Large-scale industrial production processes face increasingly tight environmental constraints, which can be addressed through costly but relatively simple end-of-pipe solutions, or through cheaper but more subtle pollution prevention approaches. Achieving the process improvements necessary for pollution prevention is challenging due to the inherent complexity and unpredictability of several types of processes found in the food processing, pharmaceuticals, biotechnology, and specialty chemical industries. We propose an iterative procedure to achieve process improvements through model-based process redesign. This procedure is based on successive convex approximations of the process performance model, where product flows and process settings are optimized for a given configuration and the solution and dual variables of this optimization problem are used to update the process configuration following a greedy capacity reallocation procedure. We implemented this procedure over a five-year period at Cerestar, a major European producer of starch products, which led to a dramatic simplification in process configuration. Reduced energy and water consumption led to an estimated $3 million annual cost savings. Moreover, the reduction in environmental impacts allowed Cerestar to maintain current production levels without investing $100 million in additional wastewater treatment capacity to comply with new environmental constraints.

1. INTRODUCTION

Environmental pressures on manufacturing firms are increasing due to tightening legislation and increasing pressure from customers and nongovernmental organizations. Skea (1995) describes possible responses firms can adopt to tightening legislation, ranging from installing end-of-pipe technology (such as building a water treatment facility, or installing gas scrubber systems), to pollution prevention (e.g., by installing clean technology or by process improvements through redesign and operational procedures).

Though pollution prevention is both economically and environmentally preferred, end-of-pipe measures have been the traditional approach for many firms, especially in the United States (Graedel and Allenby 1995, p. 80). This is due to several factors. A command-and-control regulatory environment such as that prevailing in the United States tends to prescribe specific end-of-pipe measures rather than giving firms the flexibility to find more efficient pollution prevention technologies. Moreover, it is often easier, although more costly, to install end-of-pipe measures, as by their very nature they do not normally require actual process change or significant managerial effort.

One result of this tendency towards end-of-pipe measures is that firms often associate environmental regulation with high costs. Indeed, “the add-on nature of end-of-pipe technology inevitably pushes costs upwards” (Skea 1995, p. 389), sometimes even dramatically so, as the case below illustrates. By contrast, a major advantage of pollution prevention measures, including those described here, is that they often require little or no capital investment at all.

The United States formally recognized these issues by introducing the Pollution Prevention Act in 1990, stating that “source reduction is more desirable than waste management and pollution, yet opportunities for source reduction are often not realized” (Freeman 1995, pp. 28–29). The Pollution Prevention Act also includes a hierarchy of preferred waste management approaches from source reduction, recycling, and treatment, to disposal. The framework proposed in this paper suggests how source reduction and recycling can be achieved by simplifying the process through process redesign, rather than relying on installation of new equipment.

To see why simplification can lead to process improvement and pollution prevention, it is important to recognize that many plants in the process industry have grown in a haphazard way. As the original capital investments are immense (often billions of dollars), companies tend to adapt to changes in environmental legislation or in market
demand by constantly tinkering with their processes, incrementally adding processing capacity and pipes between stages, rather than starting from scratch. This leads to excessively complex processes in which operators set control parameters (such as temperature, flow, pressure, etc.), without a clear understanding of the impact on process performance. This in turn leads to significant and unpredictable variations in productivity and environmental performance. Of course, simplification by redesign is easier said than done; it requires blending detailed process knowledge with rigorous analysis, often difficult to do when managers’ overriding concern is to keep the process operating at all cost. In addition, food processing, pharmaceutical, biotechnology, and specialty chemical industries tend to exhibit a high degree of complexity and unpredictability due to the biological agents involved (such as enzymes), and generally have tighter specifications than traditional chemical refinery models. A comprehensive discussion on complexity and challenges in predictive modeling in the process industries can be found in Camacho and Bordons (1995).

In this paper we propose a methodology, illustrated with a detailed example, of how such process redesign can be guided by an iterative model-based approach. In particular, we describe the work performed at the largest wheat starch extraction process at Cerestar, Europe’s leading producer of wheat- and corn-based starches. Vast quantities of water are used in obtaining starch from wheat. New environmental legislation, enacted by the Dutch government in 1993, required a drastic reduction of several types of wastewater contaminants commonly released by large-scale process operations. Factories, including Cerestar’s plant in Sas-van-Gent in the Netherlands, were expected to comply with these standards no later than January 2000. In response to the new standards, Cerestar conducted an engineering study, which found that these standards could be met in two ways. The first was to expand the current wastewater facility by adding a new wastewater treatment process, requiring a fixed investment of $100 million, which would allow maintaining the current level of wastewater discharge. The other option was to reduce the quantity of water discharged by the process. The engineering study found that if water discharges could be reduced by at least 30%, the current wastewater treatment system would be able to meet the new standards.

Unsurprisingly, Cerestar found the option of reducing wastewater discharge far more attractive. This, in turn, could be achieved in two ways: reducing production com- mensurately, or using fresh water more efficiently. Cutting back on production was obviously very undesirable, given the capital investment tied up in the plant and given that this particular process was becoming the bottleneck for Cerestar’s downstream operations and directly supplied the refinery and modified starch channels, Cerestar’s most profitable and high-volume products. In fact, demand for these wheat-based starch products was expected to rise even more because of market trends in Europe, as the controversy around genetically modified corn (The Economist, February 5, 2000) reduced demand for corn-based starches.

Because Cerestar had five years to implement a solution and they could build the new wastewater treatment process within a year if needed, they decided first to focus on reducing fresh water consumption without cutting back on production. To this end, Cerestar worked with the first author over a five-year period. It turned out that the profitability of the process could be increased dramatically by simplifying the process, removing unnecessary links between processing stages, identified by an iterative mathematical programming approach. The environmental and economic benefits of the project were substantial: Energy and fresh water consumption were reduced by 30% (50.4 MWH per day) and 50% (2,500 m$^3$ per day), respectively, and annual cost savings were $3 million. Cerestar’s management also gained a deeper understanding of the drivers of process performance. Most importantly, the reduction in fresh water consumption enabled Cerestar to meet the new stricter discharge specifications without investing in the $100 million expansion of the wastewater facility.

We first review some relevant literature in §2. In §3 we outline the overall methodology for process redesign. We then describe Cerestar’s process in §4. Section 5 contains the model formulation, §6 the optimization procedure. The implementation is discussed in §7. The results and insights obtained from the study are presented in §8. Section 9 summarizes key lessons.

2. LITERATURE REVIEW

There is a rich tradition of applying operations research (OR) techniques to environmental problems. Bloemhof-Ruwaard et al. (1995) and ReVelle (2000) provide reviews; a rapidly growing area is that of reverse logistics, reviewed in Fleischmann et al. (1997). Recent issues of Operations Research also include examples, such as Degraeve and Koopman’s (1998) study of methods to help decide how to achieve European air quality standards, and Stuart et al.’s (1999) application of mathematical programming to lifecycle modeling.

The OR literature on managing environmental problems in manufacturing processes is less extensive. Angell and Klassen (1999) and Corbett and Van Wassenhove (1993) discuss how proven concepts from operations management can be applied to environmental management too, but that not much work has been done yet in this direction. Corbett and Pan (2002) and Madu (1996) explore environmental applications of statistical process control for process improvement, and Greenberg (1995) describes how mathematical programming can be used for environmental quality control purposes. Gupta and Flapper (1999) discuss operational aspects of environmentally conscious manufacturing (ECM).

Other OR techniques have also been used to (re)design and improve processes. Aldowaisan and Gaafar (1999)
used linear programming to redesign a complaint resolution process. Mollaghan et al. (1998) propose combining neural networks with simulation modeling. Rajala et al. (1997) propose simulation modeling and value analysis for redesign of order management and inventory control processes. In a relatively rare application to a (semi) continuous flow process, Watson (1997) discusses how simulation was used to help design, operate, and improve a batch-process chemical facility.

However, the process improvement project described here is fundamentally more complex than those reported in these papers, due to the inability to formulate a generally valid process model. To overcome this difficulty, we develop and implement a detailed iterative procedure for process redesign, which simultaneously allows significant improvements in environmental and economic performance. As such, it is consistent with the pollution prevention approach. While the specific application is towards food processing, the procedure is general enough to apply to a variety of industries.

3. FRAMEWORK FOR MODEL-BASED PROCESS REDESIGN

In this section, we propose a framework for model-based process redesign. This framework applies to industrial multistage semi-continuous and continuous production processes with a large number of links and recycling loops between stages. This is typical of many food processing plants as well as standard and specialty plants found in (among others) the biotechnology, chemical, paper, and pharmaceuticals industries. A key characteristic of such plants is that the performance of any stage depends on many operational parameters, including pressure, temperature, humidity, raw material composition, reaction characteristics, and operator behavior. This dependence is highly complex and nonlinear, and no theoretical expressions exist to capture it. Hence, such expressions must be derived empirically, through observation of and experimentation at each stage over a period of time.

In addition, the performance of each stage also depends on the configuration of the plant: The existence of a link or a recycling loop between two stages affects the nature of the flows (for instance, the ratio of pure fresh water to impure recycled water), and hence also the performance at those stages. Adding or removing a section of piping anywhere in the facility immediately affects the performance of several stages in unpredictable ways. The expressions defining performance of those stages must then be reestimated empirically. As a result, one cannot hope to formulate an explicit model describing process performance at each stage for all possible plant configurations: This would require physically operating the plant under each of many possible configurations and estimating the corresponding performance expressions.

We propose an iterative procedure (see Figure 1) to help redesign the process despite these modeling challenges. In developing this procedure, we assume that the process is sufficiently well under control that process performance can be reasonably explained by the estimated process model, rather than being largely due to exogenous uncontrollable factors. This assumption is reasonable for the types of manufacturing process we consider. Under this assumption and based on our practical experience, we have observed that once the plant configuration is fixed and in place, the performance expressions can be estimated empirically using operational data obtained while running under that particular plant configuration, and the process can then be modeled as a nonlinear programming problem. Thus, we start with the current configuration, and estimate the performance expressions and determine the optimal pattern of flows using nonlinear programming. The solution of the nonlinear program informs us which flows are zero and hence which pipes can be removed. The dual variables guide us in the reallocation of capacity in the process. We find the point at which the marginal benefits of reallocating capacity are outweighed by the costs of the physical change. We then physically make all changes in the plant required to implement the configuration changes of removing pipes and reallocating capacity, and then re-estimate the performance expressions for the new configuration and determine a new optimal flow pattern. This procedure is repeated until no marginally profitable physical layout changes can be found.

4. PROBLEM DESCRIPTION

Cerestar is Europe’s leading manufacturer of corn and wheat-based starch products. To produce these products, Cerestar relies on highly automated and capital-intensive large-scale industrial processes. The ideas presented here were developed at their largest wheat starch extraction process, located at their flagship plant in Sas-van-Gent in the Netherlands.

Figure 1. Methodology for model-based process redesign.
Wheat from North America and Europe arrives in barges and is stored in large silos. It is then passed to the mill where it is ground and sieved in a series of steps, yielding flour of the desired quality along with derivatives such as oats, bran, and wheat skin, which are sold as animal feed. The flour is transported to the wheat starch extraction process. The objective is to extract two types of starch (A and B, where B starch is more refined than A) and to produce gluten, an important derivative used in the baking and pet food industries. The extracted A and B starch slurries (mixtures of starch and water) are refined to form a range of sugars (glucose) or modified to form specialty starches. These products are used in several industries such as paper manufacturing, food and beverage processing, pharmaceuticals, and specialty chemicals.

The process begins when the flour from the mill is mixed with water and kneaded in the dough step, to develop proteins called vital gluten, an important derivative. The water-soluble gluten is removed in the dilution step along with a small amount of fibers known as pentesones. The product flow now consists of flour made up of the A and B starch components and the non-water-soluble gluten components. The A starch, B starch, and gluten are separated by sieves of different sizes. The A and B starch streams are washed to remove impurities and are then passed on as raw material to the downstream processes such as the glucose refinery and modified starch process. The gluten is also washed at the gluten-watering step and then dried, ground, sieved, and held in silos. Each of these major steps itself consists of several smaller steps, which we refer to as stages. At each stage of this process, wastewater can be discharged and this is collected and sent to the wastewater treatment plant.

Figure 2 depicts the major steps and product flows at the wheat starch extraction process, including the downstream processes (which were not part of the analysis performed in this paper).

The profitability of the process depends on the yield of A and B starch and gluten and the fresh water and energy expended to achieve this yield. The average input to this process is 1,000 tonnes of flour per day. It is operated continuously with three 8-hour shifts per day, and is shut down only twice per year for five days for routine maintenance. Product inflows to each stage are driven by a set of critical control variables, including temperature, pressure, and processing rates (controlled by, among other things, compressor speed). These controllers are set at the beginning of each shift to optimize the yields across all products and to achieve the production levels required to meet the demand from the downstream refinery and modified starch processes. The settings of the control variables also affect energy and fresh water consumption. Energy consumption is convex increasing in the controller settings, which combined with convex-increasing pricing causes energy costs to increase steeply. Increasing energy consumption at certain stages can improve yields, but only marginally.

Increasing fresh water intake at each stage improves the yield of all three product streams. However, the costs of fresh water consumption are also convex increasing. Consequently, water streams are often recycled. Excessive recycling can lead to deterioration of product yield and quality at each stage. Measures of product quality include protein, fibers, pentesones, and impurities per unit of product; for

**Figure 2.** Product flows at the wheat starch extraction process.
each measure, an acceptable range is defined. Wastewater discharge costs are assessed by the government in proportion to the volume of fresh water consumed. This is to prevent having to install, monitor, and maintain discharge flow meters at sewer drains at several industrial sites. However, to get fresh water supply from the government-run utility, Cerestar is required to maintain a water quality compliance certificate, which needs to be updated periodically.

Three types of decisions arise in this process. The first concern is the configuration: deciding which stages to connect to fresh water and which stages to connect with recycle links. Maximum flows for each link depend on pipe and pump capacity. Configuration decisions are usually made once every six months and implemented during a maintenance shutdown. The second concern is the short-term operational control decisions made during every shift to meet current downstream demand: determining controller settings and the fresh water and recycle flows. The third is determining the amount of flour that needs to be fed in to the process, given its configuration, controller settings, and the fresh water and recycle flow rates. Initially we treat the flour input as fixed (as it is not a control parameter of the process itself), but later we show how optimizing over flour input can help Cerestar negotiate contracts with wheat suppliers.

The new wastewater standards compelled Cerestar to explore methods to reduce fresh water consumption without cutting back on production or investing in wastewater treatment capacity. The current process was highly complex, as the number of control options, interconnected flows between stages, and recycling links had grown in a haphazard way. This complexity in turn contributed to poor operational procedures, as there was no clear understanding of cause and effect or of best practices. More critically, process complexity and operational procedures combined to result in an inefficient process that consumed far higher fresh water and energy than strictly necessary. Thus, the key step to reduce these consumption levels was to simplify the process by appropriate reconfiguration and by improving operational procedures, for which we develop the mathematical model described next.

5. MODEL FORMULATION

In this section, we use the economics of the process and its operating constraints to develop an optimization model. This model is used to guide the decisions for process simplification by reconfiguration and improvement of operational procedures, required for the reduction of fresh water and energy consumption. Given that the behavior of each individual stage depends on the plant configuration in a nonlinear way for which no theoretical expressions exist, representing the process in an optimization model was a major challenge. We represent this model by a nonlinear integer program, in which for a given configuration, one can estimate the process performance constraints, and this model reduces to a continuous nonlinear program, whose solution and dual variables then can be used to develop a new configuration. To provide a precise definition of this model, let \( i \in I \) index the set of process stages, \( j \in J \), the set of process controllers at stage \( i \), and \( k \in K \), the set of product quality constraints at stage \( i \). Let index \( i = A, B, G \) when it represents the last stage of the A starch, B starch, and gluten steps, respectively. We define the following decision variables and parameters:

**Decision variables.**

- \( y_i = \begin{cases} 1 & \text{if stage } i \text{ is connected} \\ 0 & \text{otherwise} \end{cases} \)
- \( w_i, W \) Amount of fresh water supplied to stage \( i \) (m\(^3\)/hr); total fresh water consumption is given by \( W = \sum_{i \in I} w_i \).
- \( x_{i_1 i_2} = \begin{cases} 1 & \text{if a recycling link connects stage } i_1 \text{ to } i_2 \\ 0 & \text{otherwise} \end{cases} \)
- \( \alpha_{i_1 i_2}, \alpha_i \) Flow of recycled water from stage \( i_1 \) to stage \( i_2 \) (m\(^3\)/hr); the total flow of recycled water into stage \( i \) is given by \( \alpha_i = \sum_{j \in J} \alpha_{ji} \).
- \( X \) Configuration of plant, as defined by all \( y_i \) and all \( x_{i_1 i_2} \).
- \( \beta_{ij}, \beta_i, \beta \) \( \beta_{ij} \) is the value of controller \( j \) at stage \( i \); \( \beta_i = \{\beta_{i1}, \ldots, \beta_{ij}\} \) the vector of controllers at stage \( i \); \( \beta = \{\beta_1, \ldots, \beta_i\} \) the vector of vectors of controllers at each stage \( i \) in the entire process.
- \( F_A, F_B, F_G \) Outflows from the last stage of the A starch, B starch, and gluten (G) steps, respectively (m\(^3\)/hr).
- \( F_{AR}, F_{AM} \) Flows of A starch to the refinery (AR) and the modified starch plant (AM), respectively (m\(^3\)/hr).

**Parameters and functions.**

- \( \bar{w}_i \) Maximum fresh water intake at stage \( i \) (m\(^3\)/hr).
- \( \bar{\alpha}_{i_1 i_2} \) Maximum flow of recycled water from stage \( i_1 \) to stage \( i_2 \) (m\(^3\)/hr). This is determined by the capacity of the pump that generates the recycled flow from stage \( i_1 \) to stage \( i_2 \).
- \( V \) Total flour input (in kg/hr).
- \( F_R^0, F_M^0 \) Total flow required at the refinery (R) and at the modified starch plant (M) (m\(^3\)/hr).
- \( \theta_A, \theta_M \) Maximum fraction of (less refined) A starch in inflows to refinery and modified starch processes, respectively.
The profit margin defined as the difference between price and cost of materials and labor (in $/kg), the required density (in kg/m³), and the required proportion of dry substance of each of the three output flows $x \in \{A, B, G\}$ for A starch, B starch, and gluten (G), respectively.

\begin{align*}
\pi_x, \rho_x, d_x
\end{align*}

\begin{align*}
Q_k(\mathbf{b}, w_i, \alpha_i; \mathbf{x}) & \text{ Function for the } k \text{th quality measure at stage } i \text{ (in units per kg), as a function of controller settings, fresh water intake, and water flows into stage } i; \text{ the function depends on plant configuration } \mathbf{x}. \\
\beta_{ij}, \bar{\beta}_{ij} & \text{ Minimum and maximum values for } \beta_{ij} \text{ (in units).} \\
\bar{Q}_k, \bar{\bar{Q}}_k & \text{ Minimum and maximum values for } Q_k \text{ (in units per kg).} \\
C_i(\mathbf{b}; \mathbf{x}) & \text{ Flow response function for stage } i \text{ (m³/hr) representing the processing rate net of the wastewater discharge rate at that stage. This is a function (defined below) of the controller settings } \mathbf{b}, \text{ and of configuration } \mathbf{x}. \\
E(\mathbf{b}) & \text{ Total energy consumption per hour as a convex-increasing function of controller settings } \mathbf{b}, \text{ in kilowatt hours per hour (kWh/hr).} \\
\Psi_E(E) & \text{ Total energy cost per hour as a convex-increasing function of energy consumption per hour } E \text{ ($/hr).} \\
\Psi_w(W) & \text{ Total water consumption and discharge costs per hour as a convex-increasing function of fresh water consumption ($/hr).}
\end{align*}

We first describe the model, then give the mathematical formulation. The objective (1) is to maximize the total profits of the output flows for A starch, B starch, and gluten (G), taking the total energy and water costs into account. Two sets of constraints, (2) and (3), ensure consistency of mass and flows throughout the process. The mass balance constraints (2) ensure that the total mass of outflows does not exceed the total mass of flour entering the process. The flow balance constraints (3) ensure that the total outflow (water plus dry material) does not exceed the total intake of product from upstream stages and net inflow of fresh and recycled water; the total product intake from upstream stages is given by the flow response function, which we discuss below. We represent the flow balance constraint (3) as an inequality, as operators always have the option to discharge outflows directly to the wastewater treatment process. Quality constraints (4), also discussed later, require that each quality measure at stage $i$ lies between its lower and upper bounds. Constraints (3) and (4) are the ones that most complicate the model, as they depend on the configuration $\mathbf{x}$ of the plant. Constraint (5) requires that the total outflows of A and B starch must be sufficient for the refinery and modified starch processes. Constraint (6) ensures that the proportion of A starch cannot be too large as A starch is less refined than B starch. Constraint (7) is a definitional constraint. Constraint (8) establishes minimum and maximum values on a controller at a process stage. Two sets of linking constraints, (9) and (10), ensure that flows between two stages occur only when plant configuration $\mathbf{x}$ allows them, and that they do not exceed their maximum flow capacities. Finally, we need the nonnegativity and integrality conditions in (11) and (12). The problem of determining the optimal operation and configuration for the process can then be formulated by the following nonlinear integer program $P$:

\begin{align*}
\mathbf{P}: \quad Z(V) = \max_{\mathbf{x}, \mathbf{b}, w_i, \alpha_{ij}, \mathbf{F}_i, x} \left\{ \sum_{x \in \{A, B, G\}} \pi_x \rho_x d_x F_x \right\} \\
\quad - \Psi_E(E(\mathbf{b})) - \Psi_w(W) \\
\quad \sum_{x \in \{A, B, G\}} \rho_x d_x F_x \leq V \tag{1} \\
\quad F_x \leq C_x(\mathbf{b}_x, \mathbf{x}) + w_x + \sum_{i \in I_x} (\alpha_{ix} - \alpha_{xi}) \quad \forall x \in \{A, B, G\} \tag{2} \\
\quad \bar{Q}_k \leq Q_k(\mathbf{b}_k, w_x, \alpha_i; \mathbf{x}) \leq \bar{\bar{Q}}_k \quad \forall i \in I, \forall k \in K_i \tag{3} \\
\quad F_A + F_B \geq \theta_M F_R + F_M^0 \tag{4} \\
\quad F_A R + F_A M = F_A \tag{5} \\
\quad \beta_{ij} \leq \beta_{ij} \leq \bar{\beta}_{ij} \quad \forall i \in I, \forall j \in J_i \tag{6} \\
\quad w_i \leq y_i \bar{w}_i \quad \forall i \in I \tag{7} \\
\quad \alpha_{ij} \leq x_{ij} \bar{\alpha}_{ij} \quad \forall i, i_2 \in I \tag{8} \\
\quad F_A, F_B, F_G, w_i, \alpha_i \geq 0 \quad \forall i, i_2 \in I, \forall j \in J_i \tag{9} \\
\quad y_i, x_{ij}, \bar{w}_i \in [0, 1] \quad \forall i, i_2 \in I. \tag{10}
\end{align*}

In the iterative procedure, the flow response function $C_i(\mathbf{b}_x, \mathbf{x})$ in (3) and the quality constraints (4) are estimated for the given configuration $\mathbf{x}$. We used the posynomial functional form $\tilde{C}_i(\mathbf{b}_x, \mathbf{x}) = \prod_{j=1}^{n_j} y_{ij} \beta_{ij}^{\gamma_{ij}}$, which is often used to empirically model the response of a given stage to its associated control variables (Avriel and Wilde 1967, Zener 1971). Here, $y_{ij}$ represents a scaling constant, while $\eta_{ij}$ is a constant representing the elasticity or responsiveness of the flow response of stage $i$ to a unit change in controller $j$. In the implementation section we discuss how linear estimates were obtained for the quality functions in (4), to get $\tilde{Q}_k(\mathbf{b}_k, w_x, \alpha_i; \mathbf{x})$. Substitute $\tilde{C}_i(\mathbf{b}_x, \mathbf{x})$ and $\tilde{Q}_k(\mathbf{b}_k, w_x, \alpha_i; \mathbf{x})$ into (3) and (4) respectively to get:

\begin{align*}
F_i \leq \tilde{C}_i(\mathbf{b}_x, \mathbf{x}) + w_i + \sum_{i \in I_x} (\alpha_{ix} - \alpha_{xi}) \quad \forall i \in I \tag{3'} \\
\tilde{Q}_k \leq \tilde{Q}_k(\mathbf{b}_k, w_x, \alpha_i; \mathbf{x}) \leq \bar{\bar{Q}}_k \quad \forall i \in I, \forall k \in K_i. \tag{4'}
\end{align*}
Let $\hat{P}$ denote the original problem $P$ with $(3')$ and $(4')$ substituted for $(3)$ and $(4)$. Let $P(X)$ and $\hat{P}(X)$ represent problems $P$ and $\hat{P}$ respectively for a fixed configuration $X$. The following lemmas will be used in our iterative procedure and subsequent analysis.

**Lemma 1.** $\hat{P}(X)$ is a convex optimization problem if $0 < \eta_{ij} \leq 1, \gamma_{ij}, \beta_{ij} > 0, \forall i, j$.

**Proof.** To show $\hat{P}(X)$ is a convex optimization problem, we need to verify that the Hessian matrix corresponding to this problem is negative definite (Luenberger 1984). The result follows by taking the Hessian matrix and observing that it is negative definite if $0 < \eta_{ij} \leq 1, \gamma_{ij}, \beta_{ij} > 0, \forall i, j$. □

Since $\hat{P}(X)$ is a convex optimization problem, we can use standard techniques to solve the problem. The following lemma is useful for calculating the dual variables associated with the capacity constraints $(9)$ and $(10)$, which will be used in the capacity reallocation procedure in §5.

**Lemma 2.** Let $(w_i, \alpha_{i,1,2})^*$ represent the flows in the optimal solution to $\hat{P}(X)$. Then there exist vectors $\lambda^u, \lambda^a \geq 0$ such that

$$w_i^* - \bar{w}_i \leq 0 \quad \text{and} \quad \lambda_i^u (w_i^* - \bar{w}_i) = 0 \quad \forall i \in I$$

$$\alpha_{i,1,2}^* - \bar{\alpha}_{i,1,2} \leq 0 \quad \text{and} \quad \lambda_i^a (\alpha_{i,1,2}^* - \bar{\alpha}_{i,1,2}) = 0 \quad \forall i_1, i_2 \in I.$$  

**Proof.** These conditions follow directly by applying the Karush-Kuhn-Tucker conditions to $\hat{P}(X)$, which are necessary and sufficient for any convex optimization problem. □

The following lemma is useful for determining the optimal wheat input into the process.

**Lemma 3.** Let $\hat{Z}(V, X)$ represent the value of $\hat{P}(X)$ for given $V$. Then $\hat{Z}(V, X)$ is a concave function of $V$.

**Proof.** Define $z_i = \arg \max f(z)$ s.t. $A(z|X, V) \leq 0$, $\lambda^u \geq 0$ where $\lambda^a = 0$ is a constraint set which is an explicit function of $X$ and $V$. $z$ is a vector of continuous decision variables (controls and flows), and $f(z)$ is the objective function, an explicit function of $z$ alone. Define $z_i = \lambda_i z_1 + (1 - \lambda)z_2$ for any $\lambda \in [0, 1]$. Let the constraint involving $V$ be represented as $Bz \leq V$ and define $V_i = \lambda V_1 + (1 - \lambda)V_2$. By linearity of $Bz$, we have $Bz_i = B(\lambda z_1 + (1 - \lambda)z_2) \leq \lambda V_1 + (1 - \lambda)V_2 = V_i$. By convexity of the feasible region defined by the rest of the constraints $(3')$, $(4')$, $(5')$–$(10')$, feasibility of $z_i$ and $z$ also implies that $z_i$ is feasible for the other constraints; thus, $A(z|X, V) \leq 0$. In addition, the concavity of $f(z)$ implies that $f(z_{i,1}) \geq \lambda f(z_{i,1}) + (1 - \lambda)f(z_{i,2}) \forall \lambda \in [0, 1]$. Let $z_{i,1}^\ast = \arg \max f(z)$ s.t. $A(z|X, V) \leq 0$, $\lambda^u \geq 0$. Since $z_i$ is also feasible for this constraint set, $\hat{Z}(V_i, X) = f(z_{i,1}^\ast) \geq f(z_{i,1}) \geq \lambda f(z_{i,1}) + (1 - \lambda)f(z_{i,2}) = \lambda \hat{Z}(V_1, X) + (1 - \lambda)\hat{Z}(V_2, X)$. This establishes the concavity of $\hat{Z}(V, X)$ with respect to $V$. □

### 6. Procedure for Model-Based Process Redesign

In general it is impossible to solve $P$, as the exact form of the flow balance constraints $(3)$ and the quality constraints $(4)$ depends on the integer variables that jointly define the plant configuration $X$. Therefore, we developed an iterative decomposition approach based on successive convex approximations of $P$, in which Constraints $(3)$ and $(4)$ are estimated and optimal flows and controller settings are determined for a sequence of plant configurations $X_t$, where $t$ denotes the current iteration. The information from each solution is used to determine the next configuration $X_{t+1}$, using a greedy reallocation heuristic. Figure 1 summarizes the procedure. Decomposition is a natural way to solve complex problems, where one fixes the integer variables, solves the relatively easy continuous problem, and updates the integer variables. (This is the essence of Benders’ decomposition; see, e.g., Schrijver 1986, pp. 371–373.)

However, several factors conspire to make the current problem fundamentally harder:

- The original problem $P$ cannot even be formulated explicitly, as no expressions exist that capture the performance of a stage for any controller settings and any configuration.
- Even for a given configuration $X_t$, no theoretical expressions exist for the general flow balance and quality constraints $(3)$ and $(4)$. One has to estimate them empirically from the process to get the approximate problem $\hat{P}(X_t)$, which must be updated each time $X_t$ is updated.
- Worse still, the only way to obtain $(3')$ and $(4')$ for a proposed new configuration $X_{t+1}$ is to physically implement the changes required (i.e., removing and adding flow capacity throughout the plant) and then running experiments with $X_{t+1}$. This is obviously very costly and limits us to a procedure requiring very few iterations.

A further constraint was management’s desire to avoid new capital investment. This meant that to expand capacity of a link (by adding pipes or pumps, or both), this capacity had to be freed up elsewhere in the plant. So, in moving from $X_t$ to $X_{t+1}$, capacity could only be added in discrete increments of exactly those sizes that were no longer needed under $X_t$. This constraint was incorporated but is not necessary for the procedure. The approach chosen can be formalized by the following five-step procedure, summarized in Figure 3:

**Step 0.** Initialization: Start with given plant configuration $X_0$; use several months’ operating procedure to estimate Constraints $(3')$ and $(4')$.

**Step 1.** For given $X_t$ and estimated Constraints $(3')$ and $(4')$, and in light of Lemma 1, we can solve $\hat{P}(X_t)$ using the standard convex programming techniques implemented in commercial software such as gradient, sequential-unconstrained, or sequential-approximation algorithms (Hillier and Lieberman 1989).

**Step 2.** For any flow $a_{i,1,2}$ or $w_{i,t}$ that is zero in the optimal solution to $\hat{P}(X_t)$, remove the link by setting the
Figure 3. Implementation of methodology for model-based process redesign at wheat extraction process.

Step 0: initialization: determine key process controllers, estimate process model with initial configuration \( \hat{P}(X_0) \)

Step 1: find optimal flow for given configuration, i.e., solve \( \hat{P}(X_f) \)

Step 2: physically remove links with zero flows; add corresponding pipes and pumps to set of available capacity increments

Step 3a: find dual variables associated with each flow

Step 3b: add capacity increments in available set to each flow constraint, compute resulting increase in objective value

Step 3c: allocate capacity increment to constraint with highest increase in objective value; remove capacity increment from available set

Are more capacity increments available or further capacity reallocation will outweigh costs of change

Yes

Step 4: physically implement changes from Step 3 during next plant shutdown, i.e., change configuration from \( X_t \) to \( X_{t+1} \)

No

Step 5: operate process under new configuration, collecting data from operations to update process model, i.e., estimate \( \hat{P}(X_{t+1}) \)

corresponding \( x_{i1}, t+1 \) or \( y_{i1}, t+1 \) equal to zero in the next configuration \( X_{t+1} \) and add \( \hat{t}_{i1}, t+1 \) or \( \hat{w}_{i1} \) to the set \( C_i \) of available capacity increments. If flows are too small to be technically realizable (because it is not possible to run the equipment at such low flow levels), then these flows are also set to zero and the corresponding links removed.

Step 3. Reallocate the capacity available from Step 2 using the following greedy reallocation procedure.

(a) Order \( C_i \) such that \( c_i \) is its \( i \)th element. Compute dual variables corresponding to Equations (9) and (10) using the KKT conditions in Lemma 2 for all variables for which \( \alpha_{i1}, i, t \neq 0 \) or \( w_{i1}, i, t \neq 0 \). The dual variables helps us identify the active or binding constraints.

(b) For all \( c_i \in C_i \), add \( c_i \) to each binding constraint in (9) and (10), one at a time, and compute the resulting objective value \( \hat{Z}(V, X_{t+1}) \).

(c) Identify the constraint for which \( \hat{Z}(V, X_{t+1}) \) increases the most when adding capacity \( c_i \). Allocate capacity increase \( c_i \) to the constraint identified by adding \( c_i \) to the corresponding \( \hat{t}_{i1}, i, t \) or \( \hat{w}_{i1} \), and remove \( c_i \) from \( C_i \).
Repeat Steps (a)–(c) for all elements of \( \mathbf{C} \), until \( \mathbf{C} \) is empty or no further marginal increase in \( \tilde{Z}(V, \mathbf{X}_{t+1}) \) exceeds the reallocation cost.

Step 4. Physically implement the changes proposed in Steps 2 and 3 during the next plant shutdown.

Step 5. With the new configuration \( \mathbf{X}_{t+1} \) and the new flow capacities, operate the process for the next five months and keep updating the flow balance constraints (3') and quality constraints (4'). At the next plant shutdown, go back to Step 1.

This procedure has several attractive features from a practical standpoint: At every iteration it yields feasible solutions, and, if desired, it ensures that all configuration changes rely on reallocating existing capacity rather than buying new capacity, so there are no incremental costs of new capacity.

### 7. IMPLEMENTATION

We implemented the ideas presented here over a five-year period at Cerestar. The initial step was to determine the set of controllers to be regulated automatically, rather than by operators. It is important to reduce the number of controllers set by operators as this minimizes the complexity of the operation and reduces the variance of the output (Rajaram and Jaikumar 2000, 2002). To do this, we observed how inflow varied with the controller settings chosen by the process operators in the course of normal operations. We used a loglinear regression to fit the parameters of the posynomial flow response function \( C(\mathbf{B}, \mathbf{X}) = \prod_{j=1}^{n} \gamma_j \beta_j^\eta_j \). For parameter estimates that were not significant at the 5% level, we set the corresponding controllers to a fixed value derived from engineering specifications and removed them from further consideration. We also checked for collinearity using variance inflation factors (VIF), following the criteria discussed in Hair et al. (1998, pp. 191–193). Highly collinear controllers (i.e., with VIF > 10) were automatically regulated (again based on engineering specifications). As this step required significant experimentation and one-time physical changes in the plant, it took around one year to determine the minimum set of controllers. Depending on the scale used for each controller, \( \beta_j \), ranged from 1 to 10. We also found that \( \gamma_j \geq 0 \) and \( 0.4 \leq \eta_j \leq 0.9 \) for all controllers at the six stages at all iterations; this is consistent with the design specifications, as \( \eta_j \geq 1 \) for any controller implies instability in the associated control loop at that stage of the process. By Lemma 1, this implies that \( \hat{\mathbf{P}}(\mathbf{X}) \) was a convex optimization problem in this application.

Once the controllers had been chosen, we spent the first five months before the semiannual maintenance shutdown collecting data on the relationship between flows, controllers, and quality. We sampled data at each stage of the process for the relevant control variables and water flows. Because the time between changes in controller levels and the response time of the appropriate stage for a given variable was at least 10 minutes, we set the sampling frequency to 10 minutes. This yielded 6 observations per hour for each variable, so after five months, or 460 shifts, we had about 22,080 observations. Data collection was automated using software linked to the process control system. We did not conduct any formal experimental design, but used data from ongoing operations—the best we could do given the importance of running the process at full capacity at all times and the fact that no operator time was available to conduct formal experiments. Such “natural experiments” are often the only practical way of obtaining data and have also been used in the context of semiconductor manufacturing (Bohn 1995).

We used multivariate linear regression to estimate the quality constraints and loglinear regression for the flow response functions. Here again, we only included variables that were significant at the 5% level and had VIF < 10. This procedure reduced the total number of controllers in this process from 90 to 18. This in turn helped operators understand the major causal relationships between the controllers and the quality measures at a stage and facilitated the development of process knowledge.

In the remaining month before the semiannual shutdown, we used the empirical relationships between the flows, controllers, and quality measures to solve problem \( \hat{\mathbf{P}}(\mathbf{X}_0) \). Since all stages were initially connected to fresh water and had recycling interlinks with each other, we set all the binary variables defining the initial configuration \( \mathbf{X}_0 \) equal to 1. The resulting continuous nonlinear program \( \hat{\mathbf{P}}(\mathbf{X}_0) \) had over 3,000 variables and 1,538 constraints, excluding nonnegativity and integrality constraints. The majority of these constraints were 900 linking constraints (9) and (10), 450 quality constraints (4), and 180 constraints (8) on controller settings. We solved \( \hat{\mathbf{P}}(\mathbf{X}_0) \) using the DICOPT solver in GAMS (Brooke et al. 1992), and used the solution to decide which fresh water connections and recycling interlinks to close. We determined the dual variables by solving the KKT conditions (from Lemma 2) using Matlab (Math Works Inc. 1998). The set of positive dual variables associated with the fresh water and recycling flow capacity constraints (9) and (10) were used in the capacity reallocation algorithm, outlined in Step 3 of the approximation procedure in §4.

The adjusted \( R^2 \) values for the regressions for the flow response and quality constraints were all over 90%. In light of this good fit, we were confident about the configuration changes proposed by the model. However, if \( R^2 \) is low for any particular constraint, we recommend performing sensitivity analysis on the coefficients in this constraint and proposing configuration changes only if they are shown to be robust across a range of coefficient values.

During the maintenance shutdown at the end of the sixth month, the configuration changes were physically implemented. The capacity increases were achieved by using pumps and pipes available from the fresh water and recycle interlinks that were removed elsewhere. We developed a short-term operational Decision Support System (DSS) to determine controller settings and fresh water and recycle
flows at each shift. After each shift, the estimates of the flow balance and quality constraints (3’ and 4’) were updated using the most recent five months of data; data from the latest shift were added and those from the first remaining shift, five months ago, were dropped. That way, (3’) and (4’) were increasingly representative of the new configuration. The coefficient estimates did not change drastically from one iteration to the next due to the incremental nature of the configuration changes (though they did change significantly across multiple iterations), so the bias introduced by using data from before and after the most recent configuration change is likely to be minimal. The DSS then solved \( P(X_t) \) with the updated estimates, and fed the solution as inputs to the process control system. After five months, at three shifts per day and 30 days per month, all 22,080 observations were from the new configuration, so the estimates of (3’) and (4’) were as accurate as possible. At this point, we repeated the capacity reallocation procedure and physically implemented the improved configuration \( X_{t+1} \) during the next shutdown. This procedure was repeated eight times over four years until the gains from reallocating capacity were outweighed by the costs. The DSS is still used at Cerestar for operational process control.

8. RESULTS AND DISCUSSION

To better appreciate the impact of this project, compare the upper part of Figure 4, which represents the product, fresh water, and recycle flows before implementation, with the lower part, the process after five years. This transition reduced the number of control variables from 90 to 18 and the number of links by over 75%. The plant initially had some 1,000 pipes; this number was cut by around 100 in each iteration. The final fresh water consumption was around 50% less than before, while energy consumption was down 30%; Table 1 shows the percentage reduction after each successive iteration of our procedure. Together this represents over $3 million annual savings, while the total one-time costs (including fixing controllers, removing piping, reallocating capacity, and engineering time) were around $1 million. Even more significantly, the reduction in fresh water consumption enabled Cerestar to maintain the current production level without investing in the $100 million wastewater treatment capacity expansion.

In retrospect, the final configuration in Figure 4 makes good sense. Obviously, it is far simpler and hence easier to control than the initial configuration. To avoid variability propagating through the entire process, it is more important to minimize variability added to upstream stages than to downstream stages. Recycled water adds more variability as it is less pure than fresh water, but all flows (including recycled water) from a downstream stage are more pure than those from an upstream stage. Consequently, the upstream (dough) stage is fed by recycling links from stages far downstream, while the intermediate (dilution) stage is fed from a stage less far downstream (separation). In addition, there are no recycled flows from the upstream stages to the downstream stages. The final stages are critical in determining outgoing product quality, so they use fresh water. This procedure also takes into account the different degrees of variability associated with recycling water from different stages in the quality constraints. To see this, let \( i_1 \) represent an upstream stage and \( i_2 \) a downstream stage. As discussed, recycled flows from stage \( i_1 \) have greater impurities than from stage \( i_2 \). We found that the regression coefficient associated with a downstream-to-upstream recycled flow \( \alpha_{i_2i_1} \) is typically smaller than that associated with the reverse flow \( \alpha_{i_1i_2} \). Thus, an impure recycled flow from downstream to upstream is less likely to violate a quality constraint than the same level of impure recycled flow from upstream to downstream, which in turn suggests reducing these latter flows whenever possible.

This project resulted in several strategic benefits. First, we used the model to determine the optimal flour input to the process, an important variable in supplier selection, in negotiating the financial contract for the raw material wheat, in determining the required unloading capacity, and in estimating the optimal silo storage volume and the optimal capacity and production rate of the grinding mill. For instance, the optimal flour input was used to determine the grind capacity and the buffer required, given existing barge arrival patterns (which were difficult to change). To determine the optimal flour input \( V^* \), we considered problem \( \hat{P}(X_t) \) with the configuration fixed after the final iteration \((t = 8)\), and used a golden section method (Luenberger 1984) and Lemma 3. Here, the golden section method determines a locally optimal \( V^* = \arg \max_{V} \hat{Z}(V, X_t) \), which by Lemma 3 is also guaranteed to be a global optimum.

Second, we quantified the impact of downstream variability on the wheat starch extraction process. We first set the refinery flow input \( F^0_R \) equal to a fixed value \( \mu_R \) over a period of time and solved \( \hat{P}(X_t) \) in the final configuration to compute optimal profits. We then generated random \( F_R \) with mean \( \mu_R \) and variance \( \sigma^2_R \) and solved \( \hat{P}(X_t) \) for different values of \( \sigma^2_R \). The results, shown in Figure 5, suggest that even small increases in variability result in a significant decrease in profits compared to the current case with a coefficient of variation of 20%.

Second, we quantified the impact of downstream variability on the wheat starch extraction process. We first set the refinery flow input \( F^0_R \) equal to a fixed value \( \mu_R \) over a period of time and solved \( \hat{P}(X_t) \) in the final configuration to compute optimal profits. We then generated random \( F_R \) with mean \( \mu_R \) and variance \( \sigma^2_R \) and solved \( \hat{P}(X_t) \) for different values of \( \sigma^2_R \). The results, shown in Figure 5, suggest that even small increases in variability result in a significant decrease in profits compared to the current case with a coefficient of variation of 20%.

To understand why, it is important to note that the relative profit margins satisfy \( \pi_A > \pi_G > \pi_B \). Increased refinery production requires higher refinery flow input than normal from the wheat starch extraction process, so that \( F^0_R > \mu_R \), creating upside variability. To meet such upside variability, additional B starch needs to be produced by Capacity Constraint (5) and (6) at the expense of the more profitable gluten product. Conversely, decreased refinery production requires lower refinery flow input than normal, so that \( F^0_R < \mu_R \), creating downside variability. Such downside variability causes less A starch to be produced by Capacity Constraint (6) and hence a reduction in profits. In addition, because the controller values are set for a stable range of flows, changes in...
Figure 4. Wheat and product flows before and after simplification.

this range due to downstream flow variability force deviation from these values, increasing energy and fresh water consumption. This analysis provides additional justification for the initiatives known as “robust process control” developed by Rajaram et al. (1999), which reduced the coefficient of variation at the downstream refinery from 50% to 20%.

Third, as a result of the savings in energy and water consumption, the Dutch government presented Cerestar with an environmental management award and a tax subsidy valued at $10 million to set up a pilot plant for another type of product at Sas-van-Gent. Finally, the ideas gained from this project have been used to design a larger wheat starch extraction process currently under construction at Sas-van-Gent. For instance, the controllers and recycle flows which were effective in minimizing fresh water consumption, energy consumption, and meeting product quality constraints were used in the design of the new process. After startup, Cerestar has committed to use this framework to further optimize profits at the new process.

9. SUMMARY AND CONCLUSIONS
We have described a model-based iterative procedure for process redesign. Although we describe in detail how it was implemented at a major wheat starch extraction facility in
Europe, the iterative approach can be applied to any large-scale continuous or semicontinuous process in a variety of industries. The key contribution of this approach lies in formulating a sequence of tractable convex approximations of the optimization problem, which in itself cannot be formulated explicitly for several reasons.

This work has led to several important lessons. It illustrates clearly how tightening environmental constraints need not be bad: In this instance, they triggered substantial process improvement, using a pollution prevention approach rather than an end-of-pipe solution. This led to major cost reductions and improvements in environmental performance. It is important to recognize that process redesign and improvement was achieved through simplification. Although at first glance it may seem paradoxical that reducing the number of links between stages in a process could lead to improved performance, it is not uncommon to find process simplicity and process performance going hand-in-hand.

Figure 5. Impact of downstream variability at refinery on profits.

The final configuration of the process resulting from the redesign efforts described here does have intuitive appeal and may be effective in other settings, although we cannot be certain that a similar configuration would also work elsewhere without following the iterative procedure described here. An additional attractive feature of the procedure as implemented here is that it relies on reallocation of existing capacity rather than new capital investment. In fact, it avoided a major capital investment that would have been otherwise required to maintain production at current levels.

Clearly, this work opens up significant opportunities for further theoretical and practical work. The current iterative procedure needs to be tested in a wider range of settings, and further improvements can undoubtedly be made in the procedure itself. For instance, the procedure could be modified to include the value of experimentation in which configuration changes take place not because they appear to represent the most profitable immediate moves, but because they might provide the most information about particularly uncertain aspects of the process. We hope that this paper will motivate others to join this exciting line of research.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the intellectual contributions to some of the ideas presented here and the administrative support provided at various stages of the project by the following people at Cerestar: David Challenge (International Manufacturing Director), Franz Behlau (Manufacturing Director, Cerestar Benelux), Frans Van Esch (Process Development Manager, Cerestar Benelux), and Stefan Vrijland (Process Development Engineer, Cerestar Benelux). They are also grateful for financial support from the AT&T Foundation for Industrial Ecology, the Center for International Business Education and Research (CIBER) at UCLA, and the University of California Pacific Rim Research Program. This paper has also benefited from the excellent suggestions made by Benjamin Hobbs and the referees.

REFERENCES


Table 1. Percentage reduction of total fresh water and energy consumed after each successive iteration of procedure.

<table>
<thead>
<tr>
<th>Iteration Number</th>
<th>Total Fresh Water Consumed; Initial Level = 5000 m$^3$/day</th>
<th>% Reduction in Total Fresh Water Consumed</th>
<th>Total Energy Consumed; Initial Level = 168 MWH/day</th>
<th>% Reduction in Total Energy Consumed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8.0</td>
<td>4.0</td>
<td>31 3</td>
<td>5.4</td>
</tr>
<tr>
<td>2</td>
<td>9.8</td>
<td>6.3</td>
<td>11.1</td>
<td>4.4</td>
</tr>
<tr>
<td>3</td>
<td>13.3</td>
<td>4.4</td>
<td>6.3</td>
<td>4.8</td>
</tr>
<tr>
<td>4</td>
<td>11.1</td>
<td>3.5</td>
<td>6.7</td>
<td>5.1</td>
</tr>
<tr>
<td>5</td>
<td>6.7</td>
<td>4.8</td>
<td>6.7</td>
<td>5.1</td>
</tr>
<tr>
<td>6</td>
<td>7.1</td>
<td>1.3</td>
<td>7.1</td>
<td>1.3</td>
</tr>
<tr>
<td>7</td>
<td>3.8</td>
<td>5.4</td>
<td>3.8</td>
<td>5.4</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>