the total contribution margin of the product-mix configuration adopted by the company in the past year. The proposed approach gives a better solution when comparing to previous managerial decisions in 91.15% of the time, with an average gain around 5%.

This 5% increase, added to the potential increase that can be achieved using the extra capacity (3 to 5%) and the expected reduction of setup times and scrap levels during product transitions due to production line specialization, amounts to an expected increase around 10% in total contribution margin.

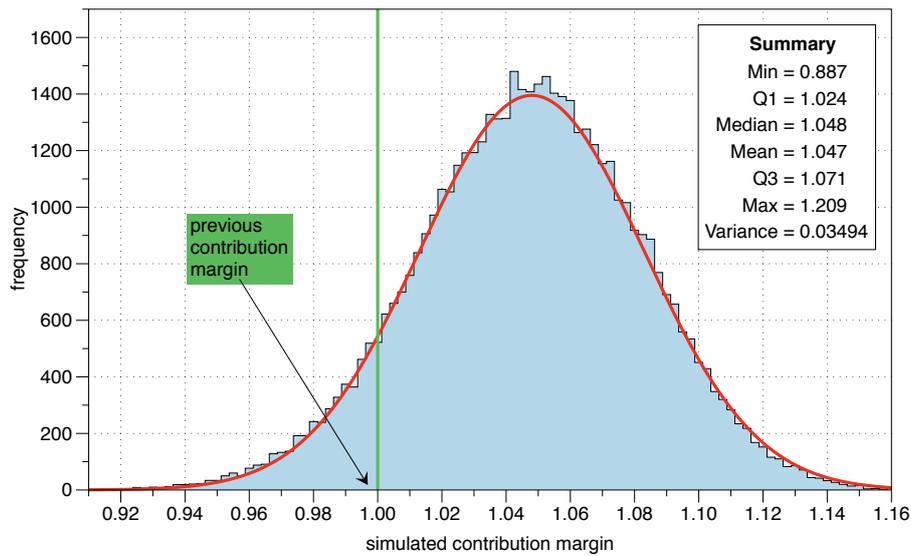


Figure 2: Total contribution margin distribution

## 8 Conclusion

The approach used to tackle the product-mix problem of the company in study gave positive results. In addition to the increase in total contribution margin, the capacity distribution among products allowed a reduction in production line load, giving the company some room to offer new products or increase the offer of its most profitable products. The reduction in the capacity utilization is, in part, given to a higher specialization of lines, achieved using restrictions that force products with similar characteristics to be allocated in the same production lines. These restrictions also bring benefits not measured by the model, such as setup time reductions and lower scrap levels throughout product transitions. Combined with computational simulation, the model showed to be a tool capable to aid in productive capacity management and planning. It can also be applied to assess varying demand levels and identify the moment in which investments have to be made to increase the performance of the production lines or even the acquisition of new production lines or facilities.

## References

- [1] S. Neoh, N. Morad, C. Lim, and Z. Abdul Aziz, “A Layered-Encoding Cascade Optimization Approach to Product-Mix Planning in High-MixLow-Volume Manufacturing,” *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, vol. 40, no. 1, pp. 133–146, 2010.
- [2] J. Birge and F. Louveaux, *Introduction to Stochastic Programming*. New York: Springer, 1997.
- [3] N. Metropolis and S. Ulam, “The Monte Carlo Method,” *Journal of the American Statistical Association*, vol. 44, no. 247, pp. 335–341, 1949.
- [4] A. Shapiro, “Stochastic programming approach to optimization under uncertainty,” *Mathematical Programming*, vol. 112, no. 1, pp. 183–220, 2007.
- [5] G. Andronico, V. Ardizzone, R. Barbera, B. Becker, R. Bruno, A. Calanducci, D. Carvalho, L. Ciuffo, M. Fargetta, E. Giorgio, G. La Rocca, A. Masoni, M. Paganoni, F. Ruggieri, and D. Scardaci, “e-Infrastructures for e-Science: A Global View,” *Journal of Grid Computing*, vol. 9, no. 2, pp. 155–184, 2011.
- [6] H. Li and S. Azarm, “An Approach for Product Line Design Selection Under Uncertainty and Competition,” *Journal of Mechanical Design*, vol. 124, no. 3, pp. 385–392, 2002.
- [7] J. Jiao and Y. Zhang, “Product portfolio identification based on association rule mining,” *Computer-Aided Design*, vol. 37, pp. 149–172, Feb. 2005.
- [8] J. W. M. Bertrand and J. C. Fransoo, “Operations management research methodologies using quantitative modeling,” *International Journal of Operations & Production Management*, vol. 22, no. 2, pp. 241–264, 2002.
- [9] M. Pidd, *Computer Simulation in Management Science*. New York: John Wiley & Sons, 4th ed., 1998.

- [10] S. Landau and B. S. Everitt, “Classification : Cluster Analysis and Discriminant Function Analysis ; Tibetan Skulls,” in *A Handbook of Statistical Analyses Using SPSS*, Chapman and Hall/CRC, 2003.
- [11] S. Sirmakessis, K. Markellos, P. Markellou, G. Mayritsakis, K. Perdikouri, A. Tsakalidis, and G. Panagopoulou, “STING: Evaluation of Scientific & Technological Innovation and Progress in Europe Through Patents,” in *1st STING User-Focus Group Meeting*, (Lausanne), 2001.
- [12] J. Macqueen, “Some Methods for Classification and Analysis of Multivariate Observations,” in *Proceedings of the fifth Berkeley Symposium on Mathematical Statistics and Probability* (L. M. Le Cam and J. Neyman, eds.), pp. 281–297, 1967.