Modeling and Analysis Strategies for Failure Amplification Method

SHUEN-LIN JENG

National Cheng Kung University, Tainan 701, Taiwan

V. ROSHAN JOSEPH and C. F. JEFF WU

Georgia Institute of Technology, Atlanta, GA 30332-0205

The failure amplification method (FAMe) uses a special type of factor known as amplification factor to amplify the failure probability so as to maximize the information in the experiment. A general strategy for model building in FAMe is proposed by using generalized linear models (GLM). The best design settings for improving the process capability are determined through carefully selected GLMs and loss functions. Two real experiments for improving the quality of printed circuit boards are used to illustrate the proposed strategy.

Key Words: Design of Experiments; Generalized Linear Models; Model Selection; Process Capability; Robust Parameter Design.

DESIGNED EXPERIMENTS are widely used for reducing failure or defect rate in manufacturing processes. The experiment is carried out by changing the factor settings according to the levels specified by the investigator and observing the number of failures. When the probability of failures is small, it can happen that very few failures occur in the experiment. With such an outcome, it is difficult or even impossible to build an adequate model and obtain optimum process settings. To overcome this difficulty, Joseph and Wu (2004) proposed a novel experimentation strategy known as failure amplification method (FAMe). In FAMe, an amplification factor is selected based on the physical knowledge of the process and is used to amplify the failures. The experiment is then performed at the amplified conditions to ensure that an adequate number of failures are observed, which will provide sufficient information for modeling, analysis, and optimization. It is assumed that, if the process is improved at the amplified conditions, then it is also improved at the normal conditions. One may raise the question: if the failure rate in the normal conditions is small, then why should we even perform an experiment? The answer is to improve the process capability. If the process capability is not improved, then even with a slight shift in the process conditions (due to some special causes), the failure rate can shoot up. This is pictorially depicted in Figure 1. Thus, it is essential to reduce the failure rate as much as we can.

The models proposed in Joseph and Wu (2004) can be simplistic and may not be adequate to deal with more complex processes. In this article, we propose new modeling strategies and provide a general framework for the analysis of experiments using FAMe. Because the experiment is performed at the amplified conditions, we need to *extrapolate* to the normal conditions for optimization. Because of this, carefully selected models should be used for the analysis. The performance of these models at the tails, i.e., regions with low failure rate, is critical for obtaining accurate results during extrapolation. Gener-

Dr. Jeng is an Associate Professor in the Department of Statistics and a member of ASA. His email address is sljeng@mail.ncku.edu.tw.

Dr. Joseph is an Assistant Professor in the H. Milton Stewart School of Industrial and Systems Engineering and a member of ASQ. His email address is roshan@isye.gatech.edu.

Dr. Wu is the Coca-Cola Chair Professor in Engineering Statistics in the H. Milton Stewart School of Industrial and Systems Engineering and a Fellow of ASQ. His email address is jeffwu@isye.gatech.edu.

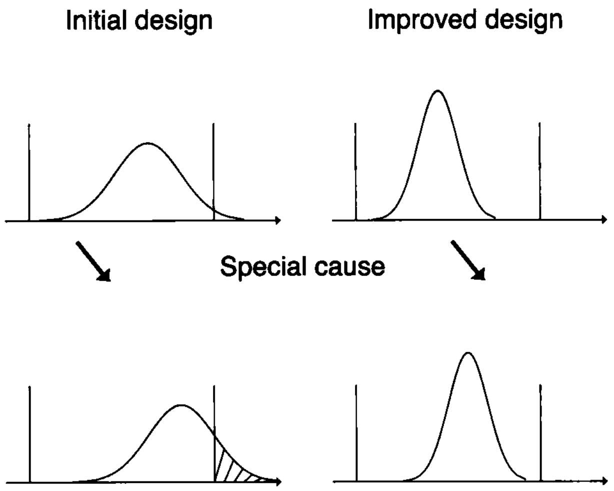

FIGURE 1. Importance of Using FAMe to Improve Process Capability.

alized linear models (GLMs) are suitable for the analysis of failure data (McCullagh and Nelder (1989), Hamada and Nelder (1997)). Our proposed strategy builds appropriate nonlinear models but avoids the complexity of using generalized nonlinear models by utilizing the availability of GLMs in standard software. It is illustrated with the analysis of two real experiments on printed circuit boards (PCBs).

Motivating Examples

We use two examples to illustrate the proposed modeling strategy. The first is an experiment on the inner layer (IL) manufacturing process of PCBs, which was analyzed by Joseph and Wu (2004). It is included here as a comparison of our proposed methods with theirs. The second example is an experiment on the outer layer (OL) manufacturing process of PCBs. The new feature of this experiment is the presence of a noise factor for which the manufacturer has no control in the production process.

In PCBs, circuits can be laid out in different layers. A double-sided (DS) PCB has two layers of circuits (top and bottom), whereas a multilayer (ML) PCB has more than two layers of circuits (4, 6, 8, $\cdots$). To manufacture an ML PCB, first inner layers are manufactured and sandwiched between two copper layers using a pressing operation. Each IL has circuits on two sides. Thus, a 4-layer PCB contains one IL, a 6-layer PCB contains two ILs, and so on. An overall view of the process is shown in Figure 2. Numerous defects are generated during the manufacturing of PCBs, of which shorts and opens in the circuits are the major ones. The first experiment was conducted to reduce shorts and opens in the ILs and the second one in the OLs.

The rejection due to IL shorts and opens is of the order of 1–2% . If we conduct an experiment with such a low failure rate, then we may observe only a few defects in the experiment, which cannot be used for estimation of the model parameters. We can overcome the problem using a large number of PCBs for each experiment, but the cost will be prohibitive. Therefore, a cost-effective strategy is to run the experiment after amplifying the failures. It is known that the shorts will increase if the spacing between the conductors is reduced and opens will increase if the line width of the conductors is reduced. The industry was mainly engaged in the production of circuits with line width/spacing 5 mil (1 mil = 0.001 inch) or higher (up to 15 mil). Therefore, to amplify the failures, the experiments were performed by creating a special circuit pattern with 3 and 4 mil. Joseph and Wu (2004) classified this type of amplification as the complexity factor method because the line width and spacing here are the complexity factors of the product. Thus, the amplification factors are line width and spacing in the following two examples.

Example 1

Eight factors were selected from the IL PCB process for experimentation. The factors and levels are

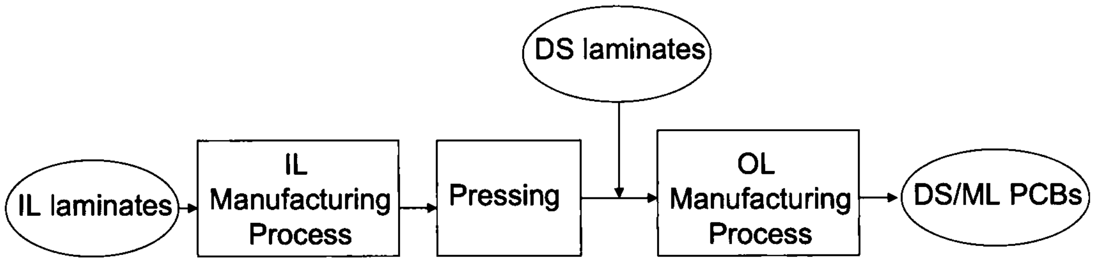

FIGURE 2. Inner Layer (IL) and Outer Layer (OL) Processes in PCB Manufacturing.

TABLE 1. Factors and Levels for the IL PCB Experiment in Example 1

Control factors	Notation	Levels 1	2	3
Preheat	x_1	No*	Yes	—
Surface preparation	x_2	Scrub*	Pumice	Chemical
Lamination speed	x_3	1.2 mpm	1.5 mpm*	1.8 mpm
Lamination pressure	x_4	20 psi	40 psi*	60 psi
Lamination temperature	x_5	95 ^{0}C	105 ^{0}C*	115 ^{0}C
Exposure energy	$x_6(m)$	14	17*	20
Developer speed	x_7	3 fpm	4 fpm*	5 fpm
ORP	x_8	500	530*	560

* Operating levels of the factors in production.

shown in Table 1. An 18-run orthogonal array given in Table 2 was used for the experiment. The experiment was conducted by processing one IL for each run. A special IL circuit pattern was designed only for the purpose of the experiment. The IL contains conductors with 3, 4, 5, 6, and 7 mil of line width and spacing. See Maruthi and Joseph (1999) for details of the experiment. There were 80 pairs of conductors on each IL. Since a pair of conductors gives rise to two opportunities for opens and one opportunity for shorts, there are a total of 160 opportunities for opens under each line width and 80 opportunities for shorts under each spacing. The data on shorts and opens from the experiment are shown in Table 2.

TABLE 2. $OA(18, 2^1 \times 3^7)$ and Data from the IL PCB Experiment in Example 1

Run	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	Opens 3	4	5	6	7	Shots 3	4	5	6	7
1	1	1	1	1	1	1	1	1	33	7	4	0	1	1	0	0	0	0
2	1	1	2	2	2	2	2	2	7	9	1	0	0	4	1	0	0	0
3	1	1	3	3	3	3	3	3	14	3	1	0	0	19	2	0	0	0
4	1	2	1	1	2	2	3	3	2	0	2	0	0	9	0	0	0	0
5	1	2	2	2	3	3	1	1	7	1	2	1	0	22	1	1	1	0
6	1	2	3	3	1	1	2	2	78	30	7	1	1	8	0	0	0	0
7	1	3	1	2	1	3	2	3	9	1	3	0	0	19	1	0	0	0
8	1	3	2	3	2	1	3	1	7	0	1	0	1	4	0	1	0	0
9	1	3	3	1	3	2	1	2	4	3	0	0	0	7	0	0	0	0
10	2	1	1	3	3	2	2	1	6	0	0	0	0	22	1	0	0	1
11	2	1	2	1	1	3	3	2	13	2	0	0	0	34	2	2	0	0
12	2	1	3	2	2	1	1	3	34	5	0	1	3	13	4	1	0	0
13	2	2	1	2	3	1	3	2	8	3	0	0	0	7	0	1	0	0
14	2	2	2	3	1	2	1	3	25	8	0	2	1	25	1	0	0	0
15	2	2	3	1	2	3	2	1	7	0	0	0	0	41	1	0	0	1
16	2	3	1	3	2	3	1	2	10	6	0	0	0	45	9	5	0	1
17	2	3	2	1	3	1	2	3	8	0	0	0	0	3	0	0	0	0
18	2	3	3	2	1	2	3	1	12	2	0	0	1	7	2	0	0	0

TABLE 3. Factors and Levels for the OL PCB Experiment in Example 2

Factors	Notation	Levels 1	Levels 2	Levels 3
Vacuum delay time	x_1	2 sec*	10 sec	—
Developer concentration	x_2	.85 %	1.0%*	1.15 %
Preheat temperature	x_3	150 ^{0}C*	250 ^{0}C	400 ^{0}C
Developer break point	x_4	30 %	50%*	70 %
Exposure step	x_5	13	16*	19
Develop pressure	x_6	1.7 bar	2.0 bar*	2.3 bar
Lamination temperature	x_7	90 ^{0}C	105 ^{0}C*	115 ^{0}C
Lamination speed	x_8	1.5 m/min	1.8 m/min*	2.1 m/min
Noise factor	N	DS/Fresh bath	ML/Old bath	—

* Operating levels of the factors in production.

Example 2

About 5% of the PCBs have shorts in the OLs and about 1% have opens. Most of the OL shorts can be reworked, whereas very few of the opens can be corrected. Thus, in OLs, an open is more serious than a short. Again, eight factors were selected from the image transfer stage of the OL manufacturing process for experimentation. The factors and levels are shown in Table 3. The same 18-run orthogonal array in Table 2 was used for the experiment. The factors are assigned to the columns in the same order given in Table 3. Outer layer process can be used for DS boards or ML boards. Because the manufacturer had no control over the type of the board, it is treated as a noise factor and was included in the experiment. There was one more important noise factor in the process. The solution in the developer was changed after processing several boards. Therefore, its performance toward the end was not as efficient as when it was new. When a cross array is used for experimentation, the number of runs will double if two noise factors are used. To reduce the run size, the method of noise-factor compounding (Taguchi (1986)) was employed. It is known that more shorts will occur with older developer bath. Moreover, ML boards have greater numbers of shorts than DS boards. Therefore, the combination of ML boards with old bath will lead to more shorts than the combination of DS boards with fresh bath. Thus, the two levels of the compounded noise factor (N) are defined as N_1: DS/Fresh bath and N_2: ML/Old bath. As in the IL experiments, a special test pattern was designed for the OLs. In the case of OL boards, there were 88 pairs of conductors. Thus, there is a total of 88 opportunities for shorts and 176 opportunities for opens for each line width and spacing. The data are shown in Table 4. More details about this experiment can be found in Maruthi et al. (1998).

Model Building

Modeling Strategy

When the responses are recorded as failure or success and are independent under some control ($\mathbf{X}$) and noise ($\mathbf{N}$) factors, a binomial model with failure probability $p(\mathbf{X}, \mathbf{N}, M)$ is a natural candidate for describing the outcome, where M denotes the amplification factor. A general model relating failure probability and the variables can be described by

$$f(p) = \lambda(\mathbf{X}, \mathbf{N}, M), \tag{1}$$

where f is the link function in the GLM and λ is an appropriate function of the variables. Now we propose a modeling strategy for choosing f and λ.

For a binomial family, the typical link functions are complimentary log-log (cloglog), logit, and probit:

$$\log\log\frac{1}{1-p}, \quad \log\frac{p}{1-p}, \quad \text{and} \quad \Phi^{-1}(p). \tag{2}$$

The canonical link of binomial model is logit. This link function can be obtained by assuming the underlying variable to follow a symmetric logistic distribution. Similarly, the complimentary log-log (cloglog) and probit links assume the underlying latent variable follows an extreme-value distribution and nor-

TABLE 4. Data from the OL PCB Experiment in Example 2

	DS/Fresh bath										ML/Old bath									
	Opens					Shorts					Opens					Shorts				
Run	3	4	5	6	7	3	4	5	6	7	3	4	5	6	7	3	4	5	6	7
1	1	0	0	0	0	30	12	3	1	2	1	1	0	0	0	42	30	6	3	7
2	15	1	0	0	0	6	5	1	0	1	1	1	0	1	1	38	12	5	3	0
3	43	3	1	0	0	2	4	2	0	0	32	2	0	0	1	18	8	6	1	1
4	2	1	0	0	0	16	8	1	0	2	1	0	0	0	1	27	13	5	3	2
5	114	4	0	0	0	2	2	0	0	0	6	1	0	0	0	17	15	9	3	3
6	4	0	0	0	0	23	8	0	0	1	1	1	1	1	0	47	20	5	3	2
7	4	0	0	0	0	32	13	4	3	4	0	1	0	0	0	41	22	5	2	4
8	20	1	0	0	0	10	1	2	1	0	3	1	0	0	0	21	10	2	1	1
9	112	4	0	0	0	16	13	1	0	0	17	1	0	0	0	43	30	7	6	1
10	115	13	1	0	0	3	1	1	1	0	12	2	0	0	0	27	9	0	3	1
11	1	0	0	0	0	18	5	2	0	5	1	0	0	0	0	46	20	9	4	5
12	23	11	12	10	9	45	15	9	2	2	2	3	2	1	1	62	37	26	16	8
13	68	2	0	0	0	8	5	1	0	0	3	1	0	0	0	18	10	2	5	1
14	5	1	0	0	0	22	4	0	0	0	1	1	0	0	1	22	19	15	7	1
15	3	0	0	0	0	33	9	4	1	1	2	2	0	0	0	42	23	11	6	5
16	7	1	0	0	0	29	7	1	1	0	5	1	0	1	1	40	27	9	4	3
17	73	1	0	0	0	13	1	3	1	0	6	2	1	1	2	40	26	7	5	3
18	1	0	0	0	0	28	12	4	4	2	1	0	1	0	0	37	16	8	5	3

mal distribution, respectively. Figure 3 shows the failure-probability plots of $p = f^{-1}(x)$ for the three link functions and will be further discussed later on. However, none of the above guarantees the best fitting for real data. One may use other reasonable cumulative distribution functions or Box–Cox transformations to obtain links for better fitting. To keep the modeling strategy simple, we stay with the popular link functions in (2). Other flexible links can be considered only when the model fit is not satisfactory. Another advantage of using these links is their availability in most statistical software.

In GLMs, a linear model is used for $\lambda(\mathbf{X}, \mathbf{N}, M)$. However, this may not be adequate for the present problem. The relationship of $f(p)$ with M can be nonlinear. Therefore, we transform M to $h(M)$ and use it in the linear model. Now denote the part of control-noise factors that interacts with $h(M)$ as $g(\mathbf{X}, \mathbf{N})$. Then the linear model in (1) can be simpli-

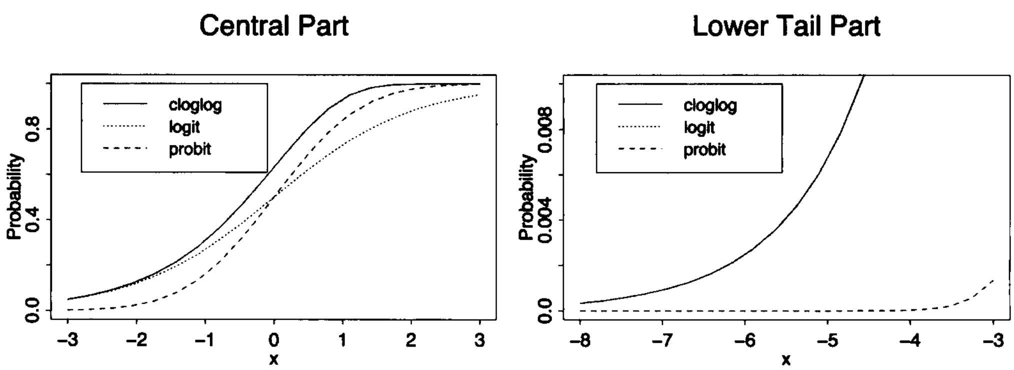

FIGURE 3. The Probability Values of Cloglog, Logit, and Probit Links (cloglog and logit overlap on the lower tail part).

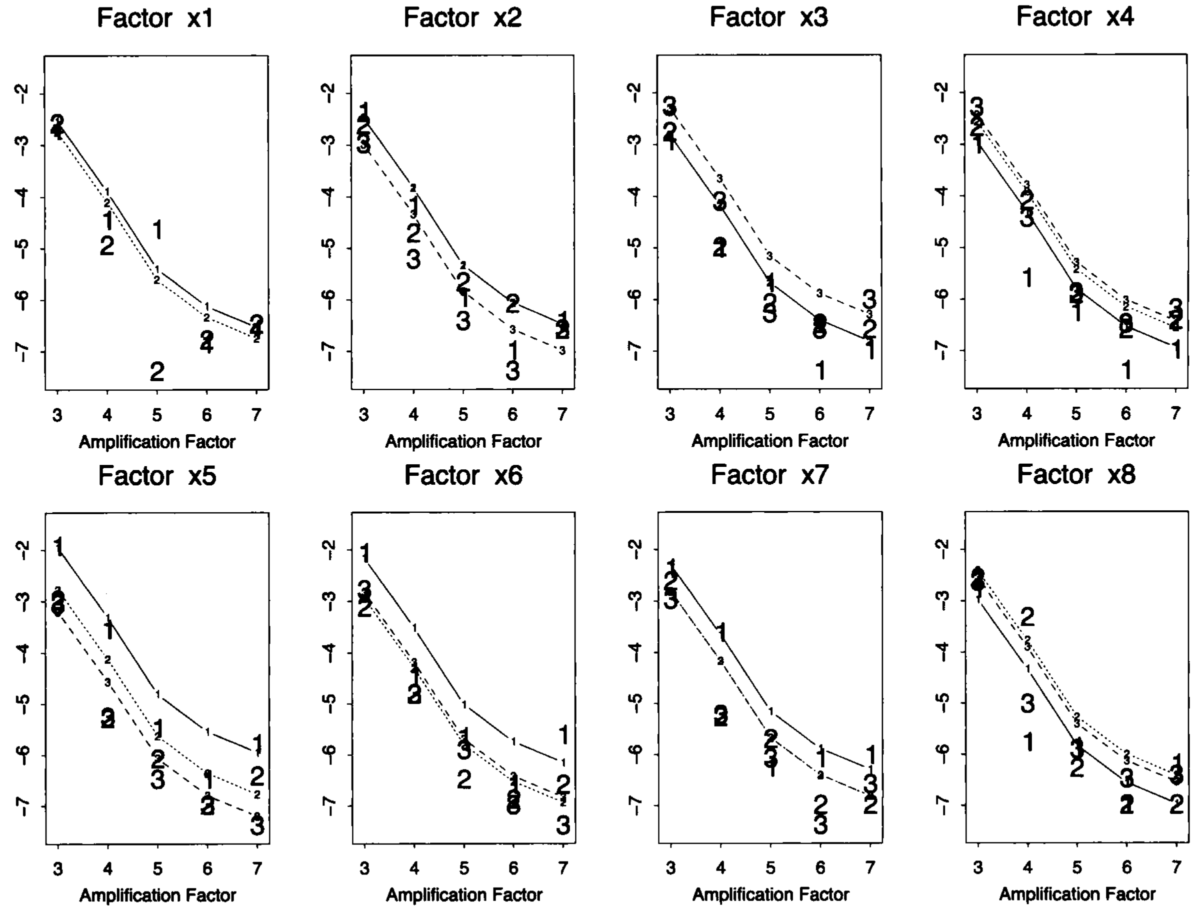

FIGURE 4. Failure Probability with Cloglog Transformation for Opens in Example 1. Average values from the experimental data are plotted along with the fitted lines from (9).

fied to

$$f(p) = \lambda(\mathbf{X},\mathbf{N}) + g(\mathbf{X},\mathbf{N})h(M). \tag{3}$$

We assume that the three-factor and higher order interactions are negligible. This suggests that $g(\mathbf{X},\mathbf{N})$ is a linear model containing only the main effects of $\mathbf{X}$ and $\mathbf{N}$, whereas $\lambda(\mathbf{X},\mathbf{N})$ can contain both main effects and two-factor interactions.

The choice of h is critical, particulary because we need to extrapolate outside the experimental range of M. Because the failure rate is likely to plateau in normal production levels of the amplification factor, simple polynomial functions will not be appropriate for h. Figure 4 shows the failure probability in Example 1 under cloglog link. Clearly, there is a nonlinear trend. Similar nonlinear trends are observed in the other link functions for Examples 1 and 2. Therefore, we should choose a monotonic nonlinear function for h. Three candidates,

$$\log M, \quad (M-\alpha_1)^{-\alpha_2}, \quad \text{and} \quad \exp(-\alpha_1(\mathrm{M}-\alpha_2)^{\alpha_3}), \tag{4}$$

which possess such nonlinear trends are thus considered for h, where α_1, α_2, and α_3 are parameters that will be determined through an iterative estimation procedure. More complicated functions can be considered but should only be used when the simpler ones fail to fit the data.

Model-Selection Procedure

There are three parts of the model that need to be chosen, namely, a link function, a nonlinear function h, and the significant control and noise-factor effects in λ and g. Figure 3 shows the differences of the three link functions. For failure-amplification problems, the lower tail should be the primary concern. Interestingly, the lower tail of cloglog and logit behave almost the same, whereas the probit link has smaller probabilities compared with the other two. These properties indicate that these three links have different scales of model parameters and that the probit link will have more distinct estimates than the other two links. We use the Akaike information criterion (AIC) for model selection because of its ability to compare nested and nonnested models and its popularity in applications. Note that AIC value equals "deviance + 2 size", where the deviance is the $-2\log$ (maximized likelihood) and size is the number of parameters fitted in the model. Because our objective is to identify a parsimonius model that gives small deviance, we should find the model with minimum AIC. Other criteria like BIC or C_p can also be used and the ensuing steps will be similar to that of using AIC.

Choosing/estimating functions f, λ, g, and h simultaneously is cumbersome and is not supported by most commercial statistical software. Therefore, we provide an iterative algorithm that foregoes the burden of programming and utilizes a standard GLM software (for example, the *glm* command in Splus: www.insightful.com and R: www.r-project.org).

Algorithm

1. Choose a function h with parameters $\boldsymbol{\alpha}$ and a link function f. Fit the model $f(p) = h(M, \boldsymbol{\alpha})$, which can be done by minimizing the objective function of AIC with respect to $\boldsymbol{\alpha}$. Denote the estimate by $\hat{\boldsymbol{\alpha}}$.
2. Use a forward selection procedure to select important factors and interactions based on the effect heredity principle to obtain λ and g in $f(p) = \lambda(\mathbf{X}, \mathbf{N}) + g(\mathbf{X}, \mathbf{N})h(M, \hat{\boldsymbol{\alpha}})$.
3. Use the functions λ and g obtained in step 2 to minimize the objective function of AIC with respect to $\boldsymbol{\alpha}$.
4. Repeat steps 2 and 3 until the model functions λ, g, and the parameter estimate $\hat{\boldsymbol{\alpha}}$ converge.
5. Repeat steps 1 through 4 for different choices of h and f. Select h and f with the lowest AIC.

Note that step 2 can be easily carried out using a standard GLM software, thus simplifying the estimation procedure. This step can be further simplified by identifying the factors in g using interaction plots of the control and noise factors against the amplification factor. Then the variable selection needs to be applied only for λ. This simplified strategy will be illustrated using the two PCB examples.

Before getting into the examples, we should provide more details about the variable selection. In the forward selection procedure, we start with the model containing no effects and then the important effects are added to the model one by one (Neter et al. (1996)). This is continued until the AIC starts increasing. Most statistical packages support the implementation of this procedure. However, in general, the model selected will not satisfy the effect heredity principle, i.e., we may select a model containing an interaction without any of its parent effects (Hamada and Wu (1992)). This is undesirable. One simple approach to overcome this problem is to repeat the forward selection procedure by manually removing the interactions for which none of the parent effects are significant.

Example 1. IL PCB Experiment

The data was analyzed by Joseph and Wu (2004) assuming the relationship

$$\log\log\frac{1}{1-p} = \lambda(\mathbf{X}_{(-6)}) + \gamma\log m + \beta\log M, \quad (5)$$

where the exposure energy (x_6) is treated as an adjustment factor denoted by m. Note that this is a special case of (3) with f as the cloglog link, g equal to a constant β, and $h(M) = \log M$. The $\lambda(\mathbf{X}_{(-6)})$ is a second-order linear model (main effects and two-factor interactions) of the control factors, excluding x_6. The two degrees of freedom of three-level factors are split into linear and quadratic components with contrasts $x_l = (-1, 0, 1)$ and $x_q = (1, -2, 1)$. The two-level factor x_1 is coded with $x_{1l} = (-1, 1)$.

Let p_1 and p_2 be the failure probabilities of opens and shorts, respectively. Denote the two amplification factors, line width and spacing, by M_1 and M_2. Note that M_1 is an amplification factor for opens and M_2 is an amplification factor for shorts. The following two models for opens (6) and shorts (7) were selected by forward selection based on AIC:

$$\begin{aligned}\log\log\frac{1}{1-p_1} &= 10.27 - .73x_{5l} + 0.57x_{4l}\\ &\quad - .33x_{2l} - .27x_{1l}x_{5q}\\ &\quad + 2.77\log m + 5.06\log M_1, \quad (6)\\ \log\log\frac{1}{1-p_2} &= -6.66 + .48x_{1l} + .20x_{4l}\\ &\quad - .15x_{1l}x_{5q} + 4.70\log m\\ &\quad + 7.66\log M_2. \quad (7)\end{aligned}$$

Now consider the proposed modeling strategy. Figure 4 shows the interaction plots of the eight control factors with the amplification factor. The par-

TABLE 5. Analysis in Example 1

Data	Link (Eq.#)	AIC	Size	Deviance
Opens	Cloglog-JW (6)	147.91	7	133.91
	Cloglog (9)	121.64	10	101.64
	Logit	122.62	11	100.62
	Probit	137.48	12	113.48
Shorts	Cloglog-JW (7)	101.08	6	89.08
	Cloglog	84.24	11	62.24
	Logit (10)	81.71	10	61.71
	Probit	85.53	12	61.53

allel behaviors of different levels (numbers) of each factor show that there is no indication of strong interaction between control and amplification factors. So the term g in (3) is set to be a constant. The same conclusion is obtained for the shorts data also (plots are not shown). Thus, a simple model is chosen for both shorts and opens:

$$f(p) = \lambda(\mathbf{X}, M) = \lambda(\mathbf{X}) + \beta h(M). \quad (8)$$

For consistency of notation with the rest of the paper, the variable, exposure energy, is denoted by x_6 instead of m.

Table 5 gives the AIC, number of parameters, and deviances of models with different link functions. Cloglog-JW denote the models (6) and (7) suggested by Joseph and Wu (2004). Other models in Table 5 are built by using the proposed modeling strategy. Clearly, models (6) and (7) are not good enough because of the high AIC (and hence high deviance). The proposed strategy provides much better models in all three links. The cloglog link for opens and the logit link for shorts have the lowest AIC. There is not much difference in terms of AIC between cloglog and logit links. Thus, if the same link should be used for both opens and shorts data, one may pick either cloglog or logit. In all the cases, the best selected function h from the three candidates in (4) was $h(M) = \exp(-\alpha_1(M - \alpha_2)^{\alpha_3})$. For numerical stability, we restricted the parameters α_1, α_2, and α_3 to lie in the range $[0, 100]$, $[0, 3]$, and $[-20, 0]$, respectively.

The following model is selected for opens:

$$\begin{aligned}\log\log\frac{1}{1-p_1} = {} & -2.63 - .82x_{5l} - .34x_{6l} + .28x_{8l} \\ & - .27x_{8q} + .15x_{6q} + .24x_{1l} \\ & + 1.06x_{5l}x_{6l} - .38x_{1l}x_{6l} \\ & - 5.27e^{-1.36(M_1-3)^{-1.16}}. \end{aligned} \quad (9)$$

The order of the terms in the model is the order in which they are selected by the regression procedure. The fitted values are plotted in Figure 4, which shows a reasonably good fit to the data. Comparing models (6) and (9), the surface preparation (x_2) is selected in model (6) but not in model (9). Model (9) also selects ORP (x_8) as an important factor with linear and quadratic effects that were absent in model (6).

The following model is selected for shorts:

$$\begin{aligned}\log\frac{p_2}{1-p_2} = {} & -1.86 + .96x_{6l} + .47x_{1l} - .18x_{7l} \\ & + .29x_{6q} + .25x_{4l} - .24x_{1l}x_{7l} \\ & - .45x_{4l}x_{6l} - .31x_{1l}x_{4l} \\ & - 715.18e^{-5.52(M_2-3)^{-.09}}. \end{aligned} \quad (10)$$

Comparing models (7) and (10), developer speed (x_7) is selected in (10) but not in (7). Moreover, the interaction terms are completely different. This indicates that the proposed procedure provides more insights than the procedure in Joseph and Wu (2004).

Example 2. OL PCB Experiment

The plot of opens data in Figure 5 shows that, after the cloglog transformation, only factors x_5 and N appear to have strong nonparallel behavior. Thus, factors x_5 and N interact with the amplification factor. The logit and probit transformation plots, which are not shown here, lead to similar results. Thus, the function g in (3) is chosen to be $\beta_1 x_{5l} + \beta_2 N + \beta_3$. Here the two-level noise factor N is coded with $N = (-1, 1)$ and the other factors are coded as in Example 1. Then the model for opens becomes

$$f(p_1) = \lambda_1(\mathbf{X}, N) + (\beta_1 x_{5l} + \beta_2 N + \beta_3)h_1(M_1).$$

Similar analysis done on shorts data shows that the amplification factor has no interaction with the control factors but has strong interaction with the noise factor. Hence, the function g is chosen to be $\beta_1 N + \beta_2$. Then the model for shorts becomes

$$f(p_2) = \lambda_2(\mathbf{X}, N) + (\beta_1 N + \beta_2)h_2(M_2).$$

The following cloglog and logit models ((11) and (12)) are selected for opens and shorts, respectively, by using the proposed modeling strategy,

$$\begin{aligned}\log\log\frac{1}{1-p_1} = {} & -3.71 + 2.19x_{5l} - .74N \\ & - .48x_{7l} - .32x_{6l} + .55x_{4l} \\ & + .14x_{3l} - .07x_{4q} + .25x_{7l}N \\ & + .26x_{6l}N - .43x_{5l}N + .21x_{3l}N \end{aligned}$$

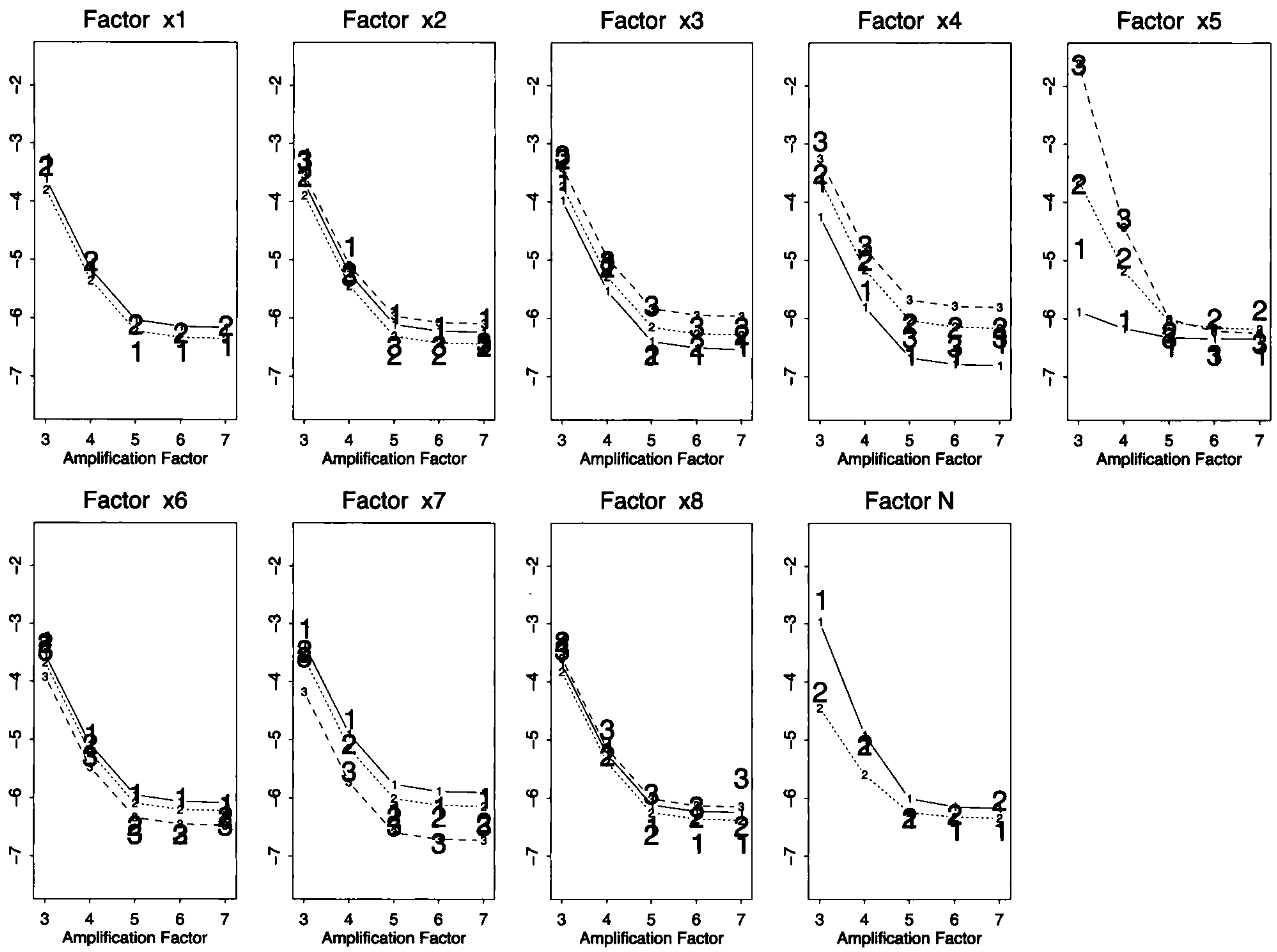

FIGURE 5. Failure Probability with Cloglog Transformation for Opens in Example 2. Average values from the experimental data are plotted along with the fitted lines from (11).

$$
\begin{aligned}
&- .29x_{4l}x_{6l} + .22x_{3l}x_{5l} \\
&+ .15Nx_{4q} + .21x_{5l}x_{4q} + .06x_{7l}x_{4q} \\
&+ (-2.09x_{5l} + .66N - 2.56) \\
&\quad \times e^{-39.54(M_1-1.85)^{-5.69}}, \qquad (11)
\end{aligned}
$$

$$
\begin{aligned}
\log \frac{p_2}{1-p_2} &= -.92 + .54N - .47x_{5l} - .27x_{7l} \\
&- .28x_{4l} + .19x_{3l} + .17x_{3q} \\
&+ .19x_{1l} + .30x_{8l} - .16x_{7l}x_{3q} \\
&+ .19x_{5l}N - .10Nx_{3q} + .09x_{1l}x_{3q} \\
&- .08x_{5l}x_{3q} - 07x_{7l}N \\
&+ (.31N - 3.77)e^{-100.00(M_2-.59)^{-3.58}}. \qquad (12)
\end{aligned}
$$

For comparison, we repeated the analysis using the models recommended in Joseph and Wu (2004). In their models, a cloglog link was used for both opens and shorts with $h(M) = \log M$ and g equal to a constant. The resulting AIC was 486.7 for opens and 272.9 for shorts, both much larger than the AIC of our models (see Table 6). This clearly shows the superiority of the proposed modeling strategy.

TABLE 6. Analysis in Example 2

Data	Link (Eq.#)	AIC	Size	Deviance
Opens	Cloglog (11)	266.33	20	220.33
	Logit	288.28	19	250.28
	Probit	311.18	22	267.18
Shorts	Cloglog	243.34	16	211.34
	Logit (12)	238.29	17	204.29
	Probit	249.61	17	215.61

TABLE 7. Optimum Settings in Example 1

Models (Eq.#)	x1	x2	x3	x4	x5	x6	x7	x8
Cloglog-JW (6) and (7)	−1	1	x	−1	0.342	−0.399	x	x
Cloglog (9) and Logit (10)	−1	x	x	−1	1	−1	−1	−1

Optimum Settings

Optimizing on shorts and opens separately may lead to conflicting levels for the factors. One approach to overcome this problem is to use a loss function to combine the two responses. We propose two loss functions:

$$c_1\lambda_1(\mathbf{X}, N, M_1) + c_2\lambda_2(\mathbf{X}, N, M_2) \qquad (13)$$

and

$$c_1p_1(\mathbf{X}, N, M_1) + c_2p_2(\mathbf{X}, N, M_2) = c_1f_1^{-1}(\lambda_1(\mathbf{X}, N, M_1)) + c_2f_2^{-1}(\lambda_2(\mathbf{X}, N, M_2)). \qquad (14)$$

The first one is a *weighted sum of link functions* and the second one is a *weighted sum of probabilities*. The former is appropriate if the same link function is used for both models, whereas the latter is appropriate if the link functions are different. The weights c_1 and c_2 can be chosen depending on the importance of the two responses. It is easier to determine the weights in (14) because the weights are directly related to the costs of rejection and rework. In the case of ILs (example 1), both shorts and opens are equally bad, and thus we take $c_1 = c_2 = .5$. In the case of OLs (example 2), most of the shorts can be reworked, whereas very few of the opens can be reworked. Therefore, PCBs with opens will be rejected, leading to greater loss. Engineers suggested that the loss due to five shorts can be considered as equivalent to the loss with one open. Therefore, we take $c_1 = 5/6$ and $c_2 = 1/6$.

Optimum settings of the control factors can be obtained by averaging the loss over the distribution of the noise factors and amplification factors. Because the link functions for shorts and opens in (9) and (10) are different, we use the loss function based on probabilities. Table 7 gives the optimum settings for Example 1 by taking expectation over the amplification factors in the range 5 to 7 mil assuming a uniform distribution. Note that, for optimization, we adopt only the levels used in production; amplification is applied only to facilitate estimation. In the table, we have also shown the optimum levels obtained by Joseph and Wu (2004). They are clearly different from the levels obtained from our models. Because our models have better fit to the data than theirs, the optimum levels obtained using our models are expected to be closer to the true optimum.

The optimum settings for Example 2 based on the models in (11) and (12) are shown in Table 8. It is obtained by averaging over the distribution of noise factor and the amplification factors (equal probability is assigned for each combination of the levels) and using the loss function in (14). Again, because the models selected based on the proposed modeling strategy are better than those selected using the strategy in Joseph and Wu (2004), the optimum settings obtained here should be better.

Figures 6 and 7 show the estimated failure probabilities for 5 to 7 mil of width (M_1) and spacing (M_2). The predicted failure probabilities are also given for 8 and 9 mil, which are not included in the experiments. We can see that the failure probabilities at the optimum settings are substantially lower than those at the existing settings. This indicates that the process capability has been greatly improved through the experimentation.

TABLE 8. Optimum Settings in Example 2

Models (Eq.#)	x1	x2	x3	x4	x5	x6	x7	x8
Cloglog (11) and Logit (12)	−1	x	−1	0.325	1	1	1	−1

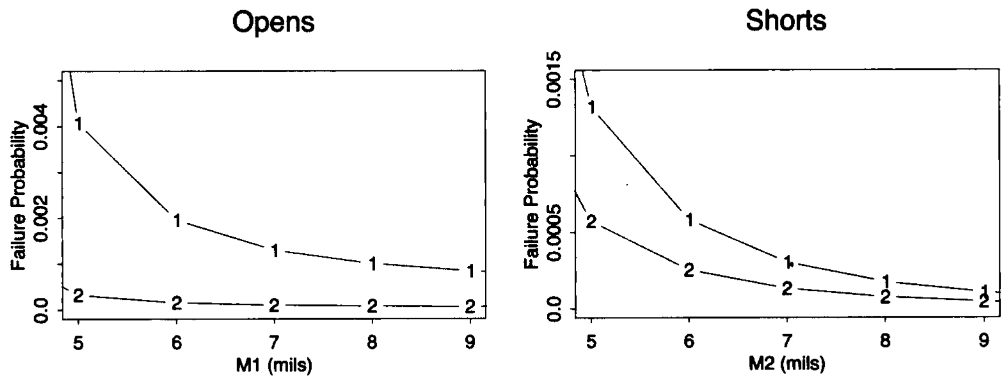

FIGURE 6. Estimated Failure Probabilities in Example 1. 1 = production setting, 2 = optimum setting. Left panel gives cloglog model for opens; right panel gives logit model for shorts.

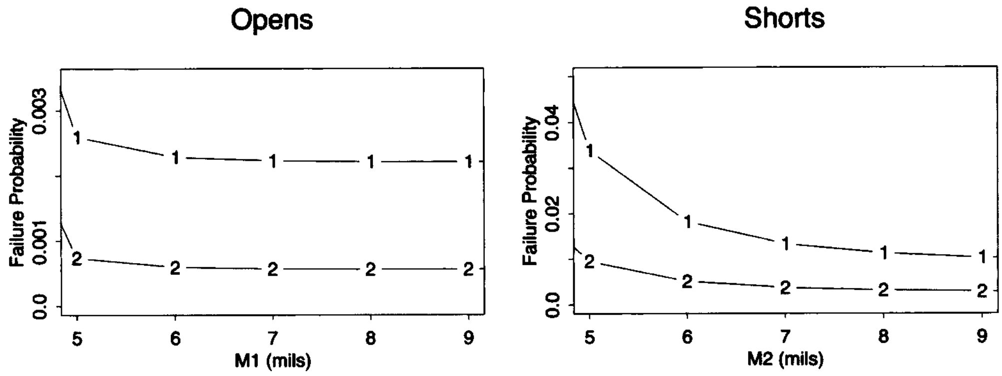

FIGURE 7. Estimated Failure Probabilities in Example 2. 1 = production setting, 2 = optimum setting. Left panel gives cloglog model for opens; right panel gives logit model for shorts.

Conclusions

FAMe points the engineers and statisticians to an avenue for resolving estimation problems in situations with low failure rate. In this article, we propose a general strategy for building models with amplification factors. The proposed strategy is illustrated using two real experiments in PCB manufacturing.

Model selection is done by combining the physical knowledge of the process and the information from the experimental data. The procedure takes advantage of widely available GLM software and tunes it to find a nonlinear function of amplification factor through a model-selection criterion, such as the AIC. This strategy can be implemented easily without a heavy burden of programming.

Two types of loss functions, link and probability, are discussed. Link loss function possesses additive property of factor effects in models from two failures modes. It is easier to measure the total impact of the factors on the failure rate and thus easier to optimize. On the other hand, the probability loss function provides the flexibility of combining models with different link functions.

Accurate failure prediction is essential for proper production planning and control. In PCB manufacturing, the failure prediction is difficult because the design of PCBs (line width, spacing, etc.) varies with customer requirements (see Joseph and Adya (2001)). The models obtained from the experiment, which are functions of line width and spacing, can be used for failure prediction. However, because there are various other failure modes, these models need to be calibrated using actual production data for better prediction.

Acknowledgments

Dr. Jeng's research was partly supported by Taiwan National Science Council grant NSC 43140F. Dr. Joseph's research was supported by U.S. National

Science Foundation grant DMI-0448774. Dr. Wu's research was supported in part by a grant from the U.S. Army Research Laboratory and the U.S. Army Research Office under contract W911NF-05-1-0264.

References

HAMADA, M and NELDER, J. A. (1997). "Generalized Linear Models for Quality-Improvement Experiments". *Journal of Quality Technology* 29, pp. 292–304.

HAMADA, M and WU, C. F. J. (1992). "Analysis of Designed Experiment with Complex Aliasing". *Journal of Quality Technology* 24, pp. 103–137.

JOSEPH, V. R. and ADYA, H. (2001). "Prediction of Yield in a Multi-Product Batch Production Environment". *Quality Engineering* 14, pp. 153–159.

JOSEPH, V. R. and WU, C. F. J. (2004). "Failure Amplification Method: An Information Maximization Approach to Categorical Response Optimization (with Discussion)". *Technometrics* 46, pp. 1–31.

MARUTHI, B. N., and JOSEPH, V. R. (1999). "Improving the Yield of Printed Circuit Boards Using Design of Experiments". *Quality Engineering* 12, pp. 259–265.

MARUTHI, B. N.; SOMASHEKAR, H. S.; SHARMA, R. K.; KUMAR, K.; and JOSEPH, V. R. (1998). "Improving the Process of Outer Layer Image Transfer in PCB Manufacturing". *Conference Proceedings of the 8th National Convention of NIQR*, Bangalore, India, pp. 1A.1–1A.8.

MCCULLAGH, P. and NELDER, J. A. (1989). *Generalized Linear Models,* 2nd ed. Chapman & Hall, New York, NY.

NETER, J.; KUTNER, M. H.; NACHTSHEIM, C. J.; and WASSERMAN, W. (1996). *Applied Linear Statistical Models.* McGraw-Hill, Boston.

TAGUCHI, G. (1986). *Introduction to Quality Engineering: Designing Quality into Products and Processes.* Asian Productivity Organization, UNIPUB/Kraus International, White Plains, NY.

WU, C. F. J. and HAMADA, M. (2000). *Experiments: Planning, Analysis and Parameter Design Optimization.* John Wiley & Sons, New York, NY.