An Efficient Algorithmic Framework for Environmental Modeling[*]

U.Diwekar [a]

[a]Vishwamitra Research Institute,
Center for Uncertain Systems: Tools for Optimization and Management,
34, N. Cass Avenue, Westmont, IL 60559

Abstract: Environmental systems often involve nonlinear models, continuous as well discrete decisions, and multiple objectives. Uncertainties are inherent in these models increasing the complexity of decision making further. This paper presents an efficient algorithmic framework for environmental decision making for large scale systems involving multiple objectives in the face of uncertainties.

Keywords: multi-objective, combinatorial, uncertainty, optimization, environmental modeling

1 INTRODUCTION

Simulation models and other design tools allow engineers to design, simulate, and optimize systems. However, there is a critical need to incorporate green engineering into the design of these systems. This calls for extending the breadth of the design process. However, this integration poses challenging problem of discrete and continuous decisions, nonlinear models, and uncertainties. Further, the goals in terms of profitability are relatively easy to define, and researchers in academics and industries have used simulators and modeling tools to achieve profitability where environmental considerations are considered as definable constraints. However, "complete ecological considerations" to be included as environmental impact objectives is a formidable task. Thus, efficient multi-objective optimization methods are necessary to handle the conflicting and different objectives involved in the problem of greener by design. Multi-objective optimization approach is particularly valuable in the context of pollution prevention, waste management,life cycle analysis, and sustainability as there are a large number of desirable and important objectives which are not easily translated into dollars. Extending the envelope from simulation of a plant to management and planning towards sustainability, and broadening the scope to include multiple objectives other than profitability, increase uncertainties.

Further, the decision making then involves discrete decisions related to selection of alternatives, as well as continuous decisions that defines the operations and design parameters (Diwekar [2003a]). Thus, at the crux of this decision making are efficient algorithms, methods, and tools for multi-objective optimization and uncertainty analysis. This algorithmic framework is described in this paper. Two real world case studies illustrate the promise of such a framework.

2 ALGORITHMIC FRAMEWORK

The algorithmic framework consists of five calculation levels as shown in the Figure 1. The basis of this framework is numerical optimization algorithms (Diwekar [2003b]) for selecting discrete and continuous decisions in the face of multiple objectives, and probabilistic uncertainty analysis to account for uncertainties and variabilities in objectives, constraints, and parameters.

Level 1, is the inner most level and corresponds to models for simulation. This level defines all possible alternatives. Currently, for the case study presented in this paper, ASPEN Plus^{TM} (AspenTech [2004]) is used for process modeling.

Level 2, Sampling loop: The diverse nature of uncertainty, such as estimation errors and process variations, can be specified in terms of probability distributions. The type of distribution chosen for an uncertain variable reflects the amount of informa-
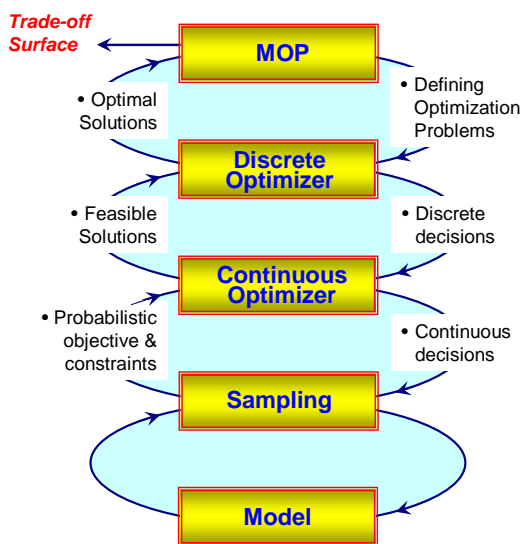
Figure 1: The Algorithmic Framework

tion that is available. For example, the uniform and log-uniform distributions represent an equal likelihood of a value lying anywhere within a specified range, on either a linear or logarithmic scale, respectively. Further, a normal (Gaussian) distribution reflects a symmetric but varying probability of a parameter value being above or below the mean value. In contrast, lognormal and some triangular distributions are skewed such that there is a higher probability of values lying on one side of the median than the other. A beta distribution provides a wide range of shapes and is a very flexible means of representing variability over a fixed range. Modified forms of these distributions, uniform* and log-uniform*, allow several intervals of the range to be distinguished. Finally, in some special cases, user-specified distributions can be used to represent any arbitrary characterization of uncertainty, including chance distribution (i.e., fixed probabilities of discrete values).

It is easier to guess the upper and lower bounds of uncertain variables and hence uniform distribution provides the first step towards uncertainty quantification. If one can identify the most likely value then traingular distributions can be used (Diwekar et al. [1997]). Recently, Kim and Diwekar [2002a] used extensive data obtained from the DECHEMA databank and obtained realistic quantification of uncertainties related to UNIFAC parameters for the chemical synthesis problem reported in the case study

section.

Once probability distributions are assigned to the uncertain parameters, the next step is to perform the sampling operation. One of the most widely used techniques for sampling from a probability distribution is the Monte Carlo sampling (MCS) technique, which is based on a pseudo-random generator to approximate a uniform distribution (i.e., having equal probability in the range from 0 to 1). The specific values for each input variable are selected by inverse transformation over the cumulative probability distribution. A Monte Carlo sampling technique also has the important property that the successive points in the sample are independent. Nevertheless, in most applications, the actual relationship between successive points in a sample has no physical significance; hence, the randomness/independence for approximating a uniform distribution is not critical. In such cases, uniformity properties plays a more critical role in sampling, as a result, constrained or stratified sampling techniques are more appealing. Latin hypercube sampling (LHS) (McKay et al. [1979]) is one form of stratified sampling that can yield more precise estimates of the distribution function. The main drawback of LHS stratification scheme is that, it is uniform in one dimension and does not provide uniformity properties in k-dimensions. Sampling based on cubature techniques or collocation techniques face similar drawback. These sampling techniques perform better for lower dimensional uncertainties. Therefore, many of these sampling techniques use correlations to transform the integral into one or two dimensions.

Efficient sampling techniques (Hammersley sequence sampling, HSS and Latin Hypercube Hammersley Sampling, LHHS) based on Hammersley points were developed by my group (Kalagnanam and Diwekar [1997],Wang et al. [2004]). HSS uses an optimal design scheme for placing the n points on a k-dimensional hypercube. This scheme ensures that the sample set is more representative of the population, showing uniformity properties of random variables in multi-dimensions, unlike Monte Carlo, LHS, and its variant, the Median Latin hypercube sampling technique. Figure 2 shows samples generated for two uniform uncertain (random variables) using MCS and HSS. It has been found that the HSS technique is at least 3 to 100 times faster than LHS and Monte Carlo techniques and hence is a preferred technique for uncertainty analysis, as well as optimization under uncertainty, and is used in this framework.

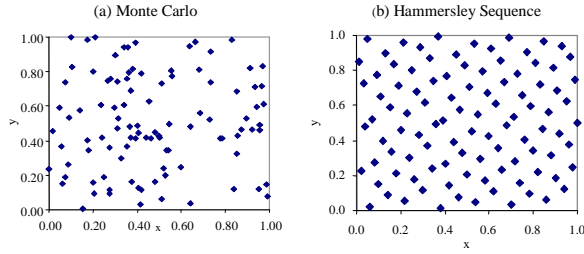**Level 3**, Continuous optimizer: This step involves

Figure 2: Samples of MCS and HSS for two uniformly distributed random variables

Figure 3: Stepwise improvement in genetic algorithms.

continuous decisions like design and operating conditions for a process. Derivative based quasi-Newton methods, where the gradient (i.e., Jacobian) is approximated based on differences in the x and $f(x)$ obtained from previous iterations, are widely used in process optimization. Among the quasi-Newton based methods, the SQP (successive quadratic programming) method is used for this framework because it requires far fewer function and gradient evaluations than other methods for highly nonlinear constrained optimization, and it does not need feasible points at intermediate iterations. Both of these properties make SQP one of the most promising techniques for problems dealing with nonlinear constraint optimization, like process simulations.

Level 4, Discrete optimizer: This involves dealing with discrete decisions such as chemical and process structural alternatives. This is the most difficult optimization step. New algorithms are designed by improving efficiency in two steps: (1) improving the discrete optimization algorithm by using quasi-random numbers like the Hammersley Sequence Sampling, and (2) by providing efficient interaction between the discrete optimization and the sampling technique for efficient stochastic optimization. These new algorithms are found to be 99 % more efficient than the traditional algorithms. For details of these algorithms, please see Kim and Diwekar [2002a]; Xu and Diwekar [2005]. Figure 2 shows the stepwise improvement in genetic algorithm, namely, (1) ESGA, and (2) HSGA for a simple problem.

Level 5, Multi-Objective Programming, (MOP):This represents the outermost loop in Figure 1.

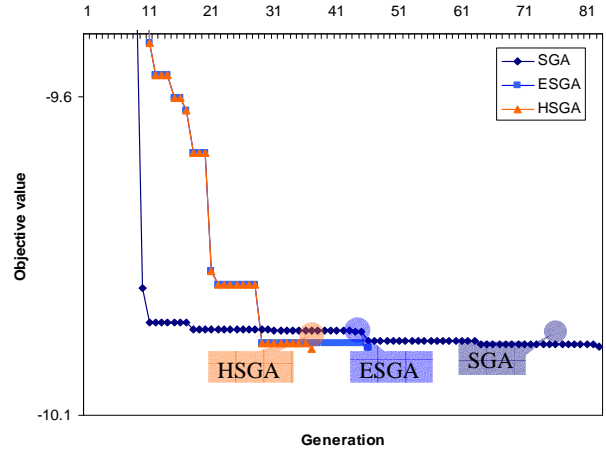A generalized Multi-objective optimization (or Multi-objective Programming) problem can be for-

mulated as follows:

$$
\begin{aligned}
\min \quad & Z = Z_i, \; i = 1, \ldots, p, \quad p \geq 2, \quad (1) \\
\text{s.t.} \quad & h(x,y) = 0, \\
& g(x,y) \leq 0,
\end{aligned}
$$

where **x** and **y** are continuous and discrete decision variables, and p is the number of objective functions. The functions $h(x,y)$ and $g(x,y)$ represent equality and inequality constraints, respectively. There are a large array of analytical techniques to solve this MOP problem; however, the MOP methods are generally divided into two basic types: preference-based methods and generating methods. Preference-based methods like goal programming attempt to quantify the decision-maker's preference, and with this information, the solution that best satisfies the decision-makers's preference is then identified (Diwekar [2003b]). As is well known, mathematics cannot isolate a unique optimum when there are multiple competing objectives. Mathematics can at most aid designers to eliminate design alternatives dominated by others, leaving a number of alternatives in what is called the Pareto set. Generating methods, such as the weighting method and the constraint method, have been developed to find the exact Pareto set or an approximation of it. For each of these designs, it is impossible to improve one objective without sacrificing the value of another relative to some other design alternative in the set. From among the dominating solutions, it is then a value judgment by the customer to select which design is the most appropriate. At
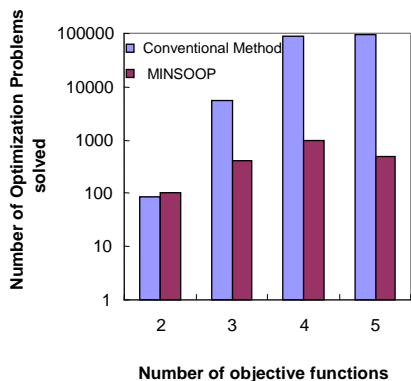
Figure 4: Computational speed-up through MIN-SOOP

issue is an effective means to find the members of the Pareto set for a design problem, especially when there are more than two or three objectives; the analysis per design requires significant computations to complete, and there are an almost uncountable number of design alternatives. A pure algorithmic approach to solving is to select one to minimize while the remaining objectives are turned into an inequality constraint with a parametric right-hand-side, L_k. The problem takes on the following form:

$$\begin{aligned} \min \quad & Z_j, \qquad\qquad\qquad (2)\\ \text{s.t.} \quad & h(x,y) = 0,\\ & g(x,y) \le 0,\\ & Z_k \le L_k, \; k = 1, \dots, j-1, j+1, \dots, p, \end{aligned}$$

where Z_j is the chosen j-th objective that is to be optimized. Solving repeatedly for different values of L_k chosen between the upper, $Z_L(j)$ and lower bounds $Z_U(j)$ leads to the Pareto set. The multi-objective optimization algorithm used in this work uses the Hammersly sequence sampling to generate combinations of the right-hand-side. The aim is to MInimize Number of Single Objective Optimization Problems (MINSOOP) by exploiting the n-dimensional uniformity of the HSS technique. Figure 4 shows how this MINSOOP (Fu and Diwekar [2004]) algorithm improves efficiency for a simple, nonlinear, convex optimization problem, as the number of objectives increases.

3   REAL WORLD CASE STUDIES

To illustrate the usefulness of this algorithmic framework, two real world case studies are pre-

sented here. The first case study is from a chemical industry and involves all the calculation loops. The second case study deals with fuel cell based hybrid power system. This case study illustrate, how uncertainties can change the trade-offs and also shows that inclusion of uncertainty analysis in the optimization framework results in solutions that are closer to reality.

3.1   Environmentally Benign Solvent Selection and Solvent Recycling

Recently, Eastman Chemicals presented a difficult separation system design problem (Kim and Diwekar [2002b]). The separation system (acetic acid-water separation) consists of an extraction column followed by a distillation column, and a decanter separating the In Process Solvent (IPS, extractive agent (solvent), and water. This process seemed to be simple and easy to operate. However, in practice, had several operational difficulties. The configuration had no degrees of freedom for improving process performance and flexibility. Therefore, variations in the feed condition led to severe process instability resulting in impure products, environmentally harmful effluent, and great economic loss. We utilized the coupled solvent selection, solvent recycling process synthesis approach made possible by the new framework to provide several environmentally benign solvents and process designs that maximized acetic acid recovery, maximized process flexibility, minimized environmental impacts in terms of LC_50, LD_50, and BioConcentration Factor (BCF), and also minimized energy consumption. Figure 5 shows the Pareto optimal solutions for the three out of the six objectives. The solutions correspond to four environmentally benign solvents. The associated process designs not only show robustness in the face of feed variations, but these designs are environmentally friendly and provide high recovery(97 % as compared to 60% base recovery for 10 % feed variation, and 98 % recovery for 5% variation where base recovery was 80 %). This figure shows that the best designs are concentrated at the right hand side and top corner. However, it should be remembered that this figure only shows three out of six objectives, we need to consider trade-offs with respect to other objectives also. However, unlike the base design, these designs provide decision maker ample choice to find suitable best designs. In this case study, we used computer-aided molecular design (CAMD) for environmentally benign solvent selection. CAMD, based on the reverse use of group contribution methods, can automatically generate promising solvent molecules from their fun-
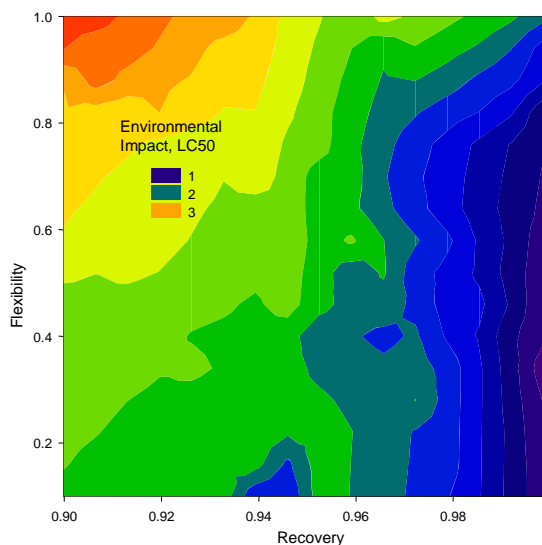
Figure 5: Pareto optimal solutions at 5% normal feed flow variations

damental building blocks or groups. The set of groups to be selected is specially designed for linear or branched hydrocarbons where aromatic, cyclic, and/or halogenated compounds are eliminated due to environmental concerns. However, this resulted in a combinatorial explosion of alternatives (e.g in our problem 6.3 X 1032 possible combinations of groups for solvent selection only). Moreover, uncertainties are inherent in these group contribution methods that needed to be characterized, quantified, and included in the framework. This is a challenging task and process synthesis added more complexity to this task. However, the new framework made this task easier.

3.2 Multi-objective Designs for a Hybrid Fuel Cell Power Plant under Uncertainty

As the first step towards a multi-objective analysis for obtaining cleaner, efficient, cost-effective and greener electricity, a fuel cell hybrid plant design case study is presented here (Subramanyan et al. [2004]). The study, which is sponsored by the National Energy Technology Laboratories (NETL), is based on a hybrid fuel cell power plant system that uses both solid oxide fuel cells (SOFC) as well as polymer electrolyte fuel cells (PEM). The aim of this case study was to illustrate the benefits of using the multi-objective optimization methods to obtain designs with minimum environmental impacts and superior performance. Figure 6a shows the Pareto
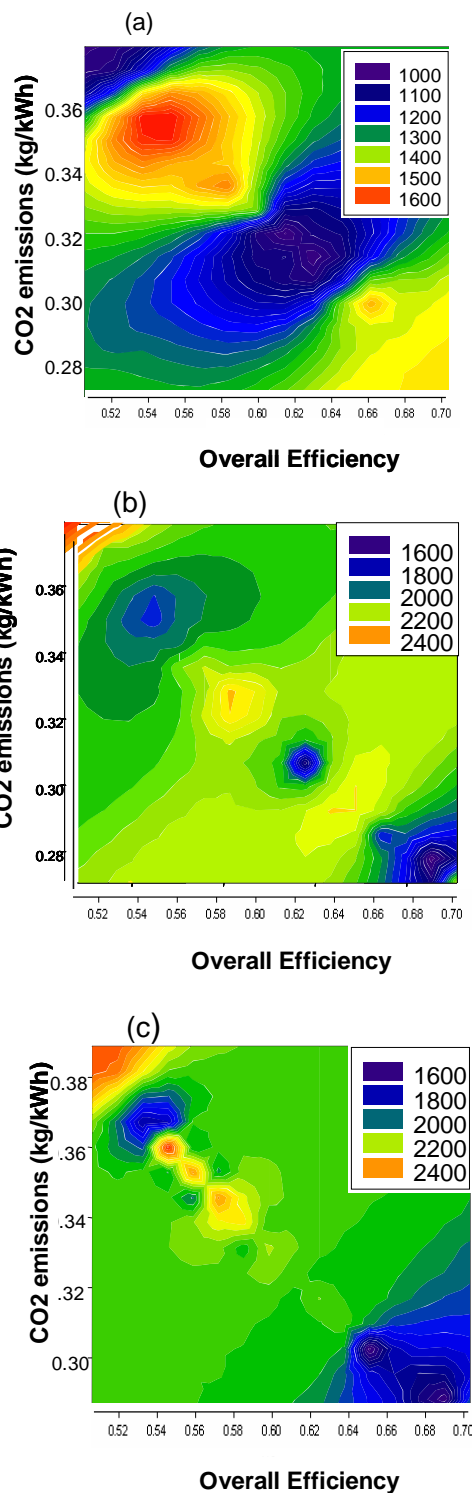

Figure 6: Pareto-optimal surface for the SOFC-PEM hybrid fuel Cell power Plant: (a) deterministic, (b) stochastic, (c) realistic

optimal solutions for minimization of cost and CO_2 emissions, and efficiency maximization. Again, these are three out of six objective functions that are considered for this work. Here the contours represent different capital costs. The question is, will these trade-offs change if models do not capture the complete physical phenomena and uncertainties are considered? Figure 6b shows such a surface for the same three objective functions when uncertainties are included in the analysis. For multi-objective optimization, these uncertainties are propagated using the HSS technique and the expected values of the objective functions are used for the analysis. It can be seen that the modeling uncertainties have considerable effect on the objectives since the trade-off surfaces are markedly different. These uncertainties can be reduced if better models are used and realistic results can be obtained. Figure 6c shows such a realistic trade-off surface when uncertainties are reduced to a minimum. After comparing the three surfaces, it is obvious that inclusion of uncertainties in the analysis (Figure 6b) resulted in trade-off surfaces that are closer to reality(Figure 6c). However, obtaining the Pareto set for a highly nonlinear system like the hybrid fuel cell power plant is computationally very expensive even with simplified models used in this exercise. Inclusion of uncertainties results in many fold increased computational intensity. Therefore, it was necessary to have the efficient methods, algorithms, and better computational power of the algorithmic framework, to address these problems.

4 SUMMARY

This paper presents an efficient algorithmic framework for environmental decision making. Environmental systems pose discrete as well as continuous optimization problems in the face of multiple objectives. Further, uncertainties are inherent in these systems making the decision making more complex. The algorithmic framework presented here involves five calculation levels: (1) modeling level, (2) stochastic analysis level where uncertainties are dealt with, (3) nonlinear optimization level for continuous decision variables, and (4) discrete optimization level, and (5) multi-objective optimization. Efficient algorithms are used at each level for solving large scale problems of significance. Two real world case studies illustrate usefulness of this framework. The first case study dealt with designing greener solvents and solvent recycling process for a chemical industry. This case study involved all calculation levels presented in the algorithmic framework. The order of magnitude increase in

computational efficiency of this framework made this case study possible. The second case study is devoted to greener energy production using fuel cell based technology. This case study illustrated the importance of including uncertainty analysis in optimization in order to obtain realistic trade-offs.

REFERENCES

AspenTech. *Aspen Plus Documentation Version 12.1-0.* AspenTech, Cambridge, MA, 2004.

Diwekar, U. Greener by design. *Environmental Science and Technology*, 37:5432–5444, 2003a.

Diwekar, U., F. C., and E. Rubin. Optimal design of advanced power systems under uncertainty. *Energy Conversion and Management Journal*, 38: 1725, 1997.

Diwekar, U. *Introduction to Applied Optimization.* Kluwer Academic Publishers, Dordrecht, 2003b.

Fu, Y. and U. Diwekar. An efficient sampling approach to multi-objective optimization. *Annals of Operations Research*, 132:109, 2004.

Kalagnanam, J. and U. Diwekar. An efficient sampling technique for off-line quality control. *Technometrics*, 39(3):308–319, 1997.

Kim, K. and U. Diwekar. Hammersley stochastic annealing: efficiency improvement for combinatorial optimization under uncertainty. *IIE Transactions*, 34:761, 2002a.

Kim, K.-J. and U. Diwekar. Integrated solvent selection and recycling for continuous processes. *Industrial and Engineering Chemistry Research*, 41:4479–4488, 2002b.

McKay, M., R. Beckman, and W. Conover. A comparison of three methods of selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21:239, 1979.

Subramanyan, K., U. Diwekar, and A. Goyal. Multi-objective optimization for hybrid fuel cells power system under uncertainty. *Journal of Power Sources*, 132:99, 2004.

Wang, R., U. Diwekar, and C. Grégorie Padró. Efficient sampling techniques for uncertainties and risk analysis. *Environmental Progress*, 23(2): 141–157, July 2004.

Xu, W. and U. Diwekar. Improved genetic algorithms for deterministic optimization and optimization under uncertainty. part i. algorithms development. *Ind. Eng. Chem. Res.*, 44:7132, 2005.