Optimization of Imaging Processes by Use of Designed Experiments and Quality Loss Functions

Allan Ames
Reading, MA
Neil Mattucci
Polaroid Corporation, Waltham, MA
Douglas Hawkins
Department of Applied Statistics
University of Minnesota, St. Paul, MN

Abstract

Imaging systems are always multivariate, typically involving response functions which relate lightness and intensities, and sometimes even families of functions, that must be optimized to produce the best possible images.

A preferred method for minimizing the work involved in empirically optimizing a product or process is that of "designed experimentation".

Until recently there has been no widely accepted method that enabled the application of designed experimentation to multivariate problems, and therefore to imaging products and processes. All this is changed with the advent of the desktop computer. We will describe a surprisingly simple and powerful, though computationally intensive, method for optimizing multivariate systems and imaging systems through use of quality loss functions and designed experiments. In addition to a better product, the method discussed will produce a more robust manufacturing process as well.

Summary of Major Topics in the Talk

- Introduction, History, and Motivation
- The origin of the univariate quality loss function
  "Good" and "Bad" characteristics
  Taguchi: *Any departure of a characteristic from nominal creates economic loss, and the greater the departure, the greater the loss.*
- Representing quality mathematically
- Extension of the univariate quality loss function to multiple variables
  (Quality loss function need not be complete to be useful!)
- Designed Experimentation and Response Surface Methodology
  (Statistical method for efficiently relating inputs to outputs)
- Defining the quality loss function in terms of process inputs

Some differences between "classical" experimental design and experimental design for use in quality loss functions.
  Classical modeling
  (You don't need to use designed experiments to use loss functions for optimization.)
- Optimizing a process or product with loss functions
- Robustness Characteristics of optimization with loss functions
- Implementation Notes
  Construction of Loss Functions
  The Need for Higher Order Designs
  Optimizations with RSM and with Loss Functions.
  Minimizing QLP on the Computer
- Review

Introduction, History and Motivation

(Many of the details of this talk may be found in a paper accepted for publication in the Journal of Quality Technology sometime in 1997.)

All of us have to deal with optimization problems in our daily lives, such as, "How do I get enough exercise and still get all my work done?" or "How do I save money for retirement, and still have some fun?"

Many in this audience have had to deal with industrial optimization problems such as making the best hard copy print with the least dye, or maximizing an output of a chemical reactor while minimizing the cost of ingredients.

The kinds of optimization problems we focus on in this talk are those where (a)the process is reasonably well defined (say, a chemical manufacturing process) with a dozen or fewer continuously variable inputs and (b) there is a large number of output variables (a dozen or more variables, or a sampled function) that must be controlled to sustain product quality.

The particular problem that led to this paper arose in the manufacture of Polaroid Spectra® instant photographic film. A Polaroid instant picture includes three components that are primarily chemical in nature: the negative, the reagent, and the positive sheet coatings. When the color negative is manufactured, its properties may be perturbed by variations (usually due to raw materials) so that, with no subsequent adjustments, the final pictures would show unacceptable run to run variations.

The cure for this problem, implemented since the days of the first Polaroid instant photographic products, was to adjust the composition of the reagent—the chemicals in the pod that are activated as the picture leaves the camera—so as to compensate for the variation in the negative. In those early days, adjustment of the chemistry was accomplished using a combination of experimentation and experience. As products evolved in complexity and improved in overall quality, the adjustment process became more complex, and in 1984 we began to hunt for more automatic methods of performing the adjustment.

We will describe here the most successful strategy that we found to cope with the problem. The first experiments on optimization using loss functions based on response surfaces derived from designed experiments were done in 1984. To give you some idea of how valuable the approach as been, By early 1994 well over 500 sets of designed experiments for multiple response optimization had been successfully completed.

Capsule Summary of Talk

At this point, we state the bottom line.

If the problem is, "How do we adjust a complex manufacturing process to achieve highest product quality and greatest process robustness?"

The proposed solution is:

1) Use experience or designed experiments in screening mode to isolate the input variables that have the most control over the worst quality problems.
2) Use higher order designed experiments to construct response surfaces for all significant process responses or characteristics.
3) From the response surfaces, construct quality loss functions to measure the loss due to departures from target in terms of process inputs.
4) Finally, find minima of the quality loss function with respect to the process inputs. These minima are the operating conditions of highest quality (least quality loss) and greatest process robustness.

Beginning of Explanation: The Univariate Loss Function

There was once a time in American manufacturing when it was believed that, for the most important characteristics of a manufactured part, there was a "band of acceptability". Within this band the product was "good" and outside the band the product was "bad". A mechanical part either fit or it did not fit.

The "good" / "bad" classification worked for the Model T assembly plant mostly because the manufacturer's job was to get the automobile out the door of the factory. After that it was the customers problem. The notion of "failure probability" was not generally understood. The belief in the utility of a region of acceptable variation prevailed for many years.

To help change the "good vs. bad" mindset, Genichi Taguchi introduced the loss function as a way to model the concept that any departure from intended targets, however small, creates economic loss. The Taguchi loss function qualitatively describes the "economic loss to society" arising from errors of all sorts in hitting a target, including random variation from inside and outside the process as well as systematic errors in the process.

Representing Quality Mathematically

Conceptually at least, one way of representing quality mathematically is to estimate the economic loss due to departure from target values of any particular parameters. When we talk about a quality loss function in this paper, usually we mean something proportional to the function that describes the "economic loss to society" of variation from specification.

Multivariate Quality Loss Function

Let us assume that the economic loss function associated with each characteristic of a product can be estimated for each use of the product. Then for a variety of independent uses of a product the losses can be weighted and summed to find the total loss to society for all variation of the product characteristics. Even where we cannot state exactly what the dollar loss associated with a deviation from target value is, we can make estimates of the importance of each target for maintenance of quality. With these estimates we can combine the separate responses into a single loss function which can be used for process optimization.

Response Surface Methodology and Designed Experiments

Assume for the moment we have some sort of batch chemical manufacturing step. Imagine we could write an equation which described the yield of the chemical reaction as a function of batch temperature and the concentration of two ingredients, C1 and C2. The result would be a function with dependent units of moles, which varied with temperature, C1, and C2. This function is a surface embedded in the space (moles, temperature, C1,C2).

In many cases we do not know the exact relation between our input and output variables, so we must approximate the response function using experimental data. To this end, we assume some sort of polynomial model, run some experiments, then least squares fit the data to the model.

The methods devised to experimentally determine coefficients in an assumed model most efficiently are called "designed experiments", because the experiments are designed to capture the maximum amount of needed information from a minimum number of runs.

Methods in which response surfaces are estimated using designed experiments are called "response surface methodology", or RSM.

For constructing quality loss functions, unlike conventional response surface methodology, we generally do not look at the significance levels of individual coefficients within the response surface polynomial. What we do care about is the accuracy of the predictions made about the product performance. This estimate cannot be made until we have constructed a complete loss function which includes all variables, not just some of them.

The thing to remember here is that response surface methodology will assist you in forming simple polynomial models of the process responses .

Let's return now to quality loss functions.

Quality Loss Function:; The Global Quality Loss Function

To help clarify our thinking, we imagine first the global quality loss function GQL, which describes the loss of product quality and product value as various characteristics of the product depart from their specified, or target, values. The function we write down does not describe the quality loss of the product, but rather uses a parabolic approximation for the loss of quality due to deviation of responses from their targets. We recognize that each characteristic, V_r, may vary from the fixed target, T_r, because of both random variation and systematic errors or design compromises. The global quality loss function GQL in terms of measured responses is:

$$GQL = \sum_{r = 1,R_g} W_r \{ V_r - T_r \}^2 \qquad (1)$$

Here, V_r is a measured response indexed on r, T_r is the target or value of V_r at optimal quality, and W_r is a weight factor. The sum is taken over all of the responses that contribute to product quality, R_g in number. The weight factors W_r are scaling parameters that relate squared error in hitting the target to economic or quality loss.

Process Related Quality Loss Functions

The next step is to shift our thinking from outputs to inputs by expressing the output variables that define quality in terms of the process input variables. This leads us to the process related quality loss function, GQLP. To actually construct this function, we need RSM or other modeling method to create a quantitative models for each response. We can then approximate the quality of the final product as a function of the process inputs.

If the process inputs are $(X_1, X_2, - -)$ and the model for the response V_r is expressed as a function of these inputs as $Y_r(X_1, X_2, - -)$ then we write equation 2 to replace 1:

$$GQLP= \sum_{r = 1,R} W_r \{ Y_r (X_1, X_2, - -) + e_r - T_r \}^2 \qquad (2)$$

In equation 2, $Y_r(X_1, X_2, - -)$ is the function of the inputs describing the response V_r and $e_r = V_r - Y_r$ is the error, both systematic and random, associated with the regression description in terms of inputs rather than in terms of responses. R is the number of process variables for which we can usefully identify a response surface polynomial Y_r. R differs from the R_g in eq. 1 because it is not always appropriate to find response surfaces for all process response variables.

If the errors, e_r, are small compared to the errors in hitting the targets, $(Y_r - T_r)$, we can neglect e_r and rewrite GQLP as:

$$QLP = \sum_{r = 1,R_g} W_r \{ Y_r (X_1, X_2, - -) - T_r \}^2 \qquad (3)$$

At this stage we have three loss functions to think about. The first, eq. 1, is the complete loss function applicable to all quantifiable characteristics of the product. The second loss function, eq. 2, includes a subset of the responses in eq. 1 for which response surfaces can be identified. Finally in eq. 3 we include only response polynomials and targets, and do not explicitly include random errors.

Process Optimization with GQLP and QLP

So long as the random effects in the process are independent of the process inputs, or alternatively so long as the dominant loss of quality arises from systematic errors in hitting targets, then the minima of the process input related quality loss function, GQLP, occur at particular values of the inputs $(X_r, X_2, - -)$. These minima define the process operating conditions that result in the least amount of quality loss for the process or product.

Robustness Characteristics

Another characteristic of the minima of GQLP comes from the definition of "minimum". At a minimum the rate of change of a function with respect to change in the inputs is zero. This raises a point of some significance: the set of process inputs which produces the highest quality also produces the greatest stability or robustness against variations in quality due to input variation.

Construction of Loss Functions

One of the difficulties of the proposed method is the lack of an obvious objective basis for choosing the weight factors, the W_r. In most practical situations, one cannot realistically estimate the "loss to society" of a departure from specification. One strategy is to find factors which will eliminate the units of measure of the squared error, that is, scale so that the "usual error" will have a value of unity, then subsequently assess the relative importance of each of these nondimensional squared error. Another method might be to scale the squared error by its expected value based on independent criteria, perhaps experimental error, then choose a multiplier of the scaled value which expresses the contribution to quality loss. This method produces a quality weighted net variance for the process. Other strategies could compare deviations from target to control limit ranges, or other process control variables.

For many imaging systems it is feasible to do testing of computer generated images with panels of viewers to find the comparative importance of various imaging errors. Whatever strategy is employed, keep in mind that the weight factors are the means to the end of finding the best process settings, and insofar as the targets can be hit, the weight factors are irrelevant.

For constructing loss functions in which a response must be maximized or minimized, we have successfully used the half parabola method of Tribus and Szonyi (1989). This method has the advantage over step functions in that it maintains continuous first derivatives which can be important for finding the optima numerically. Broad regions of acceptability that are not well described with a single parabola can be described with split parabolas.

Comments on the Need for Higher Order Designs

As use of loss functions for optimizing processes becomes more sophisticated, and both more outputs and more inputs are considered, the basic probability of finding important nonlinear interactions increases.

Similarly, as the process becomes more capable and errors get smaller, the probability increases that interactions between the inputs will be observed as significant.

Assessment of nonlinearities requires higher order designs, that is, experimental designs that can provide estimates of quadratic, or sometimes even higher, power coefficients within the response surface polynomial.

The key point is that for multivariable optimizations, the user should anticipate encountering non-linear effects and interactions which must be supported by the basic experimental designs.

For fitting a second-degree polynomials we have found the central composite designs to be particularly useful.

Some Differences Between Optimizations with RSM and with Loss Functions.

It is conventional in response surface methodology (RSM) to subject each coefficient in the polynomial describing the surface to a test for significance and to justify the use of the coefficient in the final regression. Our initial implementation using designed experiments leading to RSM's and then to loss functions implemented this practice. During our pilot studies we initiated a search for the significance levels that gave us the best predictions. After extensive testing and many heated discussions we concluded that the most reliable predictions were obtained when all coefficients were retained in the highest order design judged necessary by a knowledgeable experimenter. The statistician may ask, How could this be? Is not over paramaterization undesirable? For loss function minimization the answer seems to be "no, over parameterization is not harmful." While we are not yet certain, there may be two explanations. One is that for loss function minimization, the role of the response surface is simply to represent the experimental data, noise and all, and the more terms that are available from the regression to accomplish the representation, the better will be the final representation. A second explanation is that in the construction of a single loss function from a large number of individual surfaces, useful noise averaging effects outweigh errors introduced by overparameterization. For the reagent matching work, we estimate that we have roughly three times as many data points as there are independent responses, so the overall significance within the loss function of coefficients that are similar across multiple experiments gets higher with each experiment included in the loss function. A coefficient which might be judged insignificant for any one experiment can become extremely significant if it appears consistently within the responses most important at the loss function minimum.

Minimizing QLP on the Computer

Finding the minima of eq. 3 with respect to the process inputs requires the use of a computer. Possible methods range from simple "steepest ascent" method.to elaborate forms of Newton's method. [Chapter 8 of Fiacco and McCormack]. Dixon has useful material, and Aubin addresses directly the issue of loss function minimization.

Summary

We have presented a novel and powerful method for dealing with a problem frequently encountered in the manufacture of complex products, that of process tuning where multiple criteria must be met to achieve highest product quality. The basic strategy is to describe the response surfaces with experimentally derived polynomials which can be combined into a single loss function using known or desired targets. Minimizing the loss function with respect to process inputs locates operating conditions which produce the product of highest quality, and most stability.

Acknowledgements

The authors are indebted to S. MacDonald, S. Baistos and L. Clemente for their work on networks and process automation, to Polaroid Corporation for permission to publish this work, and to R. Eckert, E. Lindholm and other supporters of this project over the years.

References

Polaroid Spectra Film

1. Lambert, R. J. (1989). "A Technical Description of Polaroid Spectra Film" . *J. Imaging Tech.* **15** pp. 108-113
2. McCaskill, E. S. and Herchen, S. R. (1989). "Electro-chemical Properties and Mechanisms of Action of Developers in Spectra Film". *J. Imaging Tech.* **15**, pp 103-107
3. Meneghini, F. (1989). "A New Instant Photography: A Chemists's View". *J. Imaging Tech.* **15**, pp. 114-119
4. Sturge, J. M. ; Walworth,V. ; and Shepp, A., Eds. (1989). *Imaging Processes and Materials, Neblette's Eighth Edition*. Van Nostrand Reinhold, New York

Loss Functions

5. Taguchi, G. and Wu, Y. (1979). "Off-line Quality Control". Central Japan Quality Control Association, Nagaya
6. Ross, P. J. (1988). *Taguchi Techniques for Quality Engineering*. McGraw-Hill Book Company

Minimization Techniques

7. Fiacco, A. V., and McCormack, G. P. (1990). *Nonlinear Programming, Sequential Unconstrained Minimization Techniques*. SIAM, Philadelphia

8. Dixon, L. C. W.; Spedicato, E.; and Szego, G. B., Eds. (1980). *Nonlinear Optimization, Theory and Algorithms*. Birkhauser, Boston, MA,
9. Aubin, J.-P. (1984). *Explicit Methods of Optimization*. gauthier-villars (c)BORDAS, Paris

Robustness

10. Lucas, J. M. (1994). "How to Achieve a Robust Process Using Response Surface Methodology". *Journal of Quality Technology* 26, pp. 248-260

Experimental Design and Response Surfaces

11. Box, G. E. P. and Draper, N. R. (1987). *Empirical Model Building and Response Surfaces*. John Wiley and Sons, Inc., New York
12. "Response Surface Methodology,1966-1988" *Technometrics* **31**, pp 137 -157
13. Myers, R. H. (1971). *Response Surface Methodology*. Allyn and Bacon, Boston
14. Khuri, A. I. and Cornell, J. A.(1987). *Response Surfaces*. Marcel Dekker, Inc., New York, N.Y.
15. Cochran, W and Cox, G. M. (1957). *Experimental Designs* , 2nd edition, John Wiley and Sons,
16. Phadke, M. S.(1982). "Quality Engineering Using Design of Experiments". *Proceedings of the Section on Statistical Education*, American Statistical Association, pp. 11- 22

Response Surfaces for Multiple Variables

17. Del Castillo, E. and Montgomery, D. C. (1993). "A nonlinear Programming Solution to the Dual Response Problem". *Journal of Quality Technology* **25**, July pp. 199-204
18. Derringer, G. C. (1994). " A Balancing Act: Optimizing a Product's Properties". *Quality Progress* **27**, pp. 51-58.
19. Derringer, G. and Suich, R. (1980). "Simultaneous Optimization of Several Response Variables". *Journal of Quality Technology*, **12** , pp. 214 -219
20. Myers, R. H. and Carter, W. H. Jr. (1973) "Response Surface Techniques for Dual Response Systems". *Technometrics* **15**, pp. 301-317
21. Vining, G. G. and Myers, R. H. (1990). "Combining Taguchi and Response Surface Philosophies: A Dual Response Approach" . *Journal of Quality Technology* **22**, pp. 38-45

Half Parabola Loss Functions

22. Tribus, M. and Szonyi, G. (1989). "An Alternative View of the Taguchi Approach". *Quality Progress* **22**, pp. 46-52