OPERATIONS RESEARCH CENTER

working paper

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Production Allocation Remodeling System: Optimizing for Competitive Advantage in a Mature Manufacturing Industry

by

D. Bonaquist*, R. Brown**, T. Hanson*,
J. Shapiro***, V. Singhal**

OR 155-87 January 1987

* Linde, Union Carbide, Tonawanda, NY
** Resource Management Systems, Cambridge, MA
*** Massachusetts Institute of Technology, Cambridge, MA

I. INTRODUCTION

Setting production targets for geographically dispersed manufacturing sites is a common decision problem in large manufacturing companies. This is a routine decision, yet it conceals important tradeoffs between manufacturing and distribution that have a direct effect on corporate performance. The location of production clearly affects the cost of distributing products to customers, as well as service levels. Production targets also set the stage for lower level manufacturing decisions, such as production scheduling and inventory management. More generally, targets largely determine the level of utilization for manufacturing sites. At the most extreme, targets may call for a site to be shut down or mothballed.

This paper describes work we have done on a production allocation modeling system (PAMS) for the Linde Division of the Union Carbide Corporation. The system has been in use for more than a year in the company's Eastern region, and installations in other regions are underway. Work is also in progress to elevate PAMS to a national model encompassing all of Linde's important sites and customers.

Linde is a major producer of industrial gases (oxygen, nitrogen, argon, hydrogen), with numerous manufacturing sites and customers throughout the United States. The immediate purpose of PAMS is to minimize combined regional manufacturing and distribution costs over a planning period of approximately thirty days. More generally, corporate planners use it to allocate individual customer demands to geographically dispersed production sites. PAMS also optimally allocates idle time to sites in keeping with complex relationships between production costs and capabilities.

Beyond the immediate application, we believe PAMS is of general interest because it demonstrates the value of optimization in a mature industry where conventional wisdom might lead one to expect opportunities for cost reduction to be limited. In particular, it demonstrates how an integrating model can be used to bring the company's technical expertise in production and engineering to bear on the strategic goal of lowering costs to enhance competitive position.

The PAMS project also illustrates how a model and an application evolve together over the course of a project through an interplay between practical, computational, and theoretical considerations. In this case the model became both more correct and simpler -- a happy but perhaps fortuitous outcome which is by no means the rule with complex modeling applications.

The model development in PAMS is novel in that mixed integer programming (MIP) constructs for describing electricity contracts (see Bender et al (1981), Bender et al (1985) for other examples of construct analysis) are combined with manufacturing submodels and a distribution network. Moreover, the implementation successfully linked a chemical process optimization model to a mathematical programming model for tactical planning.

II. INDUSTRY BACKGROUND

Production of industrial gases is in many ways the quintessential mature manufacturing industry. The process of cryogenic distillation by which air is separated into gaseous and liquid elemental fractions has been known for over eighty years. Competing producers now operate capital intensive plants with similar intrinsic thermodynamic efficiencies; few radical breakthroughs in production technology are to be expected. Air, the sole raw material, is free and does not vary appreciably in quality. Nor is there much scope for product differentiation--except for special applications where extreme purity is essential, all liquid oxygen is very much the same.

Despite this uniformity on the supply side, however, the <u>markets</u> for industrial gases are changing, largely in response to structural changes in the national and world economy. Demand for liquid and gaseous oxygen was for many years the driving force of the industry. In recent years the rate of increase in this demand has been declining, as the centers of basic industries such as steelmaking shift offshore. On the other hand, demand for liquid nitrogen is increasing for use in food preparation, enhanced oil recovery, and other areas where a combination of very low temperature and chemical inertness is essential.

(Figure 1 shows this shift in demand over the last two decades for the industry as a whole.)

These changes have also led to a shift in the location of demand, away from the older midwestern industrial centers. For years, the bulk of the industry's production was delivered to large customers by gas pipeline, from production sites located near the customer's facility. Now, a large and increasing proportion of demand is for liquid products, which are delivered in insulated trucks or railcars to a larger number of more geographically dispersed customers.

The result has been to alter accepted premises and operating procedures. The company is no longer principally an adjunct of stable larger industries, and cannot afford to operate as if it were. This shift in the conditions underlying competition in the industry raises hazards where for decades there had been stability. It also opens up new opportunities for those companies that can been adapt to the new conditions.

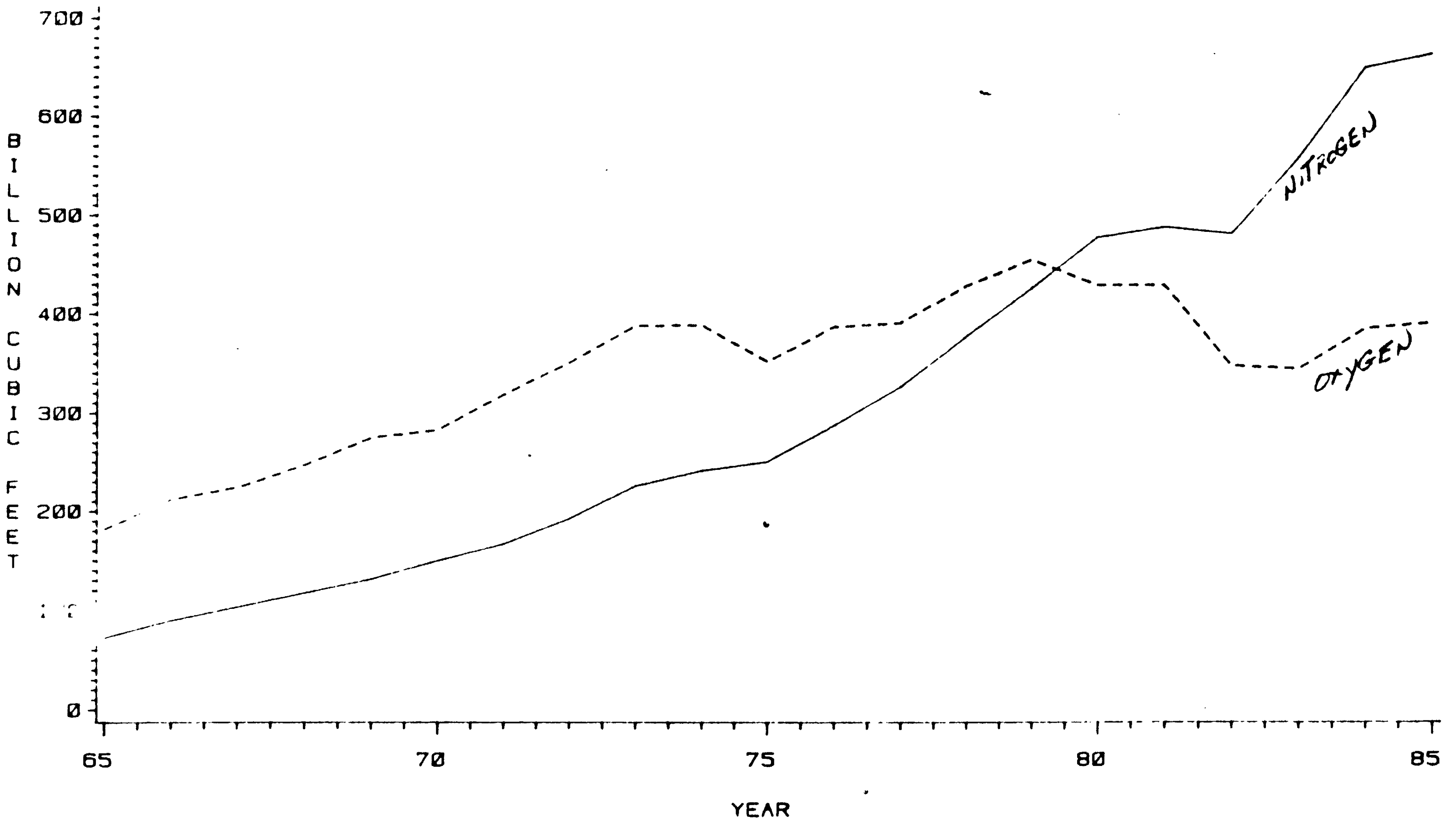

FIGURE 1.
PRODUCTION OF OXYGEN AND NITROGEN

III. PROJECT BACKGROUND

The PAMS project originated in a general desire on the part of Linde's upper management to bolster competitive position through better operation of the Linde production and distribution system. Delivered cost is one of the primary determinants of competitive advantage in this industry (the other being customer service.) The two primary elements of cost that are subject to control over the short and medium term are distribution and production. The latter are generally larger, but distribution costs are still quite significant -- typically 30% of delivered cost. It was therefore natural that Linde's attention should have focused at first on reducing each of these costs independently of the other, particularly since such an effort meshed with the current division of functional responsibilities.

For the purposes of production and distribution planning, Linde groups its customers and production sites into several large regions --East, South, Central, and so on. In principle, any site can serve any customer, provided it makes the product demanded by the customer. Within a region, known or predicted customer demands are assigned to a site through a monthly planning cycle. These demands can then be aggregated into production targets for each product at each site.

In practice, planners in the distribution function assigned customers to sites, since it was they who managed the shipment of product to customers. In making this decision, these planners made heavy use of an elaborate network model which had been developed to minimize distribution costs. This tended to favor assignment of customers to the nearest site, without taking fully into account the cost of production at the site. The production process and its economics were simply too complex to be represented well in such a model, and found no other exponent in production allocation process. Instead, region management set production rates through an informal heuristic process which attempted to reconcile forecasted product demand, relative site production costs, inventory levels, and distribution costs.

Linde also had in place a quite successful program to improve the localized efficiency of production sites. A major element of this program was the Site Optimization Map (SOM), which was developed and implemented by two of the authors. The SOM is a set of data gathering procedures and software based on nonlinear optimization techniques (including random search) to optimize the instantaneous performance of individual sites. The SOM had been developed for use at each site to determine how the site should be operated to meet given production rates while minimizing the rate of energy consumption (power demand). Although it has been and still is very successful at this localized optimization, it could not in itself determine what those rates should be.

As we reviewed Linde's procedures and tools it became clear that cost reductions in production and distribution would be at best haphazard, if not illusory, unless they were achieved in concert. Conspicuous by its absence was the ability to plan both production and distribution activities within a single, comprehensive framework to achieve the greatest overall cost reductions.

This kind of coordinated planning looked to be a relatively untapped area in which it would be possible to distinguish the company from its competitors, which generally have smaller, less complex production and distribution systems than Linde. In a competitive industry such as Linde's, cost reductions of even one or two percent can be extremely important.

IV. MODEL DEFINITION

Air separation sites produce gaseous and liquid air fractions. Gases (oxygen, nitrogen) are distributed by pipeline to customers located near the site. Liquids (oxygen, nitrogen, argon) are distributed by truck or railroad tanker. There are no joint deliveries; in fact, each vehicle is dedicated to a single product. This means that the distribution system -- and costs-- for each product are linked only through joint production at the sites.

The distribution component of PAMS is thus represented as a simple network of arcs linking production (or external supply) points to customers. In general, any site can deliver to any customer. Unit transportation cost between a site and a customer reflects the distance between the two points, and perhaps the intervening geography. The costs used in PAMS are derived from historical data and were already in use for distribution planning.

Most of the structure of the model lies in the representation of the production sites. The complexity of this representation stems from a variety of related factors, among them joint production, electricity contracts, and shut-down operation.

A. Joint Production

A site produces products jointly from the same production process -- up to five products at once. A product can be produced at any rate, within upper and lower limits that depend upon the site, the product, and the rates at which other products are being produced.

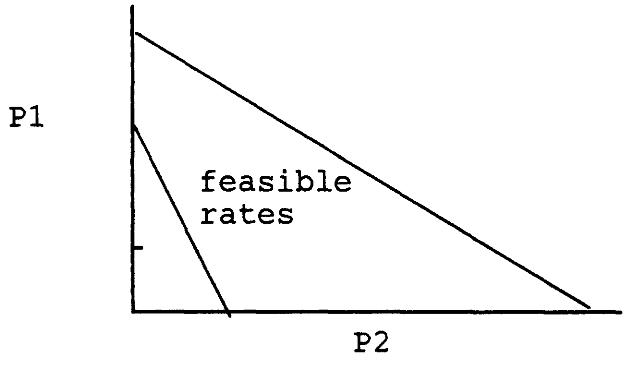

Figure 2

The (instantaneous) power demand (KW) of the site is an increasing function of production rates for all products, with strong cross terms, particularly for liquid products. There are strong theoretical and empirical reasons to believe that the surface is convex. In fact, multiple regression to a positive definitive quadratic form gives a good fit. No closed form for the function KW=f(P1,...Pn) is known.

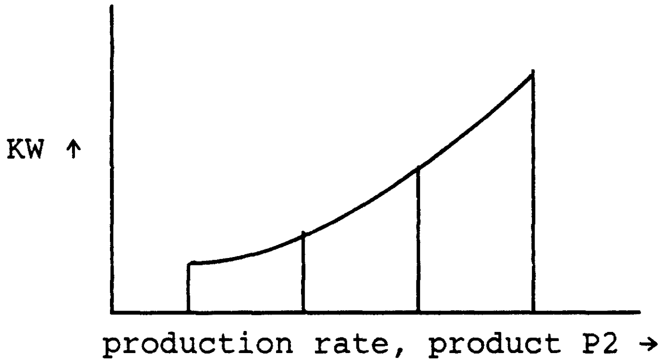

Figure 3

For modeling purposes, all our knowledge about the KW surface for a site is obtained from the SOM, which was originally developed to help site production managers operate their sites most efficiently. Given desired production rates for a set of products, the SOM will determine the minimum power demand for the site to produce at those rates. The SOM uses random search methods to find this minimum. This is standard practice in chemical engineering, where the complexity and nonlinearity of the underlying production processes makes gradient methods very difficult to implement and cumbersome to use (see Martin (1982), Wang (1978)). Random search methods also make it easier to configure the SOM to the characteristics of each site.

B. Electricity Contracts

Virtually the only variable cost is the cost of electricity used to run compressors and liquefiers, so that production cost is very closely tied to the site's power demand. A decision to assign a customer for, say, LO2 (liquid oxygen) to site A therefore implicitly alters the cost of producing LN2 (liquid nitrogen) at that site, and hence the economics of assigning an LN2 customer to site A.

But production cost at a site is not strictly a matter of thermodynamic efficiency. It is governed by contractual terms that are often quite complex and that differ, sometimes radically, from site to site.

One typical contractual feature is that the site is charged both for energy (KWH) consumption and for maximum (instantaneous) power draw (KW) during some contract billing period -- the so-called "billing demand". These costs are roughly of the same magnitude, though energy costs tend to be higher.

Under most contracts the unit cost of energy varies discontinuously by time of day. Figure 4 depicts a situation in which the day is divided into on-peak, mid-peak, and off-peak hours. Energy charges are highest during the on-peak hours, lowest during off-peak, and take on an intermediate value during mid-peak. The relative proportion of on, off, and mid peak periods in a weekday, weekend day, and holiday may all be different. Any period type may be absent from any day type.

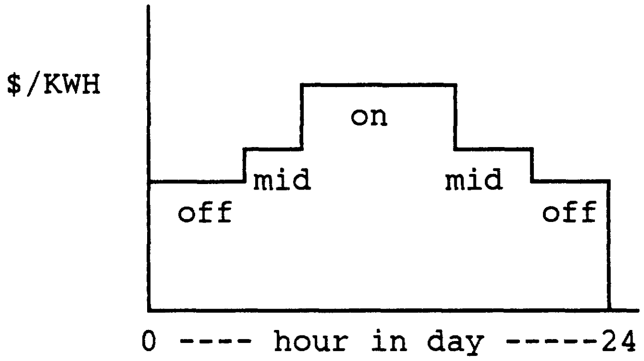

Figure 4

Time of Day Energy Pricing

The charge on billing demand ($/KW) may also be different in on, off, and mid-peak periods, although not in proportion to energy charges. Often billing demand charges are only incurred during certain periods.

Under such contractual terms, there is a strong incentive to produce at higher rates during off-peak periods, when energy and power are cheaper, and to throttle back during more expensive (i.e. on-peak) periods.

C. Shut-down Operation

Linde has excess production capacity in some regions. Gaseous products cannot be inventoried, and inventory capacity for liquids is limited. Therefore, it is often necessary to put a site into standby mode for some part of the month.

If left to itself, an LP model would choose to shut a plant down during on-peak hours, when energy and power are both most expensive. In practice, such a solution would be impractical for operational reasons. While the purpose of PAMS was not to <u>schedule</u> site production day by day or hour by hour, it was vital that the solutions be operationally feasible. It was therefore necessary to impose a kind of loose parity between the length of time the site would be shut down during on, off, and mid-peak.

D. Slates: Discretizing the Decision Space

Both energy and power costs can be very significant. Since both are directly related to power demand (KW) it was clearly important to represent these relationships with fair accuracy. One approach might have been to use quadratic programming to describe energy costs, but this was rejected for several reasons. First, there are no commercial grade QP codes capable of handling MIP constructs. Also, we had at best only an empirically derived quadratic KW function, based on regression.

Instead, we chose to discretize the space into a large number of production <u>slates</u>. A slate is a vector containing a production rate for each product, and the minimum power demand associated with operating the site to produce at those rates. The basic decision of the model, therefore, is to determine how long to operate each potential slate.

This itself would have presented little problem, since LP is quite able to represent a convex cost function. However, the cost of <u>power</u> (so called demand charge) is based on billing demand, that is, the maximum instantaneous power demand over the entire period. Thus, this cost would be incurred only by the set of production rates used during the period that resulted in the greatest power demand, regardless of how long that slate was operated.

V. MODEL FORMULATION

We present here the original MIP formulation upon which PAMS was based. Experience with MIP models drawn from this formulation, both prior to and after their application to actual planning problems, led to a number of modifications and simplifications. These are discussed briefly at the conclusion of this section. In the following section, we discuss our approach for implementing the system based on these models, and experience with the system.

Indices

i: 1 to I	index for plants
j: 0 to J	index for slates at each plant (slate 0 is plant shut-down)
k: 1 to K	index for products
m: 1 to M	index for customers

Parameters

P_{ij} = power draw for jth slate at plant i (KW)

e_i = electric energy charge at plant i ($ per KWH)

E_i = electric power demand charge at plant i ($ per KW)

c_{ikm} = cost of transporting one unit of product k from plant i to customer m ($ per cubic foot)

a_{ijk} = instantaneous production rate of product k by jth slate at plant i (cubic feet per hour)

d_{km} = demand for product k by customer m (cubic feet)

R = minimum run time for any slate at any plant (hours)

T = length of planning horizon (hours)

Variables

t_{ij} = length of time plant i uses jth slate (hours)

W_i = maximal power demand at plant i (KW)

$$x_{ij} = \begin{cases} 1 \text{ if jth slate at plant i is used at a positive level} \\ 0 \text{ otherwise} \end{cases}$$

y_{ikm} = quantity of product k shipped from plant i to customer m (cubic feet)

<u>Production Allocation Model (PAM)</u>

$$\text{minimize} \sum_{i=1}^{I} \left\{ \sum_{j=1}^{J} e_i P_{ij} t_{ij} + E_i W_i \right\} + \sum_{i=1}^{I} \sum_{k=1}^{K} c_{ikm} y_{ikm} \qquad (1)$$

Subject to:

For i = 1,...,I

$$\sum_{j=1}^{J} a_{ijk} t_{ij} - \sum_{m=1}^{M} y_{ikm} \geq 0 \quad \text{for } k = 1,\ldots,K \qquad (2)$$

$$\sum_{j=0}^{J} t_{ij} = T \qquad (3)$$

$$\left.\begin{aligned} t_{ij} - Rx_{ij} &\geq 0 && (4a) \\ t_{ij} - Tx_{ij} &\leq 0 && (4b) \\ W_i &\geq P_{ij} x_{ij} && (4c) \end{aligned}\right\} \text{ for } j = 1,\ldots,J$$

For m= 1,...,M

$$\sum_{i=1}^{I} y_{ikm} = d_{km} \quad \text{for } k = 1,\ldots,K \qquad (5)$$

$$t_{ij} \geq 0, \quad W_i \geq 0, \quad x_{ij} = 0 \text{ or } 1, \quad y_{ikm} \geq 0 \qquad (6)$$

The objective function (1) in this model is the sum of energy costs, energy power demand costs, and distribution costs. Note that energy and power costs differ from plant to plant. This is because the contracts with electric utilities vary by location, and furthermore, each plant has its unique design and operating characteristics. Note also that the slates available for use at each plant, and their costs, are uniquely associated with that plant. We have chosen the fixed number J of trial slates for each plant simply for expositional convenience.

The constraints (2) state that the total quantity shipped from each plant cannot exceed the total production. In practice, the inequality was extended to account for small quantities of beginning and allowable ending inventories. The constraints (3) state that the entire planning horizon is consumed at each plant by production time and down time (recall that slate 0 is the plant shut-down slate). The constraints (4a) and (4b) state that the time t_{ij} that the jth slate is used at plant i, if it is used at all, must lie between the conditional minimum R and the maximal allowable time T. The upper bounding constraint in (4b) is redundant in the light of constraint (3); we have included it for expositional purposes. Constraint (4c) ensures that the power demand W_i upon which the power charge is based equals the maximum of the power demand draws among slates selected by the model for plant i. The constraints (5) state that demand must be met by shipments from the plants. We note that most customers demand only one product. Thus, the total number of constraints (6) is far fewer than KM.

The specific models generated by PAMS turned out to be more complex than (PAM) for several reasons. First, the model was extended to distinguish among peak, mid-peak and off-peak operations when the electricity rates vary significantly. Plant shut-downs were modeled more extensively to ensure that shut-down periods occur contiguously. Moreover, contracts with the electric utility may be more complicated, involving, for example, terms relating to differences in power draws between peak and off-peak periods. These complications were modeled by straightforward extensions of the modeling techniques used above. Finally, for complex manufacturing sites involving several interconnected plants, the models were extended so that they would choose the plant configurations as well as the slates for each plant.

Even without these extensions, (PAM) is a large scale MIP model of the fixed charge variety. In particular, the power demand charges associated with the W_i behave in a manner similar to fixed charges. Tricks involving cutting planes on the plant objective functions derived from an optimal LP solution proved relatively effective in causing the models to produce good solutions quickly. A uniform reduction in size of the demand charges E_i relative to the energy charges e_i also caused the branch and bound to work more efficiently. A second pass through the MIP optimization with the best solution from this heuristic as incumbent required far less CPU time than that required from a cold start without an incumbent.

Feedback from users at the plants led to an important simplification that allowed the models to be still more rapidly optimized. For the purposes of monthly planning, the people running the plant prefer to employ one slate for each contract period (peak, mid-peak, off-peak). The slate suggested from an optimal solution to (PAM) for each contract period is the convex combination of the slates where the weights are the fractions of the time that a slate is used. Since the surface of the cost vs. slate function for the plants studied thus far has empirically proven to be convex, we have been able to relax the corresponding MIP constructs in optimizing the model. However, MIP constructs are still required to properly model shut-downs.

VI. **IMPLEMENTATION AND RESULTS**

PAMS was implemented for an IBM mainframe computer using the LOGS model generation language (see Brown et al (1986)) and the IBM optimization package MIP/370.

It is important to emphasize that the LOGS model generation in PAMS produces a family of models. The precise formulation of a model for a specific region consisting of several plants depends on the data passed to it. For example, depending upon whether a certain contractural element is present in the data, certain structures may or may not be present in the model. We reiterate that the model (PAM) discussed in the previous section was merely the point of departure for our implementation work, and the creation of a generator for a family of models.

The MIP models generated thus far for the Eastern Region have tended to be quite large. As many as 1000 slates for each of three plants were generated by the Site Optimization Map and included in the PAMS models. Moreover, the models incorporate upward of 1000 customers demands over a typical monthly planning horizon. Automatic customer aggregation procedures were implemented, but have not yet been extensively used. The resulting models have a few thousand rows and as many as 10,000 columns. Using the simplifications and approximations outlined above, the models are usually optimized, at least to a close first approximation, within a few CPU minutes on an IBM 3083 computer.

We believe that the use of PAMS in the Eastern Region has lead to shifts in the prevailing production and distribution patterns. However, as is often the case in real-world applications, it is difficult to substantiate this belief with experimental results, for the simple and obvious reason that PAMS is not run in an experimental context. Customer demands fluctuate from month to month, and there is no "control" process to show what would have been done in the absence of a model.

A "base case" was run early in the project, in which PAMS was used to second guess a recent month's allocation decisions. The model solutions showed an increase in distribution costs, with a decrease in production costs that more than compensates for this increase. Overall, the estimate is that PAMS produces monthly production/distribution strategies that are 1% to 2% lower in total cost than solutions that would have been obtained without it.

VII. CONCLUSIONS AND FUTURE RESEARCH

PAMS has proven itself to be a useful and important planning tool at Linde. Its success demonstrates once again that computer technology has at last reached a level of development permitting mathematical programming models to be implemented and effectively applied to business planning problems. The success of this project was also due to a felicitous blending of scientific skills and experience in chemical engineering, mathematical programming, and computer systems design and programming. Finally, the support of Linde's top management in supporting a radically new approach to planning was crucial to the project's success.

PAMS is currently being extended for use in other Linde national regions. In this regard, experimentation with the Site Optimization Map is required for those sites consisting of several production plants that can be linked in different ways. Two of the authors (Hansen and Bonaquist (1986)) have developed an MIP model for calculating slates for these more complex sites. A related area of future experimentation is to link the Site Optimization Map more directly to the PAMS models via price directed decomposition methods (see Shapiro (1979)). The idea would be to occasionally use shadow prices from the mathematical programming model to price out slates produced by the SOM, and select new slates for the PAMS model.

Once models for all Linde's regions have been developed, the intention is to construct a longer range, national model for strategic planning purposes. The types of problems to be addressed by such a model include contract negotiations with customers and electric utilities, long term plant shut-downs, and economic evaluations of new markets.

Moving in the other direction with respect to time and scope, a new project is underway to convert the production planning sub-model in PAMS to a production scheduling model. The reader may have noted that the model (PAM) selects an optimal combination of slates, but makes no attempt to schedule the sequence in which they should be used. In the short-term when we consider distinct production periods with varying demands on the plant, and recognize that inventory storage for gas is extremely limited, the sequencing of slates becomes important. These slates can be viewed as fine tuned adjustments of the tactical planning slates selected by PAMS.

Finally, generalizations of the models developed for PAMS should be applicable to other process manufacturing industries. The underlying principle in performing modeling research in this area is to better understand how to imbed process control optimization models, which provide an instantaneous prescription for the plant, in one or more mathematical programming models for production planning and scheduling. We believe the models in PAMS are an important step in this research direction.

VIII. REFERENCES

P. S. Bender, R. W. Brown, M. H. Isaac, and J. F. Shapiro, "Improving Purchasing Productivity at IBM with a Normative Decision Support System," Interfaces, 15, May-June, 1985, pp 106-115.

P. S. Bender, W. D. Northup and J. F. Shapiro, "Practical Modeling for Resource Management," Harvard Business Review, 59, March-April 1981, pp 163-173.

Bonaquist and Hansen (1986) SOM in preparation

R. W. Brown, W. D. Northup and J. F. Shapiro "LOGS: A Modeling and Optimization System for Business Planning," pp 227-241 in Computer Assisted Decision Making, edited by G. Mitra, North-Holland, 1986.

D. L. Martin and J. L. Gaddy, "Progress Optimization with the Adaptive Randomly Directed Search," AIChE Symposium Series, 78, 1982, pp 79-107.

J. F. Shapiro, Mathematical Programming: Structures and Algorithms, John Wiley and Sons, 1979.

B. Wang and R. Luus, "Reliability of Optimization Procedures for Obtaining Global Optimum," AIChE Journal, 24, 1978, pp 619-626.