OPTIMAL DESIGN OF EXPERIMENTS FOR SENSOR CALIBRATION

Selden B. Crary

Center for Integrated Sensors and Circuits
Electrical Engineering and Computer Science Department
The University of Michigan, Ann Arbor, MI 48109-2122, USA

ABSTRACT

In this paper we present a framework for the design of experiments for sensor calibration. The theoretical ideas, which are based on the mathematical theory of optimal design of experiments, are demonstrated using a new software program, I-OPT, which was developed by Crary, Snow, and Hoo.[1] This program has been used to find I-optimal designs for both a silicon capacitive absolute-pressure sensor and a silicon piezoresistive pressure sensor.

INTRODUCTION

It frequently occurs that the response of a sensor is of a known mathematical form. For example, the following is the previously reported closed-form compensation formula for a particular type of silicon capacitive absolute pressure (SCAP) sensor:[2]

$$\text{Compens. press.} = \beta_0 + \beta_1 C + \beta_2 C^2 + \beta_3 T + \beta_4 C^3 + \beta_5 T^2 + \beta_6 CT,$$

where C is the capacitance response of the sensor and T is the temperature. When such a formula is known, a calibration can be effected by making a set of measurements of the variables that appear in the compensation formula (C and T in the example), along with corresponding measurements of a calibration standard (a pressure standard in the example). Linear regression techniques can then be applied to determine best values for the coefficients, the β's in the compensation formula, and the formula used subsequently to digitally compensate individual sensors both for nonlinearities and for cross-parameter effects on the sensor's response.[2]

For reasons of economy, one is interested in accuracy and efficiency of calibration. This demands consideration of the tradeoffs between accuracy and calibration cost, which is strongly related to the number of calibration points used in the determination of the coefficients in the above formula. At present in the field of solid-state sensors, *ad hoc* procedures are used to determine the number of calibration points and the settings of the variables used in the calibration,[3,4] despite the fact that calibration is often one of the greatest single factors influencing sensor cost.

Once the functional form of the compensation formula is known, questions arise as to the best means of carrying out the required calibration. If there are *m* terms in the compensation formula, are *m* distinct calibration points adequate, as long as one avoids ill-conditioning of the matrix that must be inverted? At what values of the variables should these points be taken? Perhaps many more than *m* calibration points are required, in order to effect a sufficiently accurate calibration. Or perhaps fewer than *m* are required, since there may be correlations in sensor characteristics from sensor to sensor. As an example of this last possibility,

imagine that it had already been established that the constant term β_0 in the SCAP compensation formula was the same for all sensors, within allowable error. If this were the case, then there might be only *m-1* independent β's to be determined for each sensor. This paper draws upon the mathematical theory of optimal design of experiments to begin to establish an appropriate framework for addressing questions and issues such as these.

VARIANCE OF COMPENSATION

In general there are two types of error present in the type of sensor compensation discussed here. There is an error introduced by the noise in the calibration procedure, which in what follows will be called measurement error. There is also error introduced by the inevitable failure of the compensation formula to model the sensor perfectly. This type of error will be called model error. Given their quite distinct origins, it is reasonable to assume the independence of the contributions of measurement and model error to the variance in the compensation. This can be expressed as follows:

Variance of compensation = Measurement variance +
Model variance .

The statistics literature contains discussions of the distinctions between measurement variance and model variance.[5,6] A brief summary of three important cases follows:

(1) In cases where modeling error dominates, the best experimental designs are usually ones that spread out, more or less evenly, the fixed number of calibration points. In these cases, the common *ad hoc* procedure of spacing calibration points evenly on a rectangular grid in the domain of calibration[4] performs reasonably well. This is especially true of designs containing roughly two times as many, or more, calibration points as there are points in the model function, for in this case it is easy to define an appropriate rectangular grid.

(2) When an adequate model exists, uniformly spaced designs are not optimal in general, and in many cases significant improvements can be obtained with other methods. This is accomplished, as described below, by defining an appropriate objective function, which provides a measure of compensation error over the range of use of the sensor, and then finding the design that minimizes this objective function. Good global compensation formulas can be established for some types of sensors, using robust statistical methods such as stepwise regression,[2] and such methods are becoming more widely available in statistical software packages. Thus it can be expected that model error can be decreased significantly by application of statistical methods.

(3) If minimization of the number of calibration points is critical, and one seeks a design with the number of calibration

points equal to or somewhat larger than the number of terms in the model function, then it is useful to have available a lower bound on some well defined measure of the compensation error, as well as a means of comparing potential designs with this bound, whether or not model variance is dominant. As an example, Reference 4 describes a set of designs of experiments for calibration of a piezoresistive pressure sensor. The total error band ranges from 0.75%, when a 3-term compensation function and a 9-point design are used, to 0.04%, when a 16-term compensation function and a 25-point design are used. It appears that model error dominates for the system with the larger error band. However, it is questionable whether a 9-point design is actually necessary for this low-end sensor. Examination of alternate calibration-design options may prove useful in this and similar instances.

THEORY

We begin with a review of basic linear-regression theory. In what follows, boldface denotes vectors and matrices, a superscript T denotes the matrix transpose operation, and a circumflex denotes expected value. The mathematical model in this case is linear in the coefficients, as in

$$Y(\mathbf{x};\beta) = \beta_1 f_1(\mathbf{x}) + \beta_2 f_2(\mathbf{x}) + ... + \beta_m f_m(\mathbf{x}) \quad ,$$

and this allows for a broad class of functions, including multivariate polynomials such as $Y=\beta_1+\beta_2 x_1+\beta_3 x_2+\beta_4 x_1 x_2$ or functions with non-linear terms such as $Y=\beta_1+\beta_2 \ln x_1+\beta_3 x_1 \ln x_2$. The functions $f_i(\mathbf{x})$ are assumed to be linearly independent. A set of measurements represented by the column vector $\mathbf{Y}=(Y_1,Y_2, ... ,Y_n)^T$ is made at a set of specified values of the independent variables $\mathbf{x}$ with a set of random errors $\varepsilon = (\varepsilon_1,\varepsilon_2, ... ,\varepsilon_n)^T$, the elements of which are assumed to have zero mean and constant variance σ^2,

$$\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} f_1(\mathbf{x}_1) & f_2(\mathbf{x}_1) & \cdots & f_m(\mathbf{x}_1) \\ f_1(\mathbf{x}_2) & f_2(\mathbf{x}_2) & \cdots & f_m(\mathbf{x}_2) \\ \vdots & \vdots & \vdots & \vdots \\ f_1(\mathbf{x}_n) & f_2(\mathbf{x}_n) & \cdots & f_m(\mathbf{x}_n) \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} .$$

This can be written as $\mathbf{Y}=\mathbf{X}\beta+\varepsilon$, where $\mathbf{X}$ is called the design matrix. A key result from regression theory is that the best unbiased linear estimator of the coefficients is given by $\hat{\beta}=(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$.[7] In the special case that the number of data points exactly equals the number of terms in the regression equation, m=n, the familiar rule $(\mathbf{AB})^{-1}=\mathbf{B}^{-1}\mathbf{A}^{-1}$ can be applied, and the above equation becomes $\hat{\beta}=\mathbf{X}^{-1}\mathbf{Y}$. Furthermore, the variances in the estimates of the parameters $\hat{\beta}$ are given by $\sigma^2(\hat{\beta})=\sigma^2(\mathbf{X}^T\mathbf{X})^{-1}$, and the variance in the fit function is $\sigma^2(\hat{Y}(\mathbf{x}))=\sigma^2\mathbf{f}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{f}$, where $\mathbf{f}=(f_1(\mathbf{x}), f_2(\mathbf{x}), ... , f_m(\mathbf{x}))^T$.

OBJECTIVE FUNCTIONS

Several optimization functions are evident. Three of the most useful are the following:

D-optimality. Because the variances in the estimates of the parameters are proportional to the matrix $(\mathbf{X}^T\mathbf{X})^{-1}$, the determinant of this matrix provides a measure of the overall uncertainty of the parameter estimates. A design that minimizes the determinant of $(\mathbf{X}^T\mathbf{X})^{-1}$ is called D-optimal. It has been shown that this criterion is equivalent to minimizing the

error in the parameter estimates, in the sense of minimizing the volume of the confidence regions for finding the actual parameters.[8] D-optimality is the most widely studied optimal design criterion. It is an interesting fact that the D-optimal design is not necessarily optimal for the purpose of generating a response model, such as the compensation formula we seek for sensors. In fact, a D-optimal design can be significantly sub-optimal for such tasks.

G-optimality. A design that minimizes the worst-case expected error in prediction is called G-optimal.[9] Such a design would minimize the worst-case expected error in applying a given compensation formula, under the assumption that the functional form of the compensation formula is known.

I-optimality. When the goal of the calibration is to minimize the average variance in prediction over the range of use of the sensor, an appropriate objective function is the following:

$$\min_\omega \int_{\mathbf{x}\in X} E\{[\hat{Y}(\mathbf{x})-Y(\mathbf{x})]^2\} d\mu(\mathbf{x}) =$$

$$\min_\omega \int_{\mathbf{x}\in X} \mathbf{f}^T(\mathbf{x}) (\mathbf{X}^T\mathbf{X})^{-1} \mathbf{f}(\mathbf{x}) d\mu(\mathbf{x}) ,$$

where $\min_\omega$ indicates that the experimental design ω is sought that minimizes the integral over the set of points $\mathbf{x}\in X$ at which the sensor will be used. Weighting of different regions of the response is accomplished through the differential $d\mu(\mathbf{x})$. This criterion is called I-optimality and has been detailed in the design-of-experiments literature.[10]

AVAILABLE SOFTWARE

Finding optimal designs of experiments is a computationally intensive task, as has been well established.[10] Fortunately, recent advances in speed of computation, coupled with new algorithms such as simulated annealing,[11] are bringing the determination of optimal designs within the range of commonly available capabilities, without undue expense.[10,12]

Software for finding optimal designs of experiments was reviewed by Nachtsheim[13]. Software for finding D-optimal designs on finite grids have become widely available, since they were included in popular statistical-software systems, such as RS/Discover.[14] Meyer and Nachtsheim have discussed software for finding D-optimal designs on continuous spaces.[12] Welch's ACED software[15] finds designs approximating I- and G-optimal designs, in which potential design points are restricted to a fairly coarse grid. However, virtually nothing exists for determining I- or G-optimal designs on continuous spaces, which are the relevant criteria for sensor calibration, with the exception of a series of programs by Haines[10] that can be used for two small classes of functions.

I-OPT

In the absence of adequate software, Crary, Snow, and Hoo[1] have developed a FORTRAN program called I-OPT that finds I-optimal designs of experiments for arbitrary multivariate polynomial models. The program is essentially an extension of earlier work by Haines,[10] using a modified Fedorov procedure,[8] as well as simulated annealing for functional minimi-

zation, in order to find the optimal, or at least a near-optimal, design. They have extended the earlier work of Crary and Snow[16] by adding an efficient downhill-search algorithm, essentially their own variant of a modified-Powell search, to speed convergence once the initial annealing search has led to the basin in which the optimal solution lies, with a high degree of confidence.

EXAMPLES

Figure 1 shows the designs found with I-OPT for the compensation formula given above for a variety of allowed numbers, N, of distinct calibration points measured with equal accuracy. I-OPT has the following additional features that are useful in the context of sensor calibration: (1) the domain of calibration may differ from the domain in which the compensation formula must be accurate, as in Figure 2a; (2) weights may be applied to the domain in which the formula is applied, as in Figure 2b; (3) the calibration points may have different accuracies associated with the measurement of the relevant variables; and (4) the cost of a calibration point may vary over the region of calibration, thus allowing for optimization based on rather general cost functionals, rather than on accuracy alone.


N=7   N=9

N=12   N=14
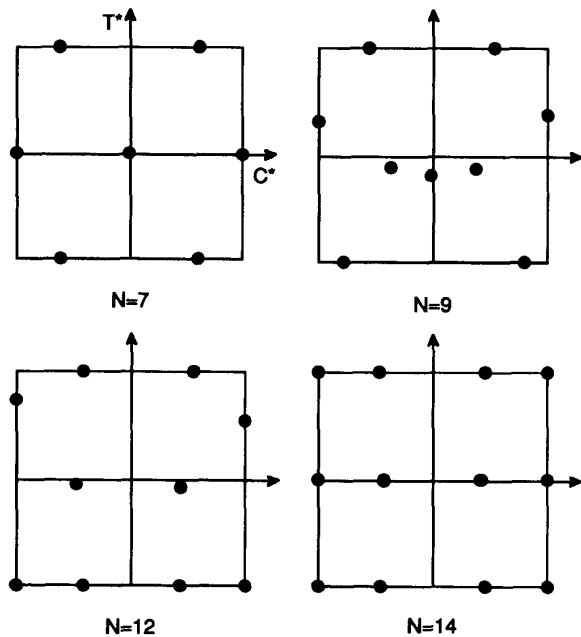
Figure 1. Several optimal designs of calibration experiments for the seven-term silicon capacitive absolute-pressure sensor response are shown, with numbers of calibration points ranging from N=7 to 14. The variables C and T have been transformed to reduced variables C* and T* so that their ranges are the closed interval [-1,1]. The objective function being minimized is the integrated expected squared error in the compensation formula (I-optimality condition), assuming that all the calibration points are made with equal accuracy. For the runs with N=12 and 14 points, the two filled circles shown nearest the center of the domain are actually double points, indicating that redundant measurements should be made at each of these points.

Figure 3 shows a 15-point design, found using an early version of I-OPT, based on the following 15-term polynomial compensation formula for a piezoresistive sensor that was presented previously:[2]

Compens. press. $= \beta_0 + \beta_1 T + \beta_2 T^3 + \beta_3 R + \beta_4 R^4 + \beta_5 R^5 + \beta_6 R^2 T +$

$\beta_7 RT^3 + \beta_8 R^3 T^2 + \beta_9 R^2 T^3 + \beta_{10} T^4 + \beta_{11} RT^4 + \beta_{12} R^3 T^4 +$

$\beta_{13} T^5 + \beta_{14} R^3 T^5$.

A comparison of designs based on different criteria is instructive. Figure 4 shows the value of the integrated variance as a function of the number of calibration points, based on the 7-term SCAP compensation function, for different criteria. The points connected by line segments represent designs found with I-OPT. Other criteria have higher values of the objective function. For N=7 and N=12, D-optimal designs that were found using Meyer and Nachtsheim's simulated annealing program[12] are seen to be significantly sub-optimal

N=7   N=7
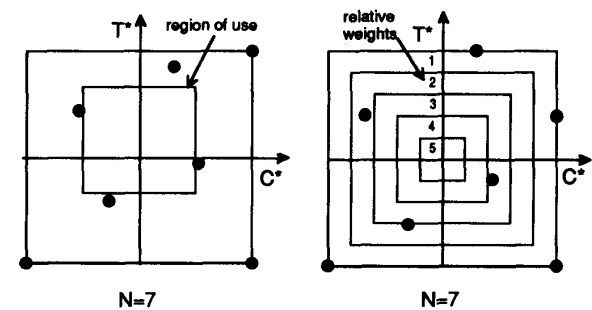
Figure 2. (a) The domain of calibration can be different from the domain where the sensor will be used. Shown here is the minimum-point (N=7) design when the region of use will be C*=[-1/2,1/2] and T*=[-1/3,2/3]. As compared to the N=7 design in Figure 1, this design places more of the calibration points near the center of the domain. (b) The region of use can also be weighted. In the example shown here, the center of the domain of use has a relative weight of five compared to the periphery of the domain, meaning that the relative cost of variance in the compensation formula in the center is being taken as five times the cost at the periphery. For this minimum-point design the relative weight is graded from center to periphery, as shown.
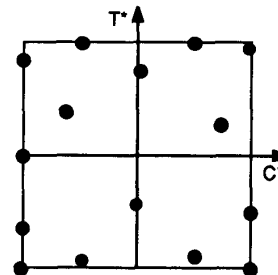
Figure 3. The 15-point I-optimal design for the 15-term polynomial model of a silicon piezoresistive pressure sensor.

406
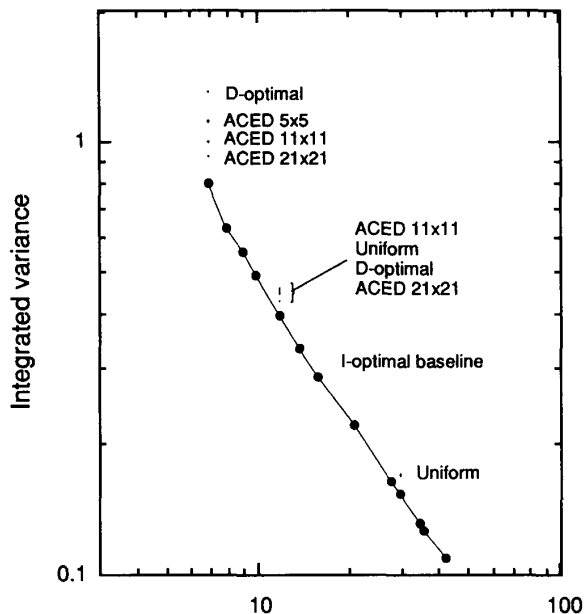
Figure 4. A comparison of designs based on various criteria. The ordinate value is the integrated variance normalized to give the average variance over the entire domain of calibration.

in the I-optimality criterion. The integrated variances of designs found with ACED's AV criterion[15] using grids of 5x5, 11x11, and 21x21 points are also shown. Finer meshes are computationally expensive in ACED and are not recommended.[15] Simulated annealing, which is used in I-OPT, overcomes this computational hurdle. Finally, uniform designs based on a rectangular grid are shown for N=4x3=12 and N=6x5=30, and these are seen to be close to optimal.

SUMMARY

Reiterating the main points: a variety of methods is available for finding designs of experiments, I-OPT establishes one possible baseline for comparison of designs, I-optimal designs are particularly attractive if the model variance is relatively small or if it is desired to keep the number of experiments to a minimum, although in this last case, comparisons should be made with designs based on other criteria.

ACKNOWLEDGMENTS

The author thanks Mark Snow and Ling Hoo for their assistance with aspects of the development of I-OPT.

REFERENCES

[1] S. B. Crary, M. E. Snow, and L. Hoo, "I-OPT: Software for Finding I-Optimal Designs of Experiments," in preparation.

[2] S. B. Crary, W. G. Baer, J. C. Cowles, and K. D. Wise, "Digital Compensation of High-Performance Silicon Pressure Transducers," Sensors and Actuators, vol. A21, pp. 70-72, 1990.

[3] H.-R. Tränkler, "Signal Processing," in Sensors: A Comprehensive Survey, W. Göpel, et al. (eds.). New York: VCH Pub., 1989, ch. 10, pp. 302-303.

[4] J. Brysek, K. Petersen, J. R. Mallon, L. Christel, and F. Pourahmadi, Silicon Sensors and Microstructures. Fremont, CA: NovaSensor, 1990, ch. 8, pp. 8.23-8.28.

[5] G. E. P. Box and N. R. Draper, "A Basis for the Selection of a Response Surface Design," J. Amer. Statist. Assoc., vol. 54, p. 622, 1959.

[6] W. J. Welch, "A Mean Squared Error Criterion for the Design of Experiments," Biometrika, vol. 70, pp. 205-213, 1983.

[7] J. Netter and W. Wasserman, Applied Linear Statistical Models. Homewood, IL: Irwin, 1974.

[8] V. V. Ferorov, Theory of Optimal Experiments. New York: Academic Press, 1972.

[9] A. Hedayat, "Study of Optimality Criteria in Design of Experiments," in Proceedings of the International Symposium on Statistics and Related Topics, M. Csörgö et al. (eds.), Ottawa, Ont., May 5-7, 1980, pp.39-56.

[10] L. M. Haines, "The Application of the Annealing Algorithm to the Construction of Exact Optimal Designs for Linear-Regression Models," Technometrics, vol. 29, pp. 439-447, 1987.

[11] S. Kirkpatrick; C. D. Gelatt, Jr.; and M. P. Vecchi; "Optimization By Simulated Annealing," Science, vol. 220, pp. 671-680, 1983; and R. H. J. M. Otten and L. P. P. P. van Ginneken, The Annealing Algorithm. Boston: Kluwer Academic, 1989.

[12] R. K. Meyer and C. J. Nachtsheim, "Constructing Exact D-Optimal Experimental Designs By Simulated Annealing," Amer. J. of Mathematical and Management Sciences, vol. 8, p. 329, 1988

[13] C. J. Nachtsheim, "Tools for Computer-Aided Design of Experiments," J. of Quality Technology, vol.19, pp. 132-160, July 1987.

[14] RS/Discover, Version 2.0 (1989), BBN Software Products Corporation, 10 Fawcett St., Cambridge, MA 02138.

[15] W. J. Welch, "Computer-Aided Design of Experiments for Response Estimation," Technometrics, vol 26, pp. 217-224, August 1984; and W. J. Welch, "ACED: Algorithms for the Construction of Experimental Designs: User's Guide, Version 1.6 (1987), available from W. J. Welch, Dept. of Statistics and Actuarial Science, Univ. of Waterloo, Waterloo, Ont., Canada N2L 3G1.

[16] M. E. Snow and S. B. Crary, "The Use of Simulated Annealing in the I-Optimal Design of Experiments," presented at the 1990 Annual Meeting of the Michigan Academy of Sciences, Arts, and Letters, Albion, MI, March 9-10, 1990 and submitted for publication in The Michigan Academician.